Whittaker-Robinson periodogram

Description

Whittaker-Robinson periodogram for a univariate series of quantitative data.

Usage

```
WRperiodogram(x, T1=2, T2=NULL, nperm=499, mult="bonferroni")
## Graph plotting function
plot.WRperio(x, prog=1, alpha=0.05, line.col="red")
```

Arguments

`x`	A vector of quantitative values for function `WRperiodogram`, or an output object of `WRperiodogram` for function `plot.WRperio`.
`T1`	First period included in the calculation (default: T1=2).
`T2`	Last period included in the calculation (default: T2=n/2).
`nperm`	Number of permutations for the tests of significance.
`mult`	Correction method for multiple testing. Choices are "sidak" or "bonferroni".
`prog`	`prog`=1 (default): use the original p-values in the plot. `prog`=2: use the p-values corrected for multiple testing. `prog`=3: progressive correction of multiple-testing.
`alpha`	Significance level for the plot; p-values smaller than or equal to alpha are represented by black symbols. Default: alpha=0.05.
`line.col`	Colour of the lines between symbols in the graph.

Details

The Whittaker-Robinson periodogram (Whittaker and Robinson, 1924) identifies periodic components in a vector of quantitative data. The vector may contain missing observations represented by NA. The periodogram statistic used in this function is the standard deviation of the means of the columns of the Buys-Ballot table (Enright, 1965). The method is also described in Legendre & Legendre (2012, Section 12.4.1).

The data must be stationary before computation of the periodogram. Stationarity is violated when there is a trend in the data or when they were obtained under contrasting environmental or experimental conditions. Users should at least test for the presence of significant linear trend in the data (using linear regression) and, if a significant trend is identified, it can be removed by computing regression residuals.

The optional graph produced by the function shows the periodogram statistics and their significance following a permutation test. The p-values may be corrected for multiple testing using either the Sidak or the Bonferroni correction, which can be applied to all values in the correlogram uniformly, or following a progressive correction.

A progressive correction means that for the first periodogram statistic, the p-value is tested against the alpha significance level without any correction; for the second statistic, the p-value is corrected for 2 simultaneous tests; and so forth until the k-th statistic, where the p-value is corrected for k simultaneous tests. This approach solves the problem of "where to stop computing a periodogram"; one goes on as long as significant values are likely to emerge, considering the fact that the tests become progressively more conservative.

In the Whittaker-Robinson periodogram, harmonics of a basic period are often found to be also significant.

Value

The function produces a table with the following columns:

`Period`	period number,
`WR.stat`	periodogram statistic,
`p-value`	p-value after permutation test,
`p.corrected`	p-value corrected for multiple testing (using the Sidak or Bonferroni method),
`p.corr.prog`	p-value after progressive correction.

When the p-values cannot be computed because of a very high proportion of missing values in the data, values of 99 are posted in the last three columns of the output table.

Author

Pierre Legendre, Université de Montréal.

References

Enright, J. T. 1965. The search for rhythmicity in biological time-series. Journal of Theoretical Biology 8: 426-468.

Legendre, P. and Legendre, L. 2012. *Numerical Ecolog*y. 3rd English ed. Elsevier, Amsterdam.

Whittaker, E. T. and G. Robinson. 1924. The calculus of observations – A treatise on numerical mathematics. Blackie & Son, London.

Examples

```
### 1. Numerical example of Subsection 12.4.1 of Legendre and Legendre (2012)

test.vec <- c(2,2,4,7,10,5,2,5,8,4,1,2,5,9,6,3)

# Periodogram with permutation tests of significance
res <- WRperiodogram(test.vec)
plot.WRperio(res)
```

```
### 2. Simulated data

periodic.component <- function(x,T,c) cos((2*pi/T)*(x+c))

n <- 500   # corresponding to 75 days, 4 observations per day
# Generate a lunar cycle, 29.5 days (T=118)
moon <- periodic.component(1:n, 118, 59)
# Generate a circadian cycle (T=4)
daily <- periodic.component(1:n, 4, 0)
# Generate a tidal cycle (T=2)
tide <- periodic.component(1:n, 2, 0)

# Periodogram of the lunar component only
res.moon <- WRperiodogram(moon, nperm=0)
res.moon <- WRperiodogram(moon, T2=130, nperm=99)
par(mfrow=c(1,2))
plot(moon)
plot.WRperio(res.moon, prog=1)

# Add the three components, plus a random normal error term
var <- 5*moon + daily + tide + rnorm(n, 0, 0.5)
# Draw a graph of a portion of the data series
par(mfrow=c(1,2))
plot(var[1:150], pch=".", cex=1)
lines(var[1:150])

# Periodogram of 'var'
res.var <- WRperiodogram(var, T2=130, nperm=99)
plot.WRperio(res.var, prog=1, line.col="blue")
# Find position of the maximum value of this periodogram
which(res.var[,2] == max(res.var[,2]))

# Replace 10% of the 500 data by NA
select <- sort(sample(1:500)[1:50])
var.na <- var
var.na[select] <- NA
res.var.na <- WRperiodogram(var.na, T2=130, nperm=99)
plot.WRperio(res.var.na, prog=1)

### 3. Test data used in the examples of the documentation file of function acf() of {stats}
# Data file "ldeaths"; time series, 6 years x 12 months of deaths in UK hospitals
ld.res.perio <- WRperiodogram(ldeaths, nperm=499)
par(mfrow=c(1,2))
plot.WRperio(ld.res.perio, prog=1) # Graph with no correction for multiple testing
plot.WRperio(ld.res.perio, prog=3) # Graph with progressive correction
acf(ldeaths)   # acf() results, for comparison
```