

SILO – Getting started guide

December 2018, version 1.1

Prepared by

Science Information Services
Science Delivery Division
Department of Environment and Science
PO Box 2454
Brisbane QLD 4001

© The State of Queensland (Department of Environment and Science) 2018.

The Queensland Government supports and encourages the dissemination and exchange of its information. The copyright in this publication is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence.



Under this licence you are free, without having to seek permission from DSITI, to use this publication in accordance with the licence terms.

You must keep intact the copyright notice and attribute the State of Queensland, Department of Environment and Science as the source of the publication.

For more information on this licence visit <https://creativecommons.org/licenses/by/4.0/>.

Disclaimer

This document has been prepared with all due diligence and care, based on the best available information at the time of publication. The department holds no responsibility for any errors or omissions within this document. Any decisions made by other parties based on this document are solely the responsibility of those parties. Information contained in this document is from a number of sources and, as such, does not necessarily represent government or departmental policy.

If you need to access this document in a language other than English, please call the Translating and Interpreting Service (TIS National) on 131 450 and ask them to telephone Library Services on +61 7 3170 5725.

Acknowledgement

This report has been prepared by the Queensland Government Department of Environment and Science. SILO datasets are constructed using observed data provided by the Australian Bureau of Meteorology.

December 2018

Contents

Introduction.....	4
Access	4
What data are available	5
Requesting point data using the web interface	7
Managing your point data requests.....	9
How to get help	10
Advanced topics	11
How to request point data using the web API	11
How to request gridded data using the web interface	14
How to request gridded data directly from the repository	15
How to work with NetCDF files	15
Hints on viewing NetCDF datasets	16

Introduction

SILO is a database of historical climate data for Australia. It is maintained by the Queensland Government and the datasets are constructed using observed data provided by the Australian Bureau of Meteorology.

SILO datasets are complete, with data from 1-Jan-1889 to yesterday, providing spatial and temporal coverage with no missing data; gaps due to missing observations are filled with data estimated from the nearest locations with observations.

The SILO home page is: <https://silo.longpaddock.qld.gov.au>

This document is to help you get started using SILO. However you should also read the technical documentation on our home page to ensure you understand the limitations of the data.

Access

All SILO data are freely available and licenced under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

To access SILO data you either need:

1. An account at an AAF¹-affiliated institution (e.g. a university or CSIRO)

- You don't need a SILO account
- Simply log in using your username/password from your home institution.

Important: you cannot log in via AAF if you have previously created a SILO account with the same email address that is associated with your AAF credentials.

or

2. A SILO account

- To create an account, go to the login page and click Register (in the top right corner)
- Your email address is used as your account name.

¹ The [Australian Access Federation](#) is an identity broker. If you have an account at an institution that is an AAF subscriber, you may be able to access services provided by other AAF member organisations by using your existing credentials from your home institution. In other words, you don't need to create a new account to access the service provided by the other organisation; you simply log in using your username/password from your home institution.

The login page is at: <https://silو.longpaddock.qld.gov.au/login>. It looks like this:

The screenshot shows the SILO login interface. At the top, the Queensland Government logo is on the left, and 'Help' and 'Contact us' links are on the right. A search bar is also present. Below this is a teal banner with the text 'SILO: Scientific information for land owners' and 'Log in | Register' links. The main content area is titled 'Log in to SILO'. It contains a login form with 'Email' and 'Password' input fields, a 'Log in' button, and links for 'Forgotten your password?' and 'Need an account?'. To the right of the form, there is a 'Continue with AUSTRALIAN ACCESS FEDERATION' button and a note: 'SILO will store your name and email address provided by AAF.'

What data are available

1. SILO provides two types of data:
 - Point sets e.g. data for every day from 1889 to yesterday at a single location, which is either:
 - One of the 18,700 locations in Australia where observations were made; or
 - A pixel from our gridded datasets (see below).
 - Gridded datasets on a $0.05^\circ \times 0.05^\circ$ grid across Australia.

Point datasets: time-series of data for single locations. Point datasets are available at:

- Bureau of Meteorology observing station locations: these datasets consist of observed data when available, and supplemented with interpolated estimates or long-term averages when observed data are either missing or considered suspect. Data are available for approximately 18,700 stations (nearly every station which has ever recorded data at any time in the past or present). These datasets were formerly known as *Patched Point* datasets.
- grid points: these datasets consist entirely of interpolated data and are extracted from the gridded datasets. Data are available at every grid cell in our gridded datasets (approximately 290,000 cells). These datasets were formerly known as *Data Drill* datasets.

Point datasets are available in several formats:

- a range of predefined or “standard” formats tailored for use in decision support tools such as APSIM. These files are plain ASCII text files.
- CSV (comma separated value) text files (commonly used in spreadsheet applications)
- JSON formatted text files.

The CSV and JSON formats only contain the variables selected by the user.

Gridded datasets: spatial datasets spanning continental Australia and some islands:

- The grids are generated by interpolating observations from the Bureau of Meteorology. Gridded data are useful for locations where there are no observations, or for viewing the spatial distribution of a given variable.
- The grids have a spatial resolution of $0.05^\circ \times 0.05^\circ$ and are stored and supplied in NetCDF format. Please see our web page for [examples](#) demonstrating how NetCDF files can be analysed and converted into other formats.

2. Data are available for the following variables:

1. Monthly time-step: monthly rainfall

2. Daily time-step:

- daily rainfall
- temperature (maximum and minimum)
- radiation (solar exposure, consisting of both direct and diffuse components)
- mean sea level pressure (9 am)
- vapour pressure (9 am)
- vapour pressure deficit (9 am)
- evapotranspiration:
 - FAO56 (Penman-Monteith) short crop estimate
 - ASCE tall crop estimate
 - Morton's areal actual estimate
 - Morton's potential estimate
 - Morton's wet-environment estimate.
- evaporation
 - observed class A pan
 - synthetic estimate
 - combination (synthetic estimate pre-1970, class A pan 1970 onwards)
 - Morton's shallow lake estimate.
- relative humidity at maximum and minimum temperatures.

For further information, see the various topics under the “[About](#)” tab in the menu bar on any page, and also the [Frequently asked questions](#) page.


Requesting point data using the web interface

1. Log in to SILO
2. Click on the “[Request point data](#)” tab in the menu bar on any page:

The screenshot shows the SILO web interface. At the top, there is a navigation bar with links for 'Home', 'About', 'Request point data', 'Gridded data', and 'API'. A search bar is also present. Below the navigation bar, the 'Request point data' section is active. It includes a 'Date range' section with 'Start date' and 'End date' input fields. A 'Data format' section allows users to choose between 'Standard' and 'Custom' formats, with a list of options including 'Standard', 'All data', 'APSIM', 'PS1', 'RAINMAN', 'CenW Forest Growth Model', 'Monthly', and 'Century'. A 'Filename prefix' section has a text input field. A 'Submit request' button is at the bottom left. On the right, there is a 'Selected locations' section with a search bar and a map of Australia showing various locations marked with circles and numbers. A legend indicates that blue circles represent 'Stations' and white dots represent 'Grid points'.

When the page loads it displays circles indicating the number of stations in rough geographic regions (as above).

3. Select one or more stations or grid points you want data for (up to 15 can be selected). Stations can be selected:
 - using the map, or
 - by entering a latitude/longitude pair, or searching for stations by name or number.

Markers  indicate station locations and white dots show the location of grid points, as shown below:



4. Enter the date range that you want data for:

- Dates can be entered as either YYYYMMDD, YYYY-MM-DD, or YYYY/MM/DD or by selecting the date using the popup calendar.
5. Data can be requested in either:
 - “standard” format (you can choose from a range of pre-defined formats)
 - custom format (you select the variables of interest and the format – either JSON or CSV).
 6. Provide a name for the file you will receive (optional)
 7. Click Submit
 8. When the requested data are available for collection, you will receive an email containing a link for downloading the data. Usually this will be after a few minutes but may be longer if there is a high system load.

Please note there is a limit on the number of locations that can be submitted in a single data request. If you need to download large volumes of data, please consider using our [API](#).

Managing your point data requests

1. You can view your point data requests by clicking on the “[My requests](#)” tab in the menu bar (at the far right) on most pages:

SILO: Scientific information for land owners

Help Contact us Search website

Home About Request point data Gridded data API My requests

Home / My requests

Current cycle overview

2018-07-01 next reset date	19 days left	6.0GB remaining	0.0GB used	1 web requests made	0 web API requests made
--------------------------------------	------------------------	---------------------------	----------------------	-------------------------------	-----------------------------------

Requests

Note that API requests are not displayed on this page.

1	<p>Reference ID: 49BMDKIET4YFQ</p> <p>Request date: 2018-06-11</p> <p>Expiry date: 2018-06-17</p>	<p>File name: silo_20180611031543897174.zip</p> <p>File size: 4.80 KB</p>	<p>download</p> <p>Show details</p>
---	---	---	-------------------------------------

2. Each user is allocated a download quota, which is the maximum volume of data that can be downloaded in a single calendar month. The quota is reset on the first day of each month.
NOTE: users may download an unlimited volume of gridded data (i.e. gridded datasets are not included in the quota).
3. The “My requests” page shows:
 - The number of point data requests made via the web interface
 - The number of point data requests made via the web API
 - The volume of data downloaded in the current month, the remaining volume of data which can still be downloaded in the current month, and when the quota will be reset
 - Detailed information about requests made via the web interface.

The “My Requests” page does not show:

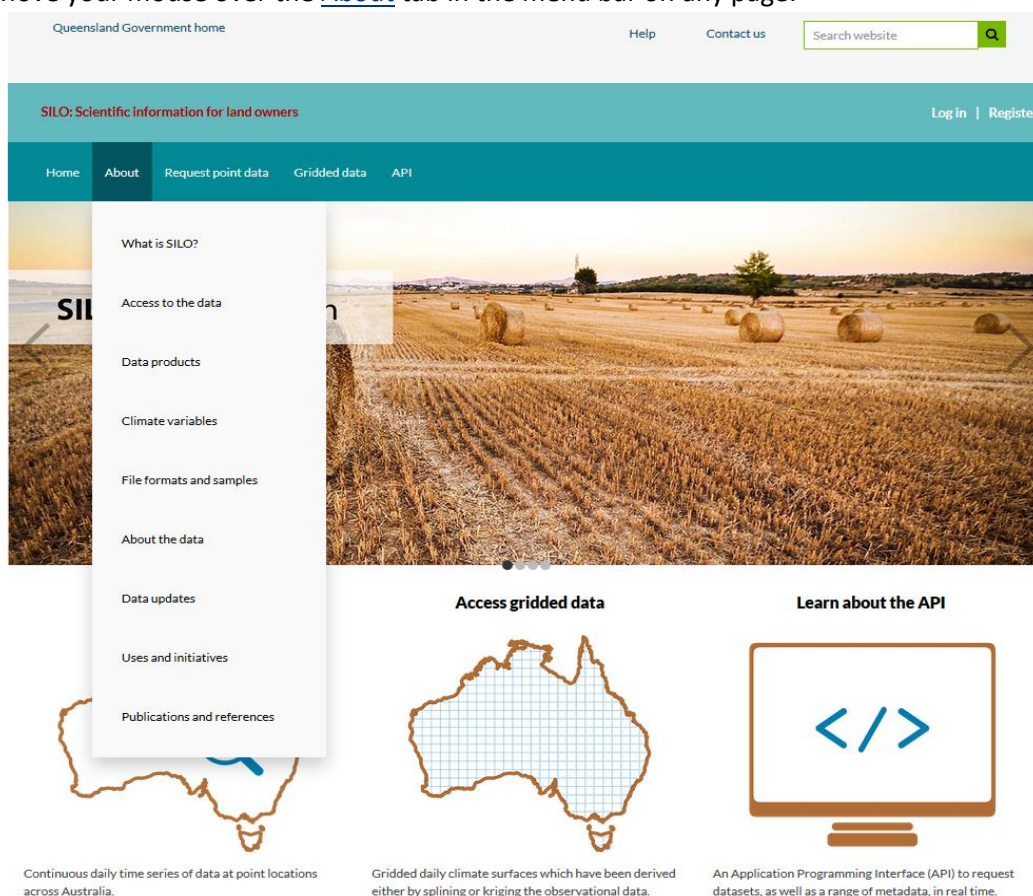
- details of point data requests made via the web API
- any information about gridded data requests.

4. Each request has a status:
 - Download – the requested dataset is available for download
 - Expired - the requested dataset is no longer available. Data are held for a period of 7 days, after which the dataset is deleted and no longer available for download.
 - Error – the requested data could not be delivered. This may be due to a system malfunction or invalid parameters in the user request.

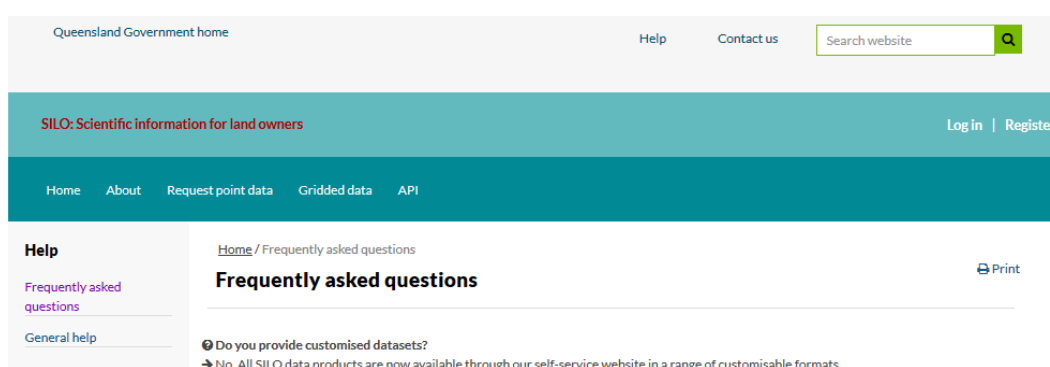
How to get help

To get help:

1. Read the “About” information
 - Move your mouse over the [About](#) tab in the menu bar on any page:



2. Check the Frequently asked questions
 - Click on [Help](#) at the top of any page



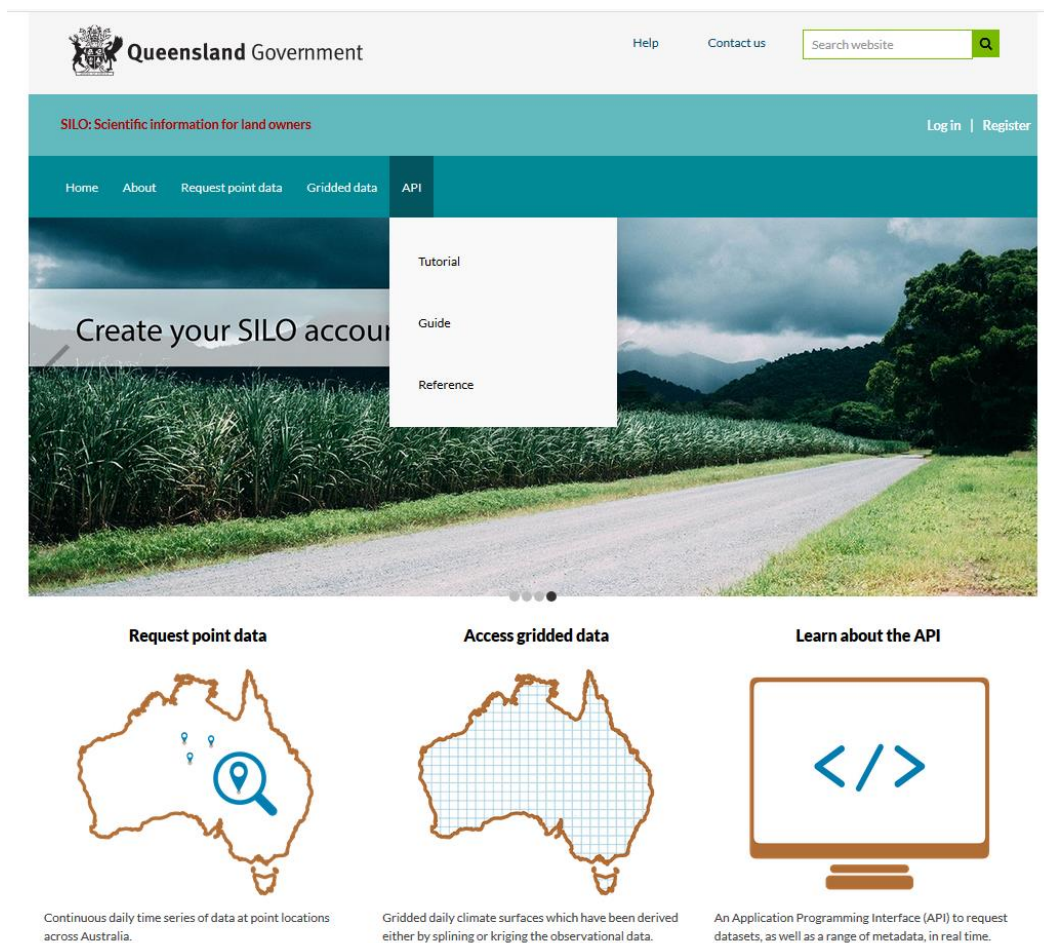
3. Read the metadata
 - Detailed metadata are provided on the Queensland Spatial Catalogue at <http://qldspatial.information.qld.gov.au/catalogueadmin> (search for “SILO”).
4. Contact us
 - To email SILO staff, click on [Contact us](#) at the top of any page
 - Please note: SILO no longer has the resources to prepare customised datasets for individual clients.

Advanced topics

This section covers getting data via our Application Programmers Interface (API) and accessing gridded data.

How to request point data using the web API

1. Read the API information
 - Move your mouse over the API tab in the menu bar on any page:



2. To use the API, you need to create an API key:
 - Log in to SILO
 - Go to the account settings drop-down menu (see the ⚙ icon in the menu bar on any page) and select “[My API keys](#)” to generate a key:

Note: you only need to generate a key once. Keys don't expire, but if you wish to split your requests up (e.g. for different projects), you can have up to four keys.

3. Encode your data request in a URL of the form:

<https://siloapi.longpaddock.qld.gov.au/pointdata?apikey=<your API key>&start=<startdate>&finish=<enddate>&format=<your format>&variables=<your variables>>

Notes:

- please see the API [tutorial](#), [guide](#) and [reference](#) for full details of the URL format
- the API can be used to query metadata (such as source codes, suppliers and variable names); search for stations by name and location; and view statistics about the amount of observed data available at a given station.

4. The data can be retrieved:

- in your web browser: simply paste the URL-encoded request into the navigation bar

Note: your browser may not display the line breaks, so you may need to right-click and select "View source".

- using command line tools, such as [wget](#) or [curl](#)
- directly by your application.

Example 1. Request data using FTP command line tools in UNIX bash:

```
#!/bin/bash
api_url="https://siloapi.longpaddock.qld.gov.au"
begin="20150101"
end="20150105"
station="1018"
lat="-30.0"
lon="135.0"
my_key="insert your API key here"

# Find stations near a given station in CSV format
curl "$api_url/stations?near=$station&radius=50&format=csv"

# Get data statistics for a given station in JSON format
curl "$api_url/stations/$station?format=json&statistics=decadal"
```

```
# List the available variables
curl "$api_url/variables"

# Get point data for one variable in CSV format
curl "$api_url/pointdata?station=$station&apikey=$my_key&start=$begin&finish=$end&format=csv&variables=daily_rain"

# Get point data for multiple variables in JSON format
curl "$api_url/pointdata?lat=$lat&lon=$lon&apikey=$my_key&start=$begin&finish=$end&format=json&variables=max_temp,min_temp"
```

Example 2. Request data in Python

```
import requests, json, csv

# Example 1. Retrieve station data in JSON format and output
params = {
    'apikey': '<insert your_API key here>',
    'format': 'json',
    'station': 1001,
    'start': '20100101',
    'finish': '20100102',
    'variables': 'max_temp,min_temp',
}

r = requests.get('https://siloapi.longpaddock.qld.gov.au/pointdata', params=params)
point_data = r.json()
print json.dumps(point_data, indent=4, sort_keys=True)

# Example 2. Retrieve metadata in CSV format and output
r = requests.get('https://siloapi.longpaddock.qld.gov.au/variables?format=csv')
metadata = r.content.decode('utf-8')
cr = csv.reader(metadata.splitlines(), delimiter=',')
my_list = list(cr)
for row in my_list:
    print(row)
```

How to request gridded data using the web interface

To request a small number of gridded datasets, you can use our web interface:

1. Move your mouse over the “Gridded data” tab in the menu bar on any page and select “[Access gridded data](#)”
2. Enter the year, select the variable and click on “Download” to retrieve the NetCDF file containing the gridded data for a single year and variable.

[Access gridded data](#)
[Python tools](#)
[NetCDF operators](#)
[Convert NetCDF files](#)

Access gridded data

SILO's gridded datasets are hosted on AWS under the AWS Public Data Program.

How can I access the data?


You can download the annual files here:

Year*
Variable*

ⓘ SILO's gridded datasets are arranged in annual blocks. Each annual file contains all of the grids for the selected year and variable.
 ⓘ Each annual file (for daily variables) is approximately 410 MB in size. The annual files for monthly rainfall are smaller (14 MB) because they only contain 12 monthly grids, instead of 365 or 366 daily grids.

How else can I access the data?

The gridded datasets can also be accessed by the following methods:



- Browsing the data collection on the [AWS Public Data website](#).
- Downloading files directly to your application or web browser by specifying the URL. The URL has the form:
`https://s3-ap-southeast-2.amazonaws.com/silo-open-data/annual/<variable_short_name>/<year>.<variable_short_name>.nc`
 and is fully described on the AWS Public Data website. For example, the 2015 data for class A pan evaporation can be downloaded using `curl` as follows:
`curl "https://s3-ap-southeast-2.amazonaws.com/silo-open-data/annual/evap_pan/2015.evap_pan.nc" --remote-name`
 The variable_short_name's are also listed on our [variables](#) page.
- Using the [AWS command line interface](#). For example, the 2005 data for monthly rainfall can be downloaded as follows:
`aws s3 cp s3://silo-open-data/annual/monthly_rain/2005.monthly_rain.nc .`
 Alternatively, all monthly rainfall files can be downloaded to your current directory as follows: `aws s3 sync s3://silo-open-data/annual/monthly_rain/ .`

Note: you do not need to log in to access gridded data.

SILO's gridded datasets are stored in annual blocks. For daily variables, all the daily grids for a given year are stored in a single file, and similarly, all twelve monthly rainfall grids for a given year are stored in a single monthly rainfall file.

How to request gridded data directly from the repository

To request a large number of gridded datasets, we recommend downloading the files directly from the data repository. The gridded datasets are hosted on AWS under the [AWS Public Data Program](#). The NetCDF files are stored in an AWS [S3](#) bucket with a directory and filename structure as follows:

```
<variable>/<year>.<variable>.nc
```

The S3 bucket is publicly accessible and the datasets can therefore be accessed using any of the standard methods for accessing data on S3, such as:

1. via URL:

The URL has the following format:

```
https://s3-ap-southeast-2.amazonaws.com/silo-open-data/annual/<variable>/<year>.<variable>.nc
```

After substituting the `year` and `variable`, datasets can be downloaded:

- in your browser (simply paste the URL into the navigation bar)
- directly to your application (see the web API Python example shown above)
- using command line tools, such as [wget](#) or [curl](#).

For example, to download the 2015 data for class A pan evaporation:

```
curl "https://s3-ap-southeast-2.amazonaws.com/silo-open-data/annual/evap_pan/2015.evap_pan.nc" --remote-name
```

2. using the [AWS Command Line Interface](#):

Example 1. To list the available variables:

```
aws s3 ls s3://silo-open-data/annual/
```

(Note the trailing forward slash)

Example 2. To download the 2012 data for radiation:

```
aws s3 cp s3://silo-open-data/annual/radiation/2012.radiation.nc .
```

(Note the trailing period)

How to work with NetCDF files

There are many open source tools available for viewing and processing NetCDF datasets. We provide an overview of some options on our [Frequently asked questions](#) page.

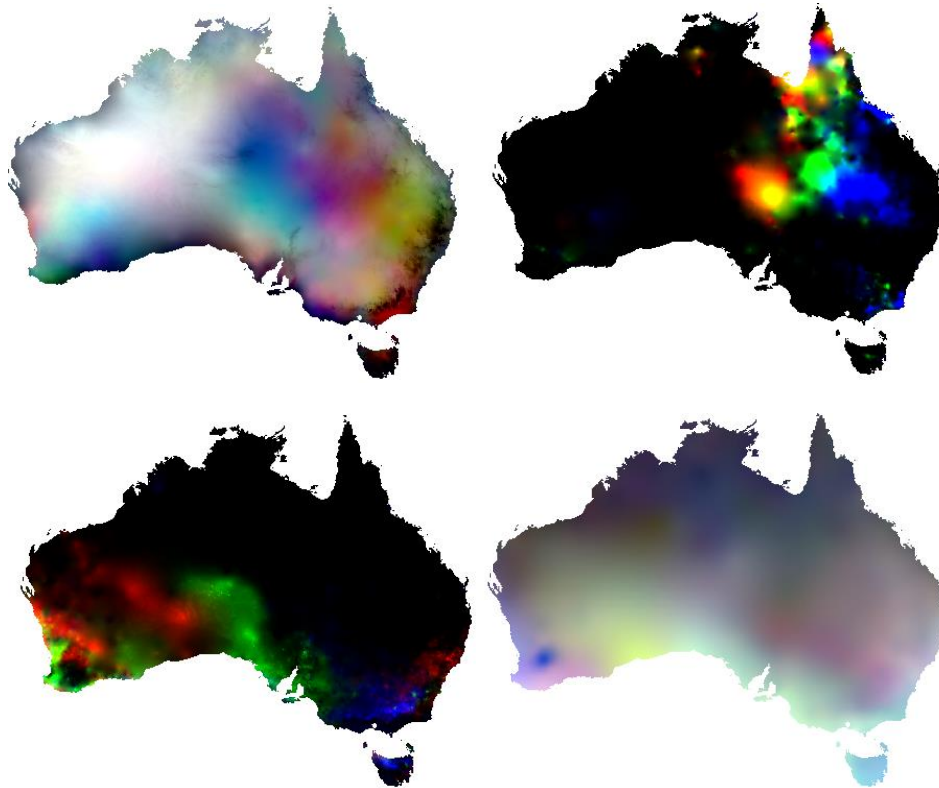
Example scripts showing how to:

1. process NetCDF files using command line tools such as NCO and CDO
2. analyse NetCDF datasets using Python
3. convert NetCDF files to other formats using GDAL

are available on our [gridded data page](#).

Hints on viewing NetCDF datasets

There are many [tools](#) available for manipulating and viewing NetCDF files. If you attempt to view the data in a NetCDF file and see anything like the following:



you may be seeing just the first three bands (e.g. grids for 1-3 January for a daily variable) mapped to the colours Red Green Blue. To view individual bands using some GIS tools you may need to change the properties of the layer from “Multiband colour” to “Singleband grey” or “Singleband pseudocolour”, and then select the band to view.