Package ‘mvtweedie’

January 7, 2026

Type Package

Title Estimate Diet Proportions Using Multivariate Tweedie Model

Version 1.2.0

Date 2025-12-19

Description Defines predict function that transforms output from a Tweedie
Generalized Linear Mixed Model (using 'glmmTMB'),
Generalized Additive Model (using 'mgcv'), or
spatio-temporal Generalized Linear Mixed Model (using package 'tinyVAST'),
and returns predicted proportions (and standard errors) across a
grouping variable from an equivalent multivariate-logit Tweedie model.
These predicted proportions can then be used for standard plotting
and diagnostics. See Thorson et al. 2022 <doi:10.1002/ecy.3637>.

Imports stats, tibble

Suggests mgcv, knitr, rmarkdown, ggplot2, glmmTMB, lattice, pdp,
raster, sp, RANN, plotrix, tweedie, abind, rnaturalearth,
rnaturalearthdata, sf, dplyr, viridisLite, tinyVAST

Depends R (>= 4.1.0)

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.3

VignetteBuilder knitr

LazyData true

URL `https://james-thorson-noaa.github.io/mvtweedie/`

NeedsCompilation no

Author James Thorson [aut, cre] (ORCID:
<`https://orcid.org/0000-0001-7415-1010`>)

Maintainer James Thorson <`James.Thorson@noaa.gov`>

Repository CRAN

Date/Publication 2026-01-07 08:50:02 UTC

Contents

`Middleton_Island_TUPU` *Middleton Island tufted puffin diets*

Description

Data to demonstrate multivariate Tweedie GAM for time-series

Usage

```
data(Middleton_Island_TUPU)
```

Details

Data sufficient to demonstrate how to use a Tweedie Generalized Additive Model to provide inference about proportions e.g. in food habits analysis, where the model output is processed to represent a multivariate logit Tweedie model.

Specifically includes Tufted Puffin (Fratercula cirrhata) bill loads sampled at Middleton Island.

- `Response` Numeric prey biomass in bill load samples
- `Year` Numeric year
- `group` factor representing prey species or category
- `SampleID` factor with a level for every sampling occasion, e.g., for use in row normalization

Value

A data long-form data frame

Author(s)

Mayumi Arimitsu

References

Hatch, S., and G. Sanger. 1992. Puffins as Samplers of Juvenile Pollock and Other Forage Fish in the Gulf of Alaska. Marine Ecology Progress Series 80: 1-14.

`mvtweedie` *Multivariate Tweedie distribution for predicting diet proportions*

Description

Using regression methods to analyze diet proportions for a marked point process

Details

Diet samples often measure a count or biomass for different prey categories. Rather than converting these data to a proportion and fitting these proportions as data, we can instead represent diet samples as an outcome from a thinnned and double-marked point process, where marks include prey category and size per encounter, and thinning represents variation in attack and capture rates and is conceptually similar to detectability/catchability in other point-count sampling analyses. Analyzing raw prey measurements (rather than proportions) allows a wide range of models (and associated off-the-shelf software), predictions can still be converted to proportions (with associated standard errors) after the model is fitted, and we can represent covariance in prey measurements within a sample using covariates that explain sample-specific attack/capture rates.

If the prey densities follow a a Poisson point-process, and prey size per encounter follows a gamma distribution, then the resulting distribution for biomass of each prey follows a multivariate Tweedie distribution. We therefore interpret the multivariate Tweedie distribution as a "process-based" model for prey samples.

References

Thorson, J. T., Arimitsu, M. L., Levi, T., & Roffler, G. H. (2022). Diet analysis using generalized linear models derived from foraging processes using R package mvtweedie. Ecology, 103(5), e3637. doi:10.1002/ecy.3637

See Also

`predict.mvtweedie` for details

`predict.mvtweedie` *Predict proportions for new data*

Description

Predict proportions and associated standard errors using a standard S3 object interface

Usage

```
## S3 method for class 'mvtweedie'
predict(object, category_name = "group", newdata, se.fit = FALSE, ...)
```

Arguments

`object`	output from `gam` or `glmmTMB`, but with `class(object)=c("mvtweedie",...)` where `...` indicates the original values for `class(object)`
`category_name`	name of column that indicates grouping variable
`newdata`	An optional data frame in which to look for variables with which to predict. If omitted, the fitted values are used.
`se.fit`	Whether to approximate the standard errors for predicted proportions
`...`	Not used

Details

After fitting Tweedie GLM using a log-link and multiple categories, we transform predicted densities to yield predicted proportions and associated SEs. This estimator for proportions arises naturally when analyzing a double-marked point process for diet samples, with marks for category and size.

`predict.mvtweedie` does this transformation for a model fitted using:

- A generalized additive model (GAM) using `gam`
- A generalized linear mixed model (GLMM) using `glmmTMB`
- A spatio-temporal generalized linear mixed model (GLMM) using `tinyVAST`

It then also calculates an approximation to the standard error for this proportion. Specifically, we calculate the proportion for each category as the density X for that category, and the sum of densities Y for all other categories:

$$p_X = \frac{X}{X+Y}$$

Assuming we have an estimator for the standard error $s(X)$ and $s(Y)$, and assuming that those estimators are independent such that $s(X+Y)^2 = s(X)^2 + s(Y)^2$, we then apply the delta method to approximate the standard error for the proportion as:

$$s(p_X)^2 = \frac{X^2}{(X+Y)^2}\left(\frac{s(X)^2}{X^2} - 2\frac{s(X)^2}{X(X+Y)} + \frac{s(X)^2+s(Y)^2}{(X+Y)^2}\right)$$

Predictions X and Y, and standard errors $s(X)$ and $s(Y)$ are then supplied by the `predict` function that is native to the software used when fitting the model.

Value

predict.mvtweedie produces a vector of predicted proportions or a list containing predicted proportions and standard errors.

Examples

```
# Load packages
library(mvtweedie)
library(mgcv)

# load data set
data( Middleton_Island_TUPU, package="mvtweedie" )

# Run Tweedie GLM
gam0 = gam(
  formula = Response ~ 0 + group,
  data = Middleton_Island_TUPU,
  family = tw
)

# Inspect results
class(gam0) = c( "mvtweedie", class(gam0) )
predict(
  gam0,
  se.fit = TRUE
)
```

`southeast_alaska_wolf` *Wolf environmental DNA in southeast Alaska*

Description

Data to demonstrate multivariate Tweedie GAM for spatial analysis

Usage

`data(southeast_alaska_wolf)`

Details

Data sufficient to demonstrate how to use a Tweedie GLM to provide inference about proportions e.g. in food habits analysis, where the model output is processed to represent a multivariate logit Tweedie model.

Specifically includes environmental DNA sampling of wolf scats obtained in Southeast Alaska.

- `Latitude` Latitude for scat sample
- `Longitude` Longitude for scat sample
- `group` prey groupings from eDNA smaples
- `Response` relative read abundance calculated as the mean proportion of DNA sequence reads from each species among samples from a given scat

Value

A data long-form data frame

Author(s)

Gretchen Roffler

References

Roffler, G. H., J. M. Allen, A. Massey, and T. Levi. 2021. Wolf Dietary Diversity in an Island Archipelago. Bulletin of the Ecological Society of America 102: 1-6. doi:10.1002/bes2.1834

Index