

# Package ‘sooty’

December 6, 2024

**Title** Data Source Catalogues Online for Southern Ocean Ecosystem Research

**Version** 0.1.0

**Description** Obtains lists of files of remote sensing collections for Southern Ocean surface properties. Commonly used data sources of sea surface temperature, sea ice concentration, and altimetry products such as sea surface height and sea surface currents are cached in object storage on the Pawsey Supercomputing Research Centre facility. Functions to retrieve these object storage catalogues are provided. The catalogues include complete collections of datasets Reynolds et al. (2008) ``NOAA Optimum Interpolation Sea Surface Temperature (OISST) Analysis, Version 2.1" <doi:10.7289/V5SQ8XB5>, Spreen et al. (2008) ``Artist Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) sea ice concentration" <doi:10.1029/2005JC003384>. In future releases helpers will be added to identify particular data collections and target specific dates for earth observation data for reading, as well as helpers to retrieve data set citation and provenance details.

**License** MIT + file LICENSE

**Encoding** UTF-8

**Language** en-US

**RoxygenNote** 7.3.2

**Imports** arrow, cachem, dplyr, memoise, curl

**URL** <https://github.com/mdsumner/sooty>

**BugReports** <https://github.com/mdsumner/sooty/issues>

**Suggests** spelling, testthat (>= 3.0.0)

**Config/testthat/edition** 3

**NeedsCompilation** no

**Author** Michael D. Sumner [aut, cre]

**Maintainer** Michael D. Sumner <michael.sumner@aad.gov.au>

**Repository** CRAN

**Date/Publication** 2024-12-06 09:00:02 UTC

## Contents

sooty\_files . . . . . 2

**Index** 3

---

sooty_files	<i>Obtain object storage catalogues as a dataframe of file/object identifiers.</i>
-------------	--

---

### Description

The object (file) catalogue of available sources is stored in Parquet format on Pawsey object storage. This function retrieves the curated catalogue, or the raw catalogue.

### Usage

```
sooty_files(curated = TRUE)
```

### Arguments

curated            logical TRUE by default, set to FALSE to return raw object catalogue

### Details

In the curated case, the returned data frame has columns 'date', 'source' which are the main useful fields, these describe the date of the data in the file, and its full URI (Uniform Resource Identifier) source on S3 object storage. There are also fields 'Bucket', 'Key', and 'protocol' from which 'source' is constructed.

The original publisher URI can be reconstructed by replacing the value of 'protocol' in 'source' with 'https://'.

The public object URI can be reconstructed by replacing the value of 'protocol' in 'source' with 'https://projects.pawsey.org.au'.

### Value

a data frame, see details

### Examples

```
if (interactive()) {
  sooty_files(FALSE)
}
```

```
sooty_files()
```

# Index

sooty\_files, [2](#)