

# Package ‘rsinaica’

February 4, 2019

**Type** Package

**Title** Download Data from Mexico's Air Quality Information System

**Version** 0.6.1

**Description** Easy-to-use functions for downloading air quality data from the Mexican National Air Quality Information System (SINAICA). Allows you to query pollution and meteorological parameters from more than a hundred monitoring stations located throughout Mexico. See <<https://sinaica.inecc.gob.mx>> for more information.

**License** MIT + file LICENSE

**URL** <https://hoyodesmog.diegovalle.net/rsinaica/>,  
<https://github.com/diegovalle/rsinaica>

**BugReports** <https://github.com/diegovalle/rsinaica/issues>

**Encoding** UTF-8

**LazyData** TRUE

**Depends** R (>= 3.2)

**Suggests** aire.zmvm, testthat, covr, knitr, rmarkdown, ggplot2, maps, mapproj, sp, ggmap, gstat, zoo, tidy

**Imports** jsonlite, httr, stringr, dplyr, lubridate

**RoxygenNote** 6.1.0

**NeedsCompilation** no

**Author** Diego Valle-Jones [aut, cre]

**Maintainer** Diego Valle-Jones <[diego@diegovalle.net](mailto:diego@diegovalle.net)>

**Repository** CRAN

**Date/Publication** 2019-02-04 21:10:03 UTC

## R topics documented:

params_sinaica	2
rsinaica	2

sinaica_param_data . . . . .	3
sinaica_station_data . . . . .	4
sinaica_station_dates . . . . .	6
sinaica_station_params . . . . .	7
stations_sinaica . . . . .	8

<b>Index</b>	<b>10</b>
--------------	-----------

---

params_sinaica	<i>Valid air quality parameters</i>
----------------	-------------------------------------

---

### Description

Valid air quality parameters

### Format

A data frame with 55 rows and 2 variables:

**param\_code** Abbreviation of the air quality parameter

**param\_name** Name of the air quality parameter

### Source

**SINAICA**

### Examples

```
head(params_sinaica)
```

---

rsinaica	<i>rsinaica package</i>
----------	-------------------------

---

### Description

Ready-made functions for downloading data from the Mexican National Air Quality Information System (SINAICA). See <https://sinaica.inecc.gob.mx> for more information.

### Details

See the README on [GitHub](#)

---

sinaica\_param\_data     *Get air quality data from all stations by parameter*

---

### Description

Download data from all stations for a single parameter by specifying a date range

### Usage

```
sinaica_param_data(parameter, start_date, end_date, type = "Crude",
  remove_extremes = FALSE)
```

### Arguments

parameter	type of parameter to download
	<ul style="list-style-type: none"> <li>• "BEN" - Benceno</li> <li>• "CH4" - Metano</li> <li>• "CN" - Carbono negro</li> <li>• "CO" - Monóxido de carbono</li> <li>• "CO2" - Dióxido de carbono</li> <li>• "DV" - Dirección del viento</li> <li>• "H2S" - Acido Sulfhídrico</li> <li>• "HCNM" - Hidrocarburos no metánicos</li> <li>• "HCT" - Hidrocarburos Totales</li> <li>• "HR" - Humedad relativa</li> <li>• "HRI" - Humedad relativa interior</li> <li>• "IUV" - Índice de radiación ultravioleta</li> <li>• "NO" - Óxido nítrico</li> <li>• "NO2" - Dióxido de nitrógeno</li> <li>• "NOx" - Óxidos de nitrógeno</li> <li>• "O3" - Ozono</li> <li>• "PB" - Presión Barométrica</li> <li>• "PM10" - Partículas menores a 10 micras</li> <li>• "PM2.5" - Partículas menores a 2.5 micras</li> <li>• "PP" - Precipitación pluvial</li> <li>• "PST" - Partículas Suspendidas totales</li> <li>• "RS" - Radiación solar</li> <li>• "SO2" - Dióxido de azufre</li> <li>• "TMP" - Temperatura</li> <li>• "TMPI" - Temperatura interior</li> <li>• "UVA" - Radiación ultravioleta A</li> <li>• "VV" - Radiación ultravioleta B</li> </ul>

	<ul style="list-style-type: none"> <li>• "XIL" - Xileno</li> </ul>
start_date	start of range in YYYY-MM-DD format
end_date	end of range from which to download data in YYYY-MM-DD format
type	The type of data to download. One of the following: <ul style="list-style-type: none"> <li>• "Crude" - Crude data that has not been validated</li> <li>• "Manual" - Manually collected data that is sent to an external lab for analysis (may no be collected daily). Mostly used for suspend particles collected by pushing air through a filter which is later sent to a lab to be weighted</li> </ul>
remove_extremes	whether to remove extreme values. For O3 all values above .2 are set to NA, for PM10 those above 600, for PM2.5 above 175, for NO2 above .21, for SO2 above .2, and for CO above 15. This is done so that the values match exactly those of the SINAICA website, but it is recommended that you use a more complicated statistical procedure to remove outliers.

### Value

data.frame with a column named *value* containing the air quality parameter values. If the data was validated the column named *date\_validated* will contain the validation date. Care should be taken when working with hourly data since each station has their own timezone (available in the [stations\\_sinaica](#) data.frame) and some stations reported the timezone in which they are located erroneously.

### Examples

```
## Not run:
## May take several seconds
df <- sinaica_param_data("O3", "2015-10-14", "2015-10-14")
head(df)

## End(Not run)
```

---

sinaica\_station\_data *Get air quality data from a single measuring station*

---

### Description

Download data from a single station by specifying a parameter and a date range

### Usage

```
sinaica_station_data(station_id, parameter, start_date, end_date,
  type = "Crude", remove_extremes = FALSE)
```

**Arguments**

station_id	the numeric code corresponding to each station. See <a href="#">stations_sinaica</a> for a list of stations and their ids.
parameter	<p>type of parameter to download</p> <ul style="list-style-type: none"> <li>• "BEN" - Benceno</li> <li>• "CH4" - Metano</li> <li>• "CN" - Carbono negro</li> <li>• "CO" - Monóxido de carbono</li> <li>• "CO2" - Dióxido de carbono</li> <li>• "DV" - Dirección del viento</li> <li>• "H2S" - Acido Sulfhídrico</li> <li>• "HCNM" - Hidrocarburos no metánicos</li> <li>• "HCT" - Hidrocarburos Totales</li> <li>• "HR" - Humedad relativa</li> <li>• "HRI" - Humedad relativa interior</li> <li>• "IUV" - Índice de radiación ultravioleta</li> <li>• "NO" - Óxido nítrico</li> <li>• "NO2" - Dióxido de nitrógeno</li> <li>• "NOx" - Óxidos de nitrógeno</li> <li>• "O3" - Ozono</li> <li>• "PB" - Presión Barométrica</li> <li>• "PM10" - Partículas menores a 10 micras</li> <li>• "PM2.5" - Partículas menores a 2.5 micras</li> <li>• "PP" - Precipitación pluvial</li> <li>• "PST" - Partículas Suspendidas totales</li> <li>• "RS" - Radiación solar</li> <li>• "SO2" - Dióxido de azufre</li> <li>• "TMP" - Temperatura</li> <li>• "TMPI" - Temperatura interior</li> <li>• "UVA" - Radiación ultravioleta A</li> <li>• "VV" - Radiación ultravioleta B</li> <li>• "XIL" - Xileno</li> </ul>
start_date	start of range in YYYY-MM-DD format
end_date	end of range from which to download data in YYYY-MM-DD format
type	<p>The type of data to download. One of the following:</p> <ul style="list-style-type: none"> <li>• "Crude" - Crude data that has not been validated</li> <li>• "Validated" - data which has undergone a validation process during which it was cleaned, verified, and validated</li> <li>• "Manual" - Manually collected data that is sent to an external lab for analysis (may no be collected daily). Mostly used for suspend particles collected by pushing air through a filter which is later sent to a lab to be weighted</li> </ul>

**remove\_extremes**

whether to remove extreme values. For O3 all values above .2 are set to NA, for PM10 those above 600, for PM2.5 above 175, for NO2 above .21, for SO2 above .2, and for CO above 15. This is done so that the values match exactly those of the SINAICA website, but it is recommended that you use a more complicated statistical procedure to remove outliers.

**Value**

data.frame with air quality data. Care should be taken when working with hourly data since each station has their own timezone (available in the `stations_sinaica` data.frame) and some stations reported the timezone in which they are located erroneously.

**See Also**

Crude data comes from <https://sinaica.inecc.gob.mx/data.php>, validated data from <https://sinaica.inecc.gob.mx/data.php?tipo=V>, and manual data from <https://sinaica.inecc.gob.mx/data.php?tipo=M>

**Examples**

```
stations_sinaica[which(stations_sinaica$station_name == "Xalostoc"), 1:5]
df <- sinaica_station_data(271, "O3", "2015-09-11", "2015-09-11", "Crude")
head(df)
```

---

`sinaica_station_dates` *Dates supported by a station*

---

**Description**

Start date and end date of the range for which SINAICA has data for an air quality station

**Usage**

```
sinaica_station_dates(station_id, type = "Crude")
```

**Arguments**

- |                         |   |
|-------------------------|---|
| <code>station_id</code> | the numeric code corresponding to each station. See <code>stations_sinaica</code> for a list of stations and their ids.   |
| <code>type</code>       | The type of data to download. One of the following: <ul style="list-style-type: none"> <li>"Crude" - Crude data that has not been validated</li> <li>"Validated" - Validated data (may not be the most up-to-date)</li> <li>"Manual" - Manual data</li> </ul> |

**Value**

a vector containing the date the station started reporting and end reporting date

**Examples**

```
## id 271 is Xalostoc. See `stations_sinaica`  
df <- sinaica_station_dates(271, "Manual")  
head(df)
```

---

sinaica\_station\_params

*Parameters supported by a station*

---

**Description**

List of air quality parameters of a measuring station for which SINAICA has data

**Usage**

```
sinaica_station_params(station_id, type = "Crude")
```

**Arguments**

station_id	the numeric code corresponding to each station. See <a href="#">stations_sinaica</a> for a list of stations and their ids.
type	The type of data to download. One of the following: <ul style="list-style-type: none"><li>• "Crude" - Crude data that has not been validated</li><li>• "Validated" - Validated data (may not be the most up-to-date)</li><li>• "Manual" - Manual data</li></ul>

**Value**

a data.frame with the parameters supported by the station

**Examples**

```
## id 271 is Xalostoc. See `stations_sinaica`  
df <- sinaica_station_params(271, "Crude")  
head(df)
```

---

stations\_sinaica      *Air quality measuring stations in Mexico*

---

### Description

This data set contains all the stations that report to the National Air Quality Information System **SINAICA**.

### Usage

```
data(stations_sinaica)
```

### Format

A data frame with 341 rows and 26 variables:

**station\_id** Numeric code of the station  
**station\_name** Name of the station  
**station\_code** Abbreviation of the station  
**network\_id** Numeric code for the network  
**network\_name** Name of the network  
**network\_code** Abbreviation of the network  
**street** street  
**ext** exterior number  
**interior** interior number  
**colonia** colonia  
**zip** zip code  
**state\_code** state code  
**municipio\_code** municipio code  
**year\_started** date the station started operations  
**altitude** altitude in meters  
**address** address  
**date\_validated** last date the station was validated  
**date\_validated2** second to last date the station was validated  
**passed\_validation** did the station pass validation  
**video** link to video of the station  
**lat** latitude  
**lon** longitude  
**date\_started** date the station started operations  
**timezone** time zone in which the station is located (may contain errors)  
**street\_view** link to Google Street View  
**video\_interior** link to video of the interior of the station



*stations\_sinaica*

9

**Source**

**SINAICA** ans Solicitud de Información 1612100005118

**Examples**

```
head(stations_sinaica)
```

# Index

[params\\_sinaica](#), [2](#)

[rsinaica](#), [2](#)

[rsinaica-package \(rsinaica\)](#), [2](#)

[sinaica\\_param\\_data](#), [3](#)

[sinaica\\_station\\_data](#), [4](#)

[sinaica\\_station\\_dates](#), [6](#)

[sinaica\\_station\\_params](#), [7](#)

[stations\\_sinaica](#), [4-7](#), [8](#)