Package ‘snirh.lab’

October 25, 2025

Type Package

Title Convert Laboratory Water-Quality Data to 'SNIRH' Import Format

Version 0.1.0

Description Convert laboratory data to the Portuguese Information System for Water Resources 'SNIRH' file format. 'SNIRH' is Portugal's national water resources information system <`https://snirh.apambiente.pt/`>. The package validates station data, converts parameters and units, and generates compliant output files for data submission.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

Language en-US

Depends R (>= 4.0.0)

Imports data.table (>= 1.14.0), cli (>= 3.0.0), utils

Suggests sf (>= 1.0.0), curl (>= 4.0.0), testthat (>= 3.0.0), rmarkdown, spelling

URL `https://github.com/lpereira-ue/snirh.lab`

BugReports `https://github.com/lpereira-ue/snirh.lab/issues`

Config/testthat/edition 3

NeedsCompilation no

Author Luís Pereira [aut, cre] (ORCID: <`https://orcid.org/0000-0002-0628-4847`>)

Maintainer Luís Pereira <`d57177@alunos.uevora.pt`>

Repository CRAN

Date/Publication 2025-10-25 12:50:03 UTC

Contents

`check_station_status` *Check status of specific SNIRH stations*

Description

Validates specific station IDs against the SNIRH database and returns their current status. This is useful for checking stations before running the full conversion process.

Usage

```
check_station_status(station_ids, matrix = "surface.water", timeout = 30)
```

Arguments

`station_ids`	Character vector of station IDs to check.
`matrix`	Character string specifying the matrix type. Currently supports "surface.water" and "biota". Default is "surface.water".
`timeout`	Numeric. Timeout in seconds for downloading data. Default is 30.

Details

This function is particularly useful for:

- Pre-validating station IDs before data conversion
- Checking why certain stations fail validation
- Getting an overview of station status for reporting

Value

A data.table with the following columns:

station_id The station ID that was checked

found Logical indicating if station exists in SNIRH

status Station status if found, NA if not found

active Logical indicating if station is active (status = "ATIVA")

See Also

`get_snirh_stations` for getting all station information

`convert_to_snirh` for the main conversion function

Examples

```
# Check status of specific stations
my_stations <- c("07G/50", "25G/07", "INVALID_ID")
status_check <- check_station_status(my_stations)
print(status_check)

# Check which stations are not active
inactive <- status_check[active == FALSE | is.na(active)]
if (nrow(inactive) > 0) {
  print("Stations requiring attention:")
  print(inactive)
  }

# Check only active stations
active_stations <- status_check[active == TRUE]
```

`convert_to_snirh` *Convert data to SNIRH file format*

Description

Cleans and converts laboratory data to the SNIRH (National Information System on Water Resources) import format. It handles data validation, unit conversions, station validation, and formatting according to SNIRH standards.

Usage

```
convert_to_snirh(data, matrix, validate_stations = NULL, timeout = 30)
```

Arguments

`data`	A data.frame or data.table containing the original laboratory data. Must contain the following columns in order: snirh_entity, station_name, station_id, sampling_date, parameter, unit, value.
`matrix`	Character string specifying the type of matrix being processed. Must be one of: "surface.water" or "biota".
`validate_stations`	Logical. Whether to validate station IDs against the SNIRH database. Defaults to TRUE for surface.water and biota matrices. Requires internet connection.
`timeout`	Numeric. Timeout in seconds for downloading station data. Default is 30 seconds.

Details

The function performs several key operations:

- Validates input data structure and removes empty rows/columns
- Validates station IDs against SNIRH database (for surface.water and biota)
- Checks for duplicate measurements (same station, date, and parameter)
- Extracts pH temperature measurements when present
- Converts measurement values to SNIRH-compatible units
- Handles measurement flags (<, >, =) and special values
- Formats output according to SNIRH import specifications

Value

A data.table formatted for SNIRH import with the following structure:

- First row contains network specification (REDE=NETWORK_NAME)
- Station identifiers (ESTACAO=STATION_ID) before each group of measurements
- Date/time stamps in DD/MM/YYYY HH:MM format
- Parameter values in SNIRH-compatible units and symbols

Station Validation

For surface.water and biota matrices, the function validates that:

- All station IDs exist in the SNIRH database
- All stations have status "ATIVA" (active)
- Internet connection is available for downloading station data

If validation fails, the function will stop and provide details about invalid stations that need to be corrected in the database.

Input Data Requirements

The input data must be a data.frame/data.table with exactly these columns:

snirh_entity Entity responsible for the data

station_name Human-readable station name

station_id Unique station identifier (must match SNIRH database)

sampling_date Date and time of sampling (POSIXct recommended)

parameter Parameter name as used in laboratory

unit Unit of measurement as used in laboratory

value Measured value (may include flags like <, >)

Parameter Conversion

Relies on an internal `parameters` dataset that maps laboratory parameter names and units to SNIRH equivalents. This dataset must contain conversion factors and SNIRH symbols for all parameters in the input data.

Examples

```
# Example data structure

lab_data <- data.table::data.table(
  snirh_entity = "APA",
  station_name = "River station 1",
  station_id = "01F/01",  # Must be valid SNIRH station ID
  sampling_date = as.POSIXct("2024-01-15 10:30:00"),
  parameter = "pH - Campo",
  unit = "Escala Sorensen",
  value = "7.2"
)

# Convert surface water data (with station validation)
snirh_data <- convert_to_snirh(lab_data, "surface.water")

# Skip station validation if needed (not recommended)
snirh_data <- convert_to_snirh(lab_data, "surface.water",
                               validate_stations = FALSE)
```

`get_snirh_stations` *Get SNIRH station information*

Description

Downloads and returns information about SNIRH monitoring stations for surface water quality. This function can be used to check station status, get available station IDs, or validate stations before data conversion.

Usage

```
get_snirh_stations(matrix = "surface.water", timeout = 30, active_only = FALSE)
```

Arguments

`matrix`	Character string specifying the matrix type. Currently supports "surface.water" and "biota" (both use the same station database).
`timeout`	Numeric. Timeout in seconds for downloading data. Default is 30.
`active_only`	Logical. If TRUE, returns only active stations (Estado = "ATIVA"). If FALSE, returns all stations. Default is FALSE.

Details

Downloads the latest station information from the SNIRH geoportal. It requires an internet connection and the ’sf’ package for processing shapefiles.

The station database includes information about:

- Station location (coordinates)
- Station status (active/inactive)
- Station metadata

Value

A data.table with station information containing:

station_id Station identifier (corresponds to "Código" in SNIRH)

status Station status (e.g., "ATIVA", "EXTINTA")

geometry Spatial geometry (if sf package is available)

Station Status

Stations can have different status values:

ATIVA Station is active and can receive new data

INATIVA Station is inactive (historical data only)

EXTINTA Station is permanently suspended and has no data

See Also

`convert_to_snirh` for the main conversion function

`check_station_status` for checking specific stations

Examples

```
# Get all surface water stations
all_stations <- get_snirh_stations("surface.water")
print(head(all_stations))

# Get only active stations
active_stations <- get_snirh_stations("surface.water", active_only = TRUE)
print(paste("Active stations:", nrow(active_stations)))

# Check if specific stations are active
my_stations <- c("07H/50", "25G/07")
station_info <- get_snirh_stations("surface.water")
station_status <- station_info[station_id %in% my_stations]
print(station_status)
```

`list_snirh_parameters` *List available SNIRH parameters*

Description

Returns a summary of available parameters in the conversion table, organized by sample type. This helps users understand what parameters can be converted to SNIRH format.

Usage

```
list_snirh_parameters(sample_type = "all", include_conversion_info = FALSE)
```

Arguments

`sample_type`	Character string specifying the sample type to filter by. Must be one of "water", "biota", "sediment", or "all". Default is "all".
`include_conversion_info`	Logical. If TRUE, includes conversion factors and unit information. Default is FALSE.

Details

This function provides an overview of the parameter conversion capabilities of the package. It can help users:

- Understand what parameters are supported
- Check parameter naming conventions
- Verify unit conversion factors
- Plan data preparation activities

Value

A data.table with parameter information. Columns depend on include_conversion_info parameter.

See Also

`parameters` for the complete parameter dataset

Examples

```
# List all water parameters
water_params <- list_snirh_parameters("water")
print(head(water_params))

# Get detailed conversion information
detailed_params <- list_snirh_parameters("water", include_conversion_info = TRUE)
print(head(detailed_params))
```

```
# Check all available sample types
all_params <- list_snirh_parameters("all")
unique_types <- unique(all_params$sample_type)
print(paste("Available sample types:", paste(unique_types, collapse = ", ")))
```

`parameters` *Parameter conversion table for SNIRH format*

Description

Dataset containing the mapping between laboratory parameter names/units and their equivalent SNIRH (Sistema Nacional de Informação de Recursos Hídricos) database format. It includes conversion factors for unit transformations and standardized symbols used in the SNIRH system.

Usage

```
parameters
```

Format

A data.table with 7 variables and multiple rows covering water quality, sediment, and biota parameters:

param_lab Character. Parameter name as provided by the laboratory. These are the original parameter names found in laboratory reports and may include special characters, accents, or laboratory-specific naming conventions.

unit_lab Character. Unit of measurement as provided by the laboratory. These represent the original units used in laboratory measurements and may vary between laboratories or analytical methods.

symbol_snirh Character. Standardized parameter symbol used in the SNIRH database. These symbols are unique identifiers that allow for consistent data storage and retrieval in the national database.

param_snirh Character. Standardized parameter name used in SNIRH. These names follow SNIRH conventions and provide consistency across different data sources and time periods.

unit_snirh Character. Standardized unit used in the SNIRH database. All measurements are converted to these standard units to ensure comparability and compliance with national monitoring standards.

factor Numeric. Conversion factor to transform laboratory units to SNIRH units. The formula is: snirh_value = lab_value * factor. For example, if converting mg/L to µg/L, the factor would be 1000.

sample_type Character. Type of sample matrix. Valid values are:

- `water`: Surface water and groundwater samples
- `biota`: Biological samples (fish, plants, etc.)
- `sediment`: Sediment samples from aquatic environments

Details

This dataset is essential for the `convert_to_snirh` function, which uses it to:

- Validate that all laboratory parameters can be converted to SNIRH format
- Apply appropriate unit conversions using the conversion factors
- Map laboratory parameter names to standardized SNIRH symbols
- Ensure data quality and consistency with national standards

The conversion factors are carefully calibrated to maintain measurement accuracy while ensuring compliance with SNIRH database requirements. Parameters without a direct SNIRH equivalent are not included in this table and will cause the conversion function to raise an error.

Parameter Categories

The parameters are organized by sample type:

Water parameters Include physical properties (temperature, pH, conductivity), chemical parameters (nutrients, metals, organic compounds), and biological indicators.

Sediment parameters Cover grain size distribution, chemical composition, contaminant levels, and organic matter content.

Biota parameters Include bioaccumulation measurements, biological indices, and organism-specific parameters.

Data Quality

This dataset is maintained according to:

- SNIRH technical specifications and data model requirements
- Portuguese water quality monitoring standards (WFD implementation)
- European Water Framework Directive requirements
- Laboratory accreditation standards (ISO 17025)

Updates

This dataset should be updated when:

- New parameters are added to SNIRH database
- Laboratory methods change, requiring new unit conversions
- SNIRH symbols or naming conventions are updated
- New sample types are introduced to the monitoring program

Source

- SNIRH database documentation: `https://snirh.apambiente.pt`
- Portuguese Environment Agency (APA) technical specifications
- Laboratory analytical method specifications
- Water Framework Directive monitoring requirements

References

- APA (2023). Critérios para a monitorização das massas de água. `https://apambiente.pt/sites/default/files/_SNIAMB_Agua/DRH/PlaneamentoOrdenamento/PGRH/2022-2027/PGRH_3_PTCONT_Monitorizacao.pdf`
- APA (2023). Critérios para a classificação das massas de água. `https://apambiente.pt/sites/default/files/_SNIAMB_Agua/DRH/PlaneamentoOrdenamento/PGRH/2022-2027/PGRH_3_PTCONT_SistemasClassificacao.pdf`
- European Commission (2000). Water Framework Directive 2000/60/EC
- ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories

See Also

`convert_to_snirh` for the main conversion function that uses this data

Examples

```
# View all available parameters for water samples
water_params <- parameters[sample_type == "water"]
print(water_params[, .(param_lab, unit_lab, param_snirh, unit_snirh)])

# Check conversion factor for a specific parameter
ph_conversion <- parameters[param_lab == "pH" & sample_type == "water"]
print(ph_conversion$factor)  # Should be 1 (no conversion needed)

# Find all parameters that require unit conversion
converted_params <- parameters[factor != 1]
print(converted_params[, .(param_lab, unit_lab, unit_snirh, factor)])

# Get SNIRH symbols for biota parameters
biota_symbols <- parameters[sample_type == "biota", unique(symbol_snirh)]
print(biota_symbols)
```

Index