Package: agTrend.ssl (via r-universe)

September 18, 2024

Title Aggregated Abundance Trends using GAMs

Version 0.0.0.9005

Date 12-20-2022

Description The package supplies a set of functions for stochastically imputing missing site-level abundance measurements so sites can be aggregated to regional abundance for trend estimation. This package is specifically intended for trend estimation for the NMFS Alaska Fisheries Science Center's assessment of Steller sea lion trends in Alaska and the U.S. west coast. Although the methodology is applicable to other applications, there are aspects of sea lion surveys hard coded into the analysis.

Depends mgcv

Imports coda, dplyr, mvnfast, purrr, readxl, rlang, stringi, tidyr

License CC0

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.1

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

RemoteUrl https://github.com/dsjohnson/agTrend.ssl

RemoteRef HEAD

RemoteSha 6bb4ecfa713f7c9ed0410ea846e94fb511e66a7e

Contents

ag.abund 2
ag.summary 2
ag.trend 3
fit.gam 3
proc.data 4
sample.abund 5

Index **6**

ag.abund *Aggregate abundance samples*

Description

Takes abundance samples created with the function `sample.abund` and aggregates them based on a factor variable in (or added) to the `sample.abund` data set.

Usage

```
ag.abund(x, ag.var)
```

Arguments

`x`	A sample abundance data frame produced by `sample.abund`
`ag.var`	A factor (or character) variable in `x` that indicates the group membership for aggregation.

Author(s)

Devin S. Johnson

ag.summary *Summarize count posterior distributions*

Description

The posterior distribution of the count sample is summarized for each site. The resulting data frame can be used for site-level plots, etc.

Usage

```
ag.summary(x, ci.prob = 0.9)
```

Arguments

`x`	A data frame created with the sampling function `sample.abund` or `ag.abund`
`ci.prob`	Probability level for credible intervals. Defaults to 0.9.

Author(s)

Devin S. Johnson

ag.trend *Estimate trends of aggregated abundance*

Description

The growth trend of is estimated from the samples obtained from the functions `sample.abund` and `ag.abund`.

Usage

```
ag.trend(x, timeframe, ci.prob = 0.9)
```

Arguments

`x`	A abundance sample data frame from `sample.abund` or `ag.abund`
`timeframe`	a 2-vector giving the start year and end year of the desired trend estimate, e.g., `timeframe = c(1989,2019)`.
`ci.prob`	The probability for the credible interval. Defaults to 0.9.

Details

The function returns a named list with elements `growth`, `fitted`, and `sample`. The `growth` element contains a table with the estimated growth of each aggregation in *percent growth* form. The `fitted` element contains a table with the fitted trendline values on the count scale. This can be used for plotting. Finally, the `sample` element contains the posterior sample. This can be used for further analysis of trends such as comparisons, etc.

Author(s)

Devin S. Johnson

fit.gam *Fit penalized GAM to each site for missing data imputation*

Description

This function takes a processed data frame with counts during survey years for each site and fits a multi-site GAM using a Tweedie response distribution.

Usage

```
fit.gam(data, obl.corr, alt.mod = NULL, warn = FALSE, debug = FALSE, ...)
```

Arguments

data	A data frame containing the sites, counts, and years.
obl.corr	Logical. Should oblique photo correction be used?
alt.mod	An alternative to the defaul model. See Details.
warn	Logical. Should fitting warnings be printed. Defaults to FALSE.
debug	Logical. If set to TRUE the function will drop into browser mode within the function upon execution. Defaults to FALSE.
...	Extra arguments passed to gam

Details

A penalized GAM model is fit to the data with a Tweedie distribution for the response distribution. The mean function for site i in year t is modeled with a smooth year term for all sites plus a factor smooth such that each site has its own smooth year function as well the mgcv formula is `mu.form = count ~ s(year, SITE, bs='fs', k=8, m=1)`. For the power and dispersion formulas, `p.form = ~ s(SITE, bs='re')` and `phi.form = ~ s(SITE, bs='re')`, so that each site has its own p and phi parameters for the Tweedie specification. Thus, for each site, the mean count is mu(i,t) and the variance is V(i,t) = phi(i)*mu(i,t)^p(i). To specify a different formula for the model, set `alt.model = list(mu=mu.form.alt, p=p.form.alt, phi=phi.form.alt)` where the `*.alt` signifies the formula for the desired alternative.

If `obl.corr = TRUE` then a model will be fitted that corrects for the approximately 3.8% reduction in the expected count if the data were collected using an oblique photo vs. a medium format vertical photo. Unlike the original agTrend package the uncertainty in this estimate is not accounted for in the model. A few initial tests revealed that is source of variation seems insignificant when compared to the natural variation of the observed counts and the model was significantly easier to fit and more robust when fixing this quantity.

Author(s)

Devin S. Johnson

---

proc.data *Process SSL count data .xlsx file*

Description

The current 'ALLCOUNTS_v(x).xlsx' file is processed into the separate, R friendly, data sets necessary for trend analysis.

Usage

```
proc.data(allcounts, age, dps)
```

Arguments

allcounts	the file path of the current allcounts excel file.
age	Age needed, must be one of 'pup' or 'nonpup'
dps	DPS needed, must be one of 'wdps' or 'edps'

Author(s)

Devin S. Johnson

---

sample.abund *Sample missing counts*

Description

Using a fitted GAM SSL model (see fit.gam) and the data used for fitting, missing survey counts are imputed for aggregated trend and abundance analysis.

Usage

```
sample.abund(
  fit,
  data,
  yrs,
  size = 1000,
  add.site.data = NULL,
  keep.par = FALSE,
  debug = FALSE
)
```

Arguments

fit	A fitted model object from a call to fit.gam.
data	Data used for model fitting.
yrs	Years used for simulation. Defaults to a sequence of years from the first to the last. But this can be any subset of years in the survey window.
size	The sample size of the random imputation draws. Defaults to 1,000.
add.site.data	An additional data set with a column lableled 'site' matching the sites used in the model fitting data that contain additional groupings of other site-level data.
keep.par	Logical. Should the sample of parameters used be retained? It is not often needed for further analysis.
debug	Enter into the function for debugging.

Author(s)

Devin S. Johnson

Index

ag.abund, *2*, 2, *3*
ag.summary, 2
ag.trend, 3

fit.gam, 3, *5*

gam, *4*

proc.data, 4

sample.abund, *2, 3*, 5