|  |  |
| --- | --- |
| TrendNPS\_Cont {TrendNPS} | R Documentation |

**Trend for Continuous Outcomes from Complex Survey Designs**

TrendNPS\_Cont is used to fit trend models of continuous outcomes for four approaches: the unreplicated linear mixed model as specified by Piepho and Ogutu (2002) that does not incorporate design weights (PO), simple linear regression of annual design-based estimates (SLRDB), weighted linear regression of annual design-based estimates (WLRDB), and probability-weighted iterative generalized least squares of a linear mixed model (PWIGLS) with 6 types of scaling. Fixed effects structure includes a term for the year for trend estimation and an optional two-level stratification factor. Random intercepts for the site and year are assumed, and an optional random site-level slope effect may be included.

**Usage**

TrendNPS\_Cont(dat,method,RE,type=NA,stratum=NA,Y=NA,lat=NA,long=NA,

stage1wt=NA,stage2wt=NA,str1prop=NA,nbhd=TRUE)

**Arguments**

|  |  |
| --- | --- |
| dat | data set for trend analysis containing columns for “Site”, “WYear” (integer-valued variable designating survey occasions used in the fixed-effects portion of the trend model),“Year” (factor variable of survey occasions used in the random-effects portion of the trend model), the continuous outcome of interest, design weights, panel weights, and other design features if needed. |
| method | trend estimation method taking the values 'PO' for the Piepho and Ogutu (2002) unreplicated linear mixed model, 'SLRDB' for simple linear regression of design-based estimates, 'WLRDB' for weighted linear regression of design-based estimates, or 'PWIGLS' = probability-weighted iterative generalized least squares (Pfeffermann et al. 1998, Asparouhov 2002). |
| slope | logical value indicating whether or not the random site-level slope effect is included in the variance components structure used for the PO and PWIGLS trend methods. Site-level and year-level random intercept terms are included as default. |
| type | scaling type used when method='PWIGLS'. Takes the values 'Aonly' (probability weighting but no scaling at either stage), 'A' (scaling for panel weights with the mean site-level design weight), 'AI' (scaling for panel weights with the mean site-level design weight, but no site-level scaling), 'B' (scaling for panel weights with the effective mean site-level design weight), 'BI' (scaling for panel weights with the effective mean site-level design weight, but no site-level scaling), or 'C' (scaling only at the year level with the inverse of the average year-level weight). |
| stratum | optional column name of two-level stratification factor. |
| Y | character string indicating the name of the column of the outcome of interest. |
| lat | latitude of each site using an equal-area projection such as UTM or Albers. |
| long | longitude of each site using an equal-area projection such as UTM or Albers. |
| stage1wt | the design weight from the original sample draw without accounting for temporal revisit designs. |
| stage2wt | the panel inclusion weight for each site each year given inclusion in the sample. |
| str1prop | the proportion of the first stratum in the population. |
| nbhd | logical indictor for using the neighborhood variance estimator when the 'WLRDB' trend method is used. If FALSE, independent random sampling is assumed to calculate the standard error for the design-based mean estimate of each survey occasion which is used for weighted linear regression. |
|  |  |

**Details**

The outcome of interest for this trend analysis must be a continuous and positive real-valued (e.g. value of 3.26) outcome. For long-term monitoring, we assume that a sampling occasion occurs no more than once per year at any given site. Furthermore, we assume that a logarithmic transformation is applied to the outcome of interest for the PO and PWIGLS trend methods, so the outcome of interest cannot take values of 0. The data set dat requires two temporal fields. “WYear” is an integer-valued variable indicating ordered survey occasions. The regression coefficient associated with this model predictor provides the basis for trend estimation and testing. The “WYear” is shifted so that the variable takes a value of 0 for the year in which the outcome of interest is least variable. See Piepho & Ogutu (2002) for more information. The “Year” variable is a factor with a unique value for each survey occasion and provides the basis for the estimation of random year-to-year variation among survey occasions. Latitude and longitude are required when the 'SLRDB' and 'WLRDB' methods are used and the neighborhood variance estimator is applied for spatially-balanced samples. If spatially-balanced sampling was not used in the original sampling design, the option nbhd option should be assigned the value of FALSE so that independent random sampling is assumed for standard errors used for weighting in the 'WLRDB' method. Stage 1 (sampling design) and Stage 2 (temporal revisit design) weights must be adjusted for any missing data or frame error prior to trend analysis. For more information on weighting adjustments for nonresponse and frame error, see Oh and Scheuren (1983), Little and Rubin (2002), and Starcevich et al. (2016). For background on the unweighted linear mixed model, see Piepho & Ogutu (2002). For information on the probability-weighted generalized least squares approach, see Pfeffermann et al. (1998), Asparouhov (2006), and Starcevich et al. (2017).

**Value**

TrendNPS\_Cont returns a vector containing the trend estimate, its standard error, the estimated variance components, and the degrees of freedom for the trend test or confidence interval construction.

An object of class "data.frame" is a list containing the following components:

|  |  |
| --- | --- |
| mu | estimated intercept from the trend model |
| trend | estimated trend of the logged outcome from the trend model. |
| SEtrend | the standard error of the trend estimate for the logged outcome. |
| sig2a | the estimated site-to-site variation. |
| sig2b | the estimated year-to-year variation. |
| sig2t | the estimated variation among site-level slopes. |
| sig2e | the estimated residual variation. |
| eta | degrees of freedom for the trend test or confidence interval construction. |

trend is the linear trend of the logged outcome and can be interpreted as multiplicative change on the original scale of the outcome of interest by calculating exp(trend).

**Fitting functions**

The unweighted linear mixed model for the PO method is obtained with the lme4 package with the following function call:

fit<-lmer(LogY ~ WYear + (1|Year) +(1|Site), data=dat).

If a two-level stratification factor is used, the PO method fits the following call:

fit<-lmer(LogY ~ WYear + Stratum + (1|Year) +(1|Site), data=dat).

For the PWIGLS method, a similar model is fit but probability-weighting is applied as defined by the type input.

**Author(s)**

Leigh Ann Starcevich and Jason Mitchell of Western EcoSystems Technology, Inc.

**References**

Asparouhov, T. (2006). General multi-level modeling with sampling weights. *Communications in Statistics – Theory and Methods* 35: 439-460.

Little, J. A. R., and D. B. Rubin. 2002. Statistical analysis with missing data*,* 2nd edition. John Wiley and Sons, Inc., New Jersey.

Oh, H. L., and F. J. Scheuren. 1983. Weighting adjustment for unit nonresponse.Pages 143-184 *in* W. G. Madow, I. Olkin, and D. B. Rubin, editors.Incomplete data and in sample surveys. (Vol. 2). Academic Press, New York.

Pfeffermann, D., C.J. Skinner, D.J. Holmes, H. Goldstein, and J. Rasbash (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society, Series B* 60(1): 23-40.

Piepho, H.P., & Ogutu, J.O. (2002). A simple mixed model for trend analysis in wildlife populations. *Journal of Agricultural, Biological, and Environmental Statistics*, 7(3), 350-360.

L. A. Starcevich, G. DiDonato, T. McDonald, and J. Mitchell. 2016. A GRTS user’s manual for the SDrawNPS package: A graphical user interface for generalized random tessellation stratified (GRTS) sampling and estimation. Natural Resource Report NPS/PWRO/NRR—2016/1233. National Park Service, Fort Collins, Colorado.

L.A. Starcevich, T. McDonald, A. Chung-MacCoubrey, A. Heard, and J. Nesmith (2017). Trend Estimation for Complex Survey Designs. Natural Resource Report NPS/xxxx/NRR—2017/xxxx. National Park Service, Fort Collins, Colorado.

**See Also**

lmer, spsurvey.

**Examples**

## Read example data set

SEKIANC\_orig = read.csv("SEKIANC\_orig.csv",header=TRUE)

## Load dependent packages

require(lme4)

require(lmerTest)

require(spsurvey)

# PO method – full random effects model

PO\_ests = TrendNPS\_Cont(dat=SEKIANC, method="PO", type=NA, slope=TRUE, stratum=NA, Y="Y", lat="Lat", long="Long",stage1wt="wgt", stage2wt="Panelwgt", str1prop=NA)

# Error: number of observations (=80) <= number of random effects (=92) for

# term (1 + WYear | Site); the random-effects parameters and the residual

# variance (or scale parameter) are probably unidentifiable

# The previous error message indicating that the random effects are not

# estimable due to too few observations. Reduce the random effects

# model by removing the random slope term.

# PO method – reduced random effects model

PO\_ests = TrendNPS\_Cont(dat=SEKIANC, method="PO", type=NA, slope=FALSE, stratum=NA, Y="Y", lat="Lat", long="Long",stage1wt="wgt", stage2wt="Panelwgt", str1prop=NA)

round(PO\_ests,4)

# mu trend SEtrend sig2a sig2t sigat sig2b sig2e eta

# 3.964 0.0194 0.0322 0.3422 0 0 0 0.0161 3.6202

# Construct a 95%-confidence interval on the trend estimate of the logged

# outcome

CI = c(PO\_ests$trend - qt(1-(0.05/2), PO\_ests$eta) \* PO\_ests$SEtrend,

PO\_ests$trend + qt(1-(0.05/2), PO\_ests$eta) \* PO\_ests$SEtrend)

round(c(PO\_ests$trend, CI),4)

# 0.0194 -0.0737 0.1125

# Construct a 95%-confidence interval on the trend estimate of the original

# outcome

round(exp(c(PO\_ests$trend, CI)),4)

# 1.0196 0.9290 1.1191

# Conduct a two-sided test of trend for an alternative hypothesis of trend in

# either direction

t\_stat = PO\_ests$trend/ PO\_ests$SEtrend

p\_value = 2\*(1-pt(abs(t\_stat), PO\_ests$eta, lower=TRUE))

round(data.frame(t\_stat,p\_value) ,4)

# 0.6038 0.5817

###########################

# WLRDB method

WLRDB\_ests = TrendNPS\_Cont(dat=SEKIANC, method="WLRDB", type=NA, slope=FALSE, stratum=NA, Y="Y", lat="Lat", long="Long",stage1wt="wgt", stage2wt="Panelwgt", str1prop=NA)

round(WLRDB\_ests,4)

# mu trend SEtrend sig2a sig2t sigat sig2b sig2e eta

# 4.1914 -0.0177 0.0492 0 0 0 0 0.0002 4

###########################

# PWIGLS Aonly

PWIGLS\_Aonly\_ests = TrendNPS\_Cont(dat=SEKIANC, method="PWIGLS", type="Aonly", slope=FALSE, stratum=NA, Y="Y", lat="Lat", long="Long",stage1wt="wgt", stage2wt="Panelwgt", str1prop=NA)

round(PWIGLS\_Aonly\_ests,4)

# mu trend SEtrend sig2a sig2t sigat sig2b sig2e eta

# 3.8812 0.0383 0.031 6.2650 0 0 0.0732 0.0167 3.6202

[Package *TrendNPS* version 1.0.0]