Package ‘avesperu’

January 9, 2026

Title Access to the List of Birds Species of Peru

Version 0.0.8

Description Allows access to the data found in the species list featured in the renowned 'List of the Birds of Peru' Plenge, M. A. (2023) <`https://sites.google.com/site/boletinunop/checklist`>. This publication stands as one of Peru's most comprehensive reviews of bird diversity. The dataset incorporates detailed species accounts and has been meticulously structured for effortless utilization within the R environment.

License MIT + file LICENSE

Suggests testthat (>= 3.0.0), xml2

Config/testthat/edition 3

Encoding UTF-8

RoxygenNote 7.3.3

URL `https://github.com/PaulESantos/avesperu`,

`https://paulesantos.github.io/avesperu/`

BugReports `https://github.com/PaulESantos/avesperu/issues`

Depends R (>= 3.5)

LazyData true

Maintainer Paul E. Santos Andrade <`paulefrens@gmail.com`>

Imports cli, parallel

NeedsCompilation no

Author Paul E. Santos Andrade [aut, cre] (ORCID: <`https://orcid.org/0000-0002-6635-0375`>)

Repository CRAN

Date/Publication 2026-01-09 00:40:02 UTC

Contents

`aves_peru_2025_v4` *aves_peru_2025_v4*

Description

The `aves_peru_2025_v4` dataset provides an updated tibble of bird species recorded in Peru, based on the most recent taxonomic revisions by the South American Checklist Committee (SACC).

Usage

```
aves_peru_2025_v4
```

Format

A tibble with 1,917 rows and 6 columns:

order_name Taxonomic order to which the bird species belongs.

family_name Taxonomic family to which the bird species belongs.

scientific_name Scientific name of the bird species.

english_name English common name of the bird species.

spanish_name Spanish common name of the bird species.

status Category indicating the species' status, based on the following codes:

- `X`: Resident species.
- `E`: Endemic species. A species is considered endemic to Peru until a record outside its borders is published.
- `NB`: Non-breeding (migratory) species. Species that occur regularly in Peru but only during their non-breeding period.
- `V`: Vagrant species. Species that occasionally occur in Peru but are not part of the usual avifauna.
- `IN`: Introduced species. Species introduced to Peru by humans (or have colonized from introduced populations elsewhere) and have established self-sustaining breeding populations.
- `H`: Hypothetical species. Records based only on observations, specimens of dubious origin, or unpublished photographs or recordings kept in private hands.
- `EX`: Extinct or extirpated species. Species that have gone extinct or have been extirpated from Peru.

Details

This version reflects dramatic taxonomic changes and category updates based on published articles, photographs, and sound recordings archived in accredited institutions. It also includes a classification criterion following the SACC guidelines. Species without a specific code are considered resident species, equivalent to the "X" category of the SACC.

- **Total species**: 1,917
- **Distribution by status**:
  - `X`: 1,547 species
  - `E`: 120 species
  - `NB`: 139 species
  - `V`: 85 species
  - `IN`: 3 species
  - `EX`: 0 species
  - `H`: 23 species

These updates reflect the SACC's continuous evaluation process, which now recognizes several former subspecies as full species.

Author(s)

Data compilation: Manuel A. Plenge Package implementation: Paul Efren Santos Andrade

References

Plenge, M. A. Version (29-09-2025) List of the birds of Peru / Lista de las aves del Perú. Unión de Ornitólogos del Perú: `https://sites.google.com/site/boletinunop/checklist`

See Also

For more information about the data, visit: `https://sites.google.com/site/boletinunop/checklist`

Examples

```
# Load the dataset
data("aves_peru_2025_v4")
```

`aves_peru_2025_v5` *aves_peru_2025_v5*

Description

The `aves_peru_2025_v5` dataset provides the most current and comprehensive tibble of bird species recorded in Peru, based on the latest taxonomic revisions by the South American Checklist Committee (SACC) as of December 22, 2025.

Usage

```
aves_peru_2025_v5
```

Format

A tibble with 1,919 rows and 9 columns:

order_name Character. Taxonomic order to which the bird species belongs.

family_name Character. Taxonomic family to which the bird species belongs.

genus Character. Genus name of the bird species.

species_epithet Character. Specific epithet (species name without genus).

scientific_name Character. Complete scientific name of the bird species (binomial nomenclature: genus + species epithet).

english_name Character. Common name in English.

spanish_name Character. Common name in Spanish (Peruvian usage).

status Character. Conservation and occurrence status category in Spanish. See Details section for complete descriptions.

status_code Character. Original SACC status code. Values: X, E, NB, V, IN, U, EX. See Details section for code definitions.

Details

All records are based on published evidence (specimens, photographs, videos, or recordings) deposited in accredited institutional collections. The dataset follows strict inclusion criteria established by the SACC and the Unión de Ornitólogos del Perú (UNOP).

Dataset Summary:

- **Total species**: 1,919
- **Version date**: December 29, 2025
- **SACC baseline date**: December 22, 2025

Distribution by Status:

Status	Code	Count	Description

Residente	X	~1,547	Resident breeding species
Endémico	E	~120	Endemic to Peru
Migratorio	NB	~140	Non-breeding migrants
Divagante	V	~85	Vagrant species
Introducido	IN	3	Introduced species
No confirmado	U	~23	Unconfirmed records
Extirpado	EX	0	Extirpated species

Status Categories (Detailed):

Residente (X - Resident):
Species that breed in Peru and maintain permanent or seasonal populations. This is the default category for species without a specific status code.

Endémico (E - Endemic):
Species whose entire known range is within Peru. A species is considered endemic until a published record documents its occurrence outside Peruvian borders.

Migratorio (NB - Non-breeding):
Species that occur regularly in Peru but only during their non-breeding period. These are typically austral or boreal migrants that breed elsewhere.

Divagante (V - Vagrant):
Species that occur occasionally in Peru and are not part of the regular avifauna. These represent extralimital records or irregular visitors.

Introducido (IN - Introduced):
Species introduced to Peru by humans (directly or colonized from introduced populations elsewhere) that have established self-sustaining breeding populations.

No confirmado (U - Unconfirmed):
Records that lack definitive published evidence. This includes:

- Sight records without corroborating physical evidence
- Specimens of dubious or uncertain origin
- Unpublished photographs or recordings in private collections

Extirpado (EX - Extirpated/Extinct):
Species that have gone extinct globally or have been extirpated from Peru.

Taxonomic Authority:

The taxonomic sequence and species limits follow the South American Checklist Committee (SACC) of the American Ornithological Society, reflecting the committee's decisions through December 22, 2025.

Note

This dataset is updated periodically as new species are documented and taxonomic revisions are published. Check the UNOP website for the most current version.

Author(s)

Data compilation: Manuel A. Plenge & Fernando Angulo Package implementation: Paul Efren Santos Andrade

Source

Data compiled by Manuel A. Plenge and Fernando Angulo (UNOP). For corrections or updates, contact: chamaepetes@gmail.com

References

Plenge, M. A. & F. Angulo. Version 29-12-2025. Lista de las aves del Perú / List of the birds of Peru. Unión de Ornitólogos del Perú: `https://sites.google.com/site/boletinunop/checklist`

See Also

- UNOP Checklist: `https://sites.google.com/site/boletinunop/checklist`
- SACC: `http://www.museum.lsu.edu/~Remsen/SACCBaseline.htm`
- `search_avesperu` for species name validation

Examples

```
# Load the dataset
data("aves_peru_2025_v5")

# View structure
str(aves_peru_2025_v5)

# Summary by status
table(aves_peru_2025_v5$status)
```

---

search_avesperu   *Search for Bird Species Data in the Birds of Peru Dataset*

Description

This function searches for bird species information in the dataset provided by the `avesperu` package, given a list of species names. It supports approximate (fuzzy) matching to handle typographical errors or minor variations in species names using optimized `agrep()` matching. The function is optimized for both small and large lists through intelligent pre-filtering and optional parallel processing, while maintaining exact `agrep()` precision.

Usage

```
search_avesperu(
  splist,
  max_distance = 0.1,
  return_details = FALSE,
  batch_size = 100,
  parallel = TRUE,
  n_cores = NULL
)
```

Arguments

`splist`	A character vector or factor containing the scientific names of bird species to search for. Names can include minor variations or typos.
`max_distance`	Numeric. The maximum allowable distance for fuzzy matching. Can be either: • A proportion between 0 and 1 (e.g., 0.1 = 10\ • An integer representing the maximum number of character differences Default: 0.1.
`return_details`	Logical. If `FALSE` (default), returns only a character vector of species status. If `TRUE`, returns a detailed data frame with complete reconciliation information including taxonomic data and matching distances.
`batch_size`	Integer. Number of species to process per batch when handling large lists. Useful for memory management and progress tracking. Default: 100 species per batch.
`parallel`	Logical. Should parallel processing be used for large lists? Automatically disabled for small lists. Requires the `parallel` package. Default: `TRUE`.
`n_cores`	Integer or `NULL`. Number of CPU cores to use for parallel processing. If `NULL` (default), uses `detectCores() - 1` to leave one core free for system operations.

Details

The function performs the following steps:

1. Validates input and converts factors to character vectors
2. Standardizes species names using `standardize_names()`
3. Identifies and reports duplicate entries in the input list
4. Uses intelligent pre-filtering to reduce search space:
    - Filters by string length (mathematically guaranteed to preserve matches)
    - Optionally filters by first character for very large candidate sets
5. Performs precise `agrep()` fuzzy matching on filtered candidates
6. Calculates exact edit distances using `adist()`
7. Selects the best match (minimum distance) for each query
8. For large lists (>batch_size), processes in batches with optional parallelization

Value

The return value depends on the `return_details` parameter:

If return_details = FALSE (default):

A character vector with the same length as `splist`, containing the conservation/occurrence status for each species. `NA` values indicate no match was found.

If return_details = TRUE:

A data frame (tibble-compatible) with the following columns:

name_submitted Character. The species name provided as input (standardized).

accepted_name Character. The closest matching species name from the database, or `NA` if no match found within `max_distance`.

order_name Character. The taxonomic order of the matched species.

family_name Character. The taxonomic family of the matched species.

english_name Character. Common name in English.

spanish_name Character. Common name in Spanish.

status Character. Conservation or occurrence status (e.g., "Endemic", "Resident", "Migrant", "Vagrant").

dist Character. Edit distance between submitted and matched names. Lower values indicate better matches. `NA` if no match found.

Warning

For very large lists (>10,000 species) with parallel processing enabled, ensure sufficient system memory is available. Each parallel worker maintains a copy of the reference database (~5-10 MB).

See Also

`agrep` for the underlying fuzzy matching algorithm

Examples

```
## Not run:
# Basic usage - returns status vector
splist <- c("Falco sparverius", "Tinamus osgodi", "Crypturellus soui")
status <- search_avesperu(splist)
print(status)

# Get detailed reconciliation information
details <- search_avesperu(splist, return_details = TRUE)
print(details)

# Exact matching only (no fuzzy matching)
exact_results <- search_avesperu(splist, max_distance = 0)

# Handle species with typos
typo_list <- c("Falco sparveruis", "Tinamus osgoodi", "Crypturellus sui")
corrected <- search_avesperu(typo_list, return_details = TRUE)

# View submitted vs accepted names
print(corrected[, c("name_submitted", "accepted_name", "dist")])


## End(Not run)
```

`show_progress` *Determine whether to show progress bar Return logical TRUE/FALSE depending on options and interactive session*

Description

Determine whether to show progress bar Return logical TRUE/FALSE depending on options and interactive session

Usage

```
show_progress()
```

Index