Package: SppTrend (via r-universe)

May 9, 2026

Title Analyzing Linear Trends in Species Occurrence Data

Version 0.4

Description Provides a methodology to analyze how species occurrences change over time, particularly in relation to spatial and thermal factors. It facilitates the development of explanatory hypotheses about the impact of environmental shifts on species by analyzing historical presence data that includes temporal and geographic information. Approach described in Lobo et al., 2023 <doi:10.1002/ece3.10674>.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.3

Suggests testthat (>= 3.0.0), devtools

Imports dplyr, tidyr, terra, stats, stringr, tidyselect, ggplot2, patchwork, sf

Config/testthat/edition 3

NeedsCompilation no

Author Mario Mingarro [aut, cre] (ORCID: <https://orcid.org/0000-0003-3977-7944>), Emilio García-Roselló [aut] (ORCID: <https://orcid.org/0000-0003-0305-770X>), Jorge M. Lobo [aut] (ORCID: <https://orcid.org/0000-0002-3152-4769>)

Maintainer Mario Mingarro <mario_mingarro@mncn.csic.es>

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

Date/Publication 2026-02-07 12:50:02 UTC

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

RemoteRef HEAD

RemoteSha 850d406c711a037b48f080c14644027443008326

Contents

`get_elevation` *Extract elevation from DEM*

Description

This function retrieves elevation values from a Digital Elevation Model (DEM) based on their geographic coordinates (`lon`/`lat`).

Usage

```
get_elevation(data, dem_file)
```

Arguments

`data`	A `data frame` containing species records. Must include `lon`, `lat`, `year`, and `month` columns.
`dem_file`	Full `character` path to the downloaded DEM raster file.

Value

The input data frame `data` with a new column (`ele`) containing the extracted elevation values.

`get_era5_tme` *Extract temperature data from ERA5 NetCDF file*

Description

This function retrieves mean monthly air temperature values associated with species occurrence records based on their geographic coordinates (`lon`/`lat`) and sampling date (`year`/`month`).

Usage

```
get_era5_tme(data, nc_file)
```

Arguments

- `data` — A `data frame` containing species records. Must include `lon`, `lat`, `year`, and `month` columns.
- `nc_file` — Full `character` path to the downloaded ERA5-Land raster (`.nc`) file.

Value

The input data frame `data` with a new column named (`tme`), containing the temperature values.

`get_fast_info` *Quick visual diagnostic of the input data*

Description

This function provides a quick visual diagnostic of the input data. It generates a map showing the spatial distribution of occurrence records together with a time-series plot derived from a NetCDF environmental dataste, including a linear trend analysis. Using the geographic coordinates of the occurrence records, the function extracts the complete climate time-series (from the earliest to the latest year represented in the data) for the corresponding occupied cells. All temperature values from occupied cells are then added annually to estimate and visualise the overall temperature trend (including slope and associated p-value). This diagnostic step allows users to quickly assess the climate trajectory of the regions where the species have been recorded and to evaluate whether sufficient temporal and environmental variation is present for subsequent analyses.

Usage

```
get_fast_info(data, nc_file)
```

Arguments

- `data` — A data frame containing species records. Must include `lon`, `lat`, `year`, and `month` columns.
- `nc_file` — Full `character` path to the downloaded ERA5-Land raster (`.nc`) file.

Value

Invisibly returns a composite plot. Displays a composite plot showing the geographic distribution and the thermal trend with its corresponding global slope and p-value.

`overall_trend` *Overall trend analysis*

Description

Calculates the overall temporal trend (OT) of selected response variables across the entire dataset. This trend integrates both environmental change and the cumulative effects of sampling bias, and serves as a neutral reference against which species-specific temporal trends are evaluated.

Usage

```
overall_trend(data, predictor, responses)
```

Arguments

`data`	A `data frame` containing the variables for the model, including `species`, `year`, `month`, `lon`, `lat`, `tme` and `ele`.
`predictor`	A `character`vector of predictor variable names representing a temporal variable (`year_month`).
`responses`	A `character` vector of response variable names to analyze.

Details

Longitude (`lon`) values are transformed to a 0-360 range to ensure statistical consistency near the antimeridian. A key feature of this function is its specialized handling of latitude. Because the Equator is set at 0, latitude values in the Southern Hemisphere are negative. To ensure that a direction shift is interpreted consistently across the globe (where a negative increase in the South corresponds to a positive increase in the North), the function employs two complementary approaches: Hemispheric split: It divides the records based on their location (`lat < 0` for `South` and `lat > 0` for `North`) and performs separate analyses for each. Global analysis: It performs an analysis using the complete dataset (`Global`) by transforming all latitudes into absolute values (`abs(lat)`). This allows for a unified global trend estimation. Note that this hemispheric division and absolute transformation logic is applied exclusively to the latitude (`lat`) variable.

Value

A data frame with trend statistics, including:

- `responses`: The name of the variable analyzed.
- `trend`: Slope of the linear model (rate of change over time).
- `t`: t-statistic of the model.
- `pvalue`: Statistical significance of the overall trend.
- `ci_95_max`, `ci_95_min`: 95\
- `n`: Sample size for the specific species/hemisphere subset
- `hemisphere`: Geographic context (`North`, `South`, or `Both` for global comparison).

Examples

```
data <- data.frame(
   species = sample(paste0("spp_", 1:10), 500, replace = TRUE),
   year = sample(1950:2020, 500, replace = TRUE),
   month = sample(1:12, 500, replace = TRUE),
   lon = runif(500, -10, 20),
   lat = runif(500, 30, 70),
   tme = rnorm(500, 15, 10)
)

data$year_month <- data$year + data$month * 0.075

predictor <- "year_month"
responses <- c("lat", "lon", "tme")

overall_trend_result <- overall_trend(data, predictor, responses)

print(overall_trend_result)
```

---

spp_strategy *Classify species ecological strategies*

Description

This function analyses the outputs of `spp_trend()` to classify species into distinct spatial or thermal response categories based on the direction and statistical significance of their species-specific trends relative to the overall trend. The function incorporates hemisphere-specific logic to correctly interpret poleward shifts in latitude and can also be applied to classify elevational trends.

Usage

```
spp_strategy(spp_trend_result, sig_level = 0.05, responses = responses)
```

Arguments

`spp_trend_result`
: A data frame containing trend indicators per species, typically generated by the `spp_trend` function. It should include columns such as:

- `species`: Name of the analyzed species.
- `responses`: Name of the analyzed variable.
- `trend`: Estimated slope of the linear model.
- `t`: t-statistic for the species-specific trend.
- `pvalue`: Statistical significance of the species-specific trend.
- `dif_t`: t-statistic of the interaction term, indicating the magnitude of the difference between the species trend and the Overall Trend (OT).

- dif_pvalue: p-values of the interaction term. A low value indicates a significant deviation from the general trend.
- n: Total number of occurrence records (sample size) for the specific species.
- hemisphere: Geographical subset (North, South, or Global) used to ensure latitudinal symmetry in the analysis.

sig_level
: The numeric significance level to use for classifying trends as significant. Defaults to 0.05. See Bonferroni correction 0.05/length(species).

responses
: A character vector of response variable names to analyze (c("lat", "lon", "tme", "ele")). The function will create classification columns for responses present in this vector and in the responses column of spp_trend_result.

Details

This function takes the trend analysis results from spp_trend and classifies each species' response based on the significance of its trend and how it differs from the general trend. Applied Bonferroni correction to avoid false positives (Type I errors) due to multiple comparisons when analyzing many species. The classification identifying three possible spatial responses and three thermal responses:

- **Spatial Responses:**
  - **Spatial Adaptation (SA):** A significant positive temporal trend in the spatial position of species occurrences. In the context of climate change, this pattern is commonly associated with a poleward shift, corresponding to a northward displacement (towards higher latitude values) in the Northern Hemisphere and southward displacement (towards lower latitude values) in the Southern Hemisphere, as species expand into newly suitable areas.
  - **Spatial Discordance (SD):** A significant negative temporal trend in the spatial position of species occurrences. In the context of climate change, this pattern is often associated with an equatorward shift and may arise when other ecological and anthropogenic factors influence species distributions independently of, or in opposition to, climate-driven range shifts.
  - **Spatial Conformity (SC):** A spatial response pattern in which the species-specific temporal trend does not differ significantly from the overall trend. Species showing spatial conformance share the same bias structure as the complete dataset, preventing the inference of a distinct, species-specific response to climate change at the scale of analysis.
- **Thermal Responses:**
  - **Thermal Tolerance (TT):** A thermal response pattern characterised by a significant positive temporal trend in the temperature conditions under which species are observed, relative to the overall trend. This pattern suggest an increased likelihood of occurrence under warmer conditions and an apparent capacity to tolerate rising temperatures through physiological, behavioural, and evolutionary mechanisms.
  - **Thermal Adjustment (TA):** A thermal response characterised by a significant negative temporal trend in the temperature conditions associated with species occurrences, relative to the overall trend. This indicates and increasing association with cooler temperature conditions over time, potentially reflecting microevolutionary change or phenotypic adjustment.
  - **Thermal Conformity (TC):** A thermal response pattern in which species-specific temperature trends do not differ significantly from the overall trend. Species showing thermal conformance share the same background thermal signal as the complete dataset, preventing the formulation of specific hypotheses regarding climate-driven thermal responses.

Note: The interpretation of longitude trends assumes that if transformation was applied in `spp_trend`*, it used the Antimeridian as 0.*

Value

A `data frame` summarizing the ecological strategy of each species for each analyzed response variable. The table includes:

- Species name
- Hemisphere
- Sample size
- Classification columns for:
  - Spatial (latitude, longitude and elevation if present) responses. Spatial Adaptation `SA`, Spatial Discordance `SD`, Spatial Conformity `SC`
  - Thermal (temperature if present) responses. Thermal Tolerance `TT`, Thermal Adjustment `TA`, Thermal Conformity `TC`

Classification for spatial responses (`lat`, `lon`, `ele`) are classified as `Spatial_lat`, `Spatial_lon` and `Spatial_ele`. Thermal responses (`tme`) are classified as `Thermal_tme`.

Examples

```
# Assuming spp_trends_results is a data frame generated by spp_trend()

spp_trends_results <- data.frame(
  species = paste0("spp_", 1:10),
  responses = rep(c("lat", "lon", "tme"), length.out = 30),
  trend = runif(30, -0.5, 0.5),
  t = runif(30, -2, 2),
  pvalue = runif(30, 0, 1),
  dif_t = runif(30, -1, 1.5),
  dif_pvalue = runif(30, 0.001, 0.9),
  n = round(runif(30, 40, 60)),
  hemisphere = sample(c("North", "South", "Global"), 30, replace = TRUE)
)

spp <- unique(spp_trends_results$species)
sig_level <- 0.05 / length(spp) # Bonferroni correction
responses_to_analyze <- c("lat", "lon", "tme")

spp_strategy_results <- spp_strategy(spp_trends_results,
                                     sig_level = sig_level,
                                     responses = responses_to_analyze)

print(spp_strategy_results)
```

spp_trend *Individual trend analysis*

Description

Estimates the species-specific temporal trends for each selected response variable and statistically compares them with the overall temporal trend derived from the complete dataset. It compares individual species' trajectories against the OT using the interaction term of the `lm()`.

Usage

```
spp_trend(data, spp, predictor, responses, n_min = 50)
```

Arguments

`data`	A `data frame` containing the variables for the model, including `species`, `year`, `month`, `lon`, `lat`, `tme` and/or `ele`.
`spp`	A `character` vector of unique species names.
`predictor`	A `character` vector of predictor variable names representing a temporal variable (`year_month`).
`responses`	A `character` vector of response variable names to analyze.
`n_min`	Minimum `numeric` number of presences required for a species in each hemisphere (or globally for species in both hemispheres) to perform the analysis.

Details

The function fits linear models for each species and compares them to the general trend using an interaction model (`response ~ predictor * group`). Longitude (`lon`) values are transformed to a 0-360 range to ensure statistical consistency near the antimeridian. A key feature of this function is its specialized handling of latitude. Because the Equator is set at 0, latitude values in the Southern Hemisphere are negative. To ensure that a direction shift is interpreted consistently across the globe (where a negative increase in the South corresponds to a positive increase in the North), the function employs two complementary approaches: Hemispheric split: It divides the records based on their location (`lat < 0` for `South` and `lat > 0` for `North`) and performs separate analyses for each. Global analysis: It performs an analysis using the complete dataset (`Global`) by transforming all latitudes into absolute values (`abs(lat)`). This allows for a unified global trend estimation. Note that this hemispheric division and absolute transformation logic is applied exclusively to the latitude (`lat`) variable.

Value

A data frame with trend statistics, including:

- `species`: Name of the analyzed species.
- `responses`: Name of the variable analyzed.
- `trend`: Slope of the linear model (rate of change over time).

- `t`: t-statistic for the species-specific trend.
- `pvalue`: Statistical significance of the species trend.
- `ci_95_max`, `ci_95_min`: 95% confidence interval bounds for the slope.
- `dif_t`: t-statistic of the interaction term (species vs. baseline).
- `dif_pvalue`: p-values of the interaction term. A low value indicates a significant deviation from the general trend.
- `n`: Sample size for the specific species/hemisphere subset
- `hemisphere`: Geographic context (`North`, `South`, or `Global` for global comparison).

Examples

```
data <- data.frame(
   species = sample(paste0("spp_", 1:10), 500, replace = TRUE),
   year = sample(1950:2020, 500, replace = TRUE),
   month = sample(1:12, 500, replace = TRUE),
   lon = runif(500, -10, 20),
   lat = runif(500, 30, 70),
   tme = rnorm(500, 15, 10)
)

data$year_month <- data$year + data$month * 0.075

predictor <- "year_month"
responses <- c("lat", "lon", "tme")

spp <- unique(data$species)

spp_trend_result <- spp_trend(data, spp, predictor, responses, n_min = 50)

print(spp_trend_result)
```

Index