Package: Linda (via r-universe)

July 5, 2026

Type Package

Title Line Transect-Based Nearest Neighbor Distance Analysis

Version 1.3

Date 2026-06-28

Description Conducting Line transect-based one-dimensional nearest neighbor distance and conducting hypothesis testing related to local distributional aggregation pattern of species. The reason that such a package is needed is because traditional two-dimensional nearest neighbor distance is not applicable when biodiversity data are sampled via optimal ecological survey methods, like line transects. In comparison to the entire studied region, line transect-collected local biodiversity data are spatially constrained and sampling-limited. To this end, two-dimensional nearest neighbor distance would tend to over-estimate distributional aggregation pattern of species when using this limited biodiversity information. Accordingly, one-dimensional nearest neighbor distance is needed and the associated statistical testing should be established for analyzing line transect-derived biodiversity data.

Maintainer Youhua Chen <`chenyh@cib.ac.cn`>

License GPL-3

Imports doParallel, foreach,iterators,parallel,stats

Encoding UTF-8

NeedsCompilation no

Author Youhua Chen [aut, cre], Tsung-Jen Shen [aut], Xiaoqin Shi [ctb]

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

Date/Publication 2026-07-04 07:20:02 UTC

RemoteUrl https://github.com/cran/Linda

RemoteRef HEAD

RemoteSha abf69cbd12f634485a069135d87ba0f3a05796aa

Contents

`Linda-package` *Line Transect-Based Nearest Neighbor Distance Analysis*

Description

Linda-an R package for conducting (Lin)e transect-based nearest neighbor (d)istance (a)nalysis

Details

Linda is an R package for calculating one-dimensional nearest neighbor distance (NND) and conducting hypothesis testing related to local distributional aggregation pattern of species. The reason that such a package is needed is because traditional two-dimensional NND is not applicable when biodiversity data are sampled via optimal ecological survey methods, like line transects. In comparison to the entire studied region, line transect-collected local biodiversity data are spatially constrained and sampling-limited. To this end, two-dimensional NND would tend to over-estimate distributional aggregation pattern of species when using this limited biodiversity information. Accordingly, one-dimensional NND is needed and the associated statistical testing should be established for analyzing line transect-derived biodiversity data. Our package, Linda, is developed to fulfill this biodiversity-inference task.

Author(s)

Youhua Chen (Chengdu Institute of Biology, Chinese Academy of Sciences);
Maintainer:
Youhua Chen <chenyh@cib.ac.cn>

References

Xiaoqin Shi, Yongbin Wu, Qi Xiao, Youhua Chen (2026) Linda: an R package using Line transect-based nearest neighbor distance analysis to infer distributional aggregation pattern of species. Plant Diversity.

Examples

```
x=cbind(1,runif(100))
x=rbind(x,cbind(2,runif(100)))
x=rbind(x,cbind(3,runif(100)))
lxy=cbind(x,1)
lxy[,2]=sort(lxy[,2]) #sequentially sampled in an economic way
LNND(lxy)
#in empirical data, because we assume individuals have been recorded in sequential order
#so the original data do not need to be sorted.
#By contrast, in simulated data, if we assumed individuals are recorded in sequential order
#and in an economic way
#sort() function should be used.
```

`CECI` *CECI is a function to compute the Clark and Evans' competition index*

Description

CECI is a function to compute the Clark and Evans' competition index, it computes the first nearest neighbor distance for all the spatial points using extensive two-dimensional circular area searching method.

Usage

```
CECI(xy, area = NULL, method = "NP")
```

Arguments

`xy`	xy is a two-column matrix, containing x and y coordinates of recorded organisms in a single line transect
`area`	area is the size of given study area, the default is NULL and will be estimated from xy data.
`method`	method is the method to compute the index, "NP" is the default, indicating the loop-way computation; "P" indicates parallel computing.

Value

It returns the following quantities:

`R`	the Clark and Evans' ratio value, which is the average of the observed NND versus the expected NND
`t1`	the average of the observed NND for all distributional points
`t2`	the expected NND under the perfect regularity pattern, details can be referred to Clark and Evans (1954).
`c`	the Z score value for testing significance
`p`	the p value for testing significance

Author(s)

Tsung-Jen Shen & Youhua Chen

References

Clark P, Evans F (1954) Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology, 35, 445–453.

See Also

`NND` `LNND`

Examples

```
xy=cbind(runif(100),runif(100))
plot(xy)
CECI(xy)
```

`LNND` *a weighted function to compute one-dimensional nearest neighboring distance and conduct hypothesis testing for multiple line transects*

Description

find nearest neighbor distance for each distributional point from multiple line transects and conduct hypothesis testing

Usage

```
LNND(lxy)
```

Arguments

`lxy`	lxy is three-column matrix, the first column is the line-transect ID, the second and third columns are x- and y-coordinate respectively. In line transect sampling, usually individuals of species are recorded in a sequential order, so the second and third columns should have been ordered in a time-forward way (backward is fine). Moreover, for the first column, IDs of different line transects should have been sorted according to time-forward sampling ordered.

Value

It returns the following quantities:

`LR`	the one-dimensional NND ratio, which is the average of the observed one-dimensional NND versus the expected NND,weighted over different line transects

`t1`	the average of the observed NND for all distributional points over different line transects
`t2`	the expected NND under one-dimensional perfect regularity pattern, weighted over different line transects. More details can be referred to Chen et al. (2025)
`sig`	the pool-level standard error over different line transects
`c`	the Z score value for testing significance
`p`	the p value for testing significance
`df`	sample size, i.e., the total number of distributional points for analysis

Note

in empirical data provided by the users themselves, because we assume individuals have been recorded in a sequential order and different line transects have been ordered in a sequential order in lxy matrix. so the original data lxy do not need to be sorted and just be used as input directly.

Author(s)

Youhua Chen

References

Xiaoqin Shi, Yongbin Wu, Qi Xiao, Youhua Chen (2026) Linda: an R package using Line transect-based nearest neighbor distance analysis to infer distributional aggregation pattern of species. Plant Diversity.

See Also

`NND`

Examples

```
x=cbind(1,runif(100))
x=rbind(x,cbind(2,runif(100)))
x=rbind(x,cbind(3,runif(100)))
lxy=cbind(x,1)
lxy[,2]=sort(lxy[,2]) #sequentially sampled in an economic way
LNND(lxy)
#in empirical data, because we assume individuals have been recorded in sequential order
#so the original data do not need to be sorted.
#By contrast, in simulated data, if we assumed individuals are recorded in sequential order
#and in an economic way
#sort() function should be used.
```

`NND` *main function to compute one-dimensional nearest neighboring distance and conduct hypothesis testing for a single line transects*

Description

find nearest neighbor distance for each distributional point from a single line transect and conduct hypothesis testing

Usage

```
NND(xy, L = NULL)
```

Arguments

`xy`	xy is two column matrix, the first is x coordinate, and the second is y coordinate. because for line transect sampling, individuals are collected from sequential sampling, so xy should be ordered in a time-forward way (backward is fine) In other words, rows of xy matrix should be sorted according to sampling sequential orders.
`L`	L is the length of the line transect, can be given or calculated from the original data (if L=NULL). L=NULL is recommended and set as a default, because the given line transect length might over-estimate the aggregation pattern!

Value

It returns the following quantities:

`R`	the one-dimensional NND ratio, which is the average of the observed one-dimensional NND versus the expected NND for a single targeted line transect
`ra`	the average of the observed NND for all distributional points for a single targeted line transect
`re`	the expected NND under one-dimensional perfect regularity pattern for a single line transect. More details can be referred to Chen et al. (2025)
`sig`	standard error for the targeted line transect
`c`	the Z score value for testing significance
`p`	the p value for testing significance
`df`	sample size, i.e., the total number of distributional points for analysis
`L`	the estimated or given line transect length

Note

in empirical data provided by the users themselves, because we assume individuals have been recorded in xy matrix in a sequential order, so the original data xy do not need to be sorted and just be used as input directly.

Author(s)

Youhua Chen

References

Xiaoqin Shi, Yongbin Wu, Qi Xiao, Youhua Chen (2026) Linda: an R package using Line transect-based nearest neighbor distance analysis to infer distributional aggregation pattern of species. Plant Diversity.

See Also

LNND

Examples

```
xy=cbind(sort(runif(100)),1) #sequentially sampled in an economic way
NND(xy)
#in empirical data, because we assume individuals have been recorded in sequential order
#so the original data do not need to be sorted.
#By contrast, in simulated data, if we assumed individuals are recorded in sequential order
#and in an economic way
#sort() function should be used.
```

`pi.est` *estimation of the inertia parameter pi for measuring the degree of non-independence of any two adjacent individuals along a line transect*

Description

estimation of the inertia parameter pi, more details can be referred to Chen et al. (2023). This parameter has been closely related to Moran's I index, more details can be referred to Chen and Shen (2020).

Usage

```
pi.est(z)
```

Arguments

`z`	z is a vector of species labels for sequentially sampled individuals when walking across a line transect

Details

suppose we collect five individuals of three species along a line transect in sequential oder as "ABCAA", then the vector z=c("A","B","C","A","A")

Value

the estimation of the inertia parameter is returned, and this value is bounded between 0 and 1.

Note

in empirical data provided by the users themselves, because we assume individuals have been recorded in the vector z in a sequential order, so the original dataset z do not need to be sorted and just be used as input directly.

Author(s)

Tsung-Jen Shen

References

Chen et al. (2023) Biodiversity survey and estimation for line-transect sampling. Frontiers in Plant Science, 14: 1159090. Chen and Shen (2020) Unifying conspecific-encounter index v and Morans' I index. Ecography, 43, 1902-1904.

See Also

`NND` `LNND` `v.est`

Examples

```
z=sample(1:5,100,replace=TRUE)
pi.est(z)
```

`v.est` *conspecific-encounter index for line transect-collecte biodiversity data*

Description

this is the line transect-derived conspecific encounter index, which is also the estimator of pi by Solow (2000). Note that conspecific-encounter index for line transect sampling is completely different to the conventional Simpson diversity index, which is proposed as a conspecific-encounter index under random sampling.

Usage

```
v.est(z)
```

Arguments

`z`	z is a vector of species labels for sequentially sampled individuals when walking across a line transect

Details

suppose we collect five individuals of three species along a line transect in sequential oder as "AB-CAA", then the vector z=c("A","B","C","A","A")

Value

the conspecific-encounter index value is returned, and this value is bounded between 0 and 1.

Note

in empirical data provided by the users themselves, because we assume individuals have been recorded in the vector z in a sequential order, so the original dataset z do not need to be sorted and just be used as input directly.

Author(s)

Youhua Chen & Tsung-Jen Shen

References

Chen et al. (2019) Inferring multispecies distributional aggregation level from limited line transect-derived biodiversity data. Methods in Ecology and Evolution, 10, 1015-1023. Solow A (2000) The effect of dependence on estimating sample coverage. Environmetrics, 11, 245-249.

See Also

`NND` `LNND` `pi.est`

Examples

```
z=sample(1:5,100,replace=TRUE)
v.est(z)
```

Index