Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"
Consiglio Nazionale delle Ricerche

ISTI Technical Reports

A tool for the temporal analysis of sea surface temperature maps

Oscar Papini, CNR-ISTI, Pisa, Italy

ISTI-TR-2021/011

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"
Consiglio Nazionale delle Ricerche

A tool for the temporal analysis of sea surface temperature maps

Papini O.

ISTI-TR-2021/011

Abstract
This document describes the usage of a tool that produces plots of the evolution of the sea surface temperature in a specified space-time window, extracting data from a series of NetCDF files.

Sea surface temperature, Image analysis tool

Citation
Papini O., *A tool for the temporal analysis of sea surface temperature maps*, ISTI Technical Reports 2021/011. DOI: 10.32079/ISTI-TR-2021/011

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"
Area della Ricerca CNR di Pisa
Via G. Moruzzi 1
56124 Pisa Italy
http://www.isti.cnr.it

A tool for the temporal analysis of sea surface temperature maps

Oscar Papini

9th September 2021

This document describes the usage of a tool, written in Python 3 (using the `tkinter` module), that produces plots of the evolution of the sea surface temperature (SST) in a specified space-time window, extracting data from a series of NetCDF files, whose internal structure may follow either EUMETSAT's conventions [1] or NASA's ones [2]. In particular, this tool has been used to extract information from SST maps obtained from EUMETSAT's METOP-B satellite (via its AVHRR sensor) and NASA's Aqua satellite (via its MODIS sensor).

1 Main window

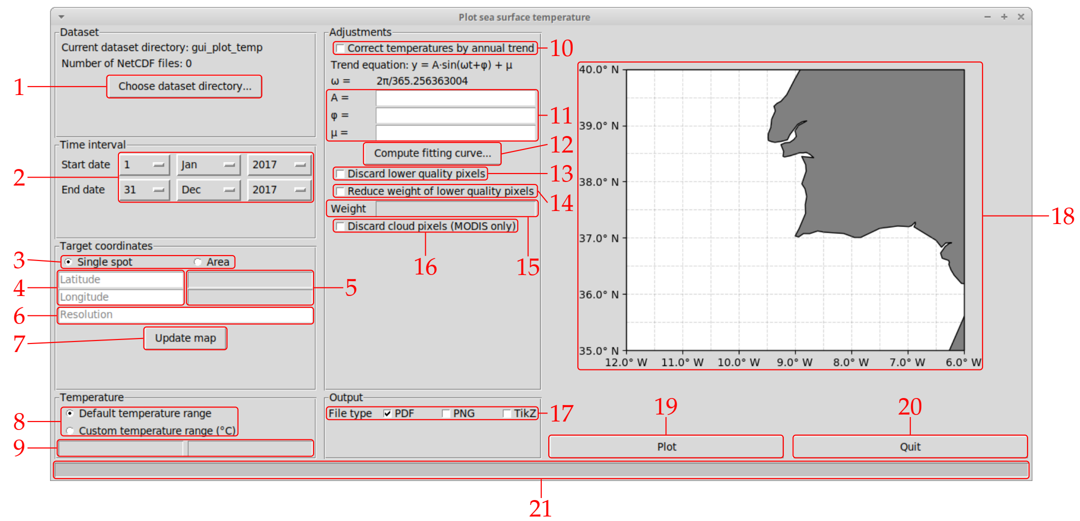


Figure 1: Main window of the GUI.

This section shows the various tools that appear in the main window of the GUI (see Figure 1).

"Dataset" frame. The "Choose dataset directory…" button (1) opens a *Select directory* window. Once chosen, the name of the directory and the number of NetCDF files in it appear in the frame. The default directory is the one that contains the GUI script.

"Time interval" frame. It is possible to select the start date and the end date of the plot (2). At the moment the script does not check coherence (e.g. 31st April is accepted), nor if the NetCDF files are compatible with the dates.

"Target coordinates" frame. The script works in two modes, "Single spot" and "Area", which can be selected with the relative radiobutton (3) (see Figure 2).

(a) "Single spot" mode.

(b) "Area" mode.

Figure 2: Changing mode in the "Target coordinates" frame.

In the "Single spot" mode, the script plots the temperature measured in the point with specified coordinates[1] (4). In particular, if the point has coordinates $(\mathrm{lat}, \mathrm{lon})$, then the temperature for a single NetCDF file is computed by averaging the temperature values contained in the small square

$$\left[\mathrm{lat} - \frac{r}{2}, \mathrm{lat} + \frac{r}{2}\right] \times \left[\mathrm{lon} - \frac{r}{2}, \mathrm{lon} + \frac{r}{2}\right]$$

where r is the specified resolution (6), in degrees.
In the "Area" mode, the script produces a plot (*spaghetti plot*) by superimposing the temperature graphs of the points contained in a specified grid, obtained from the parameters (4)–(6) in the following way: let $(\mathrm{lat}, \mathrm{lon})$ be the minimal latitude/longitude (4), $(\mathrm{LAT}, \mathrm{LON})$ the maximal latitude/longitude (5), and r the resolution (6); then a point in the grid has coordinates (i, j) with

$$i \in \{\mathrm{lat}, \mathrm{lat} + r, \mathrm{lat} + 2r, \ldots, \mathrm{lat} + kr\},$$
$$j \in \{\mathrm{lon}, \mathrm{lon} + r, \mathrm{lon} + 2r, \ldots, \mathrm{lon} + hr\},$$

where k is such that $\mathrm{lat} + (k+1)r \geqslant \mathrm{LAT}$ and h is such that $\mathrm{lon} + (h+1)r \geqslant \mathrm{LON}$. The output contains a graph for each point (i, j) and the temperature for such a point is computed as average over the small square $[i, i+r] \times [j, j+r]$. Figure 3 shows an example of a *spaghetti plot*.

[1]Please note that positive numbers denote N latitude and E longitude, while negative numbers denote S latitude and W longitude; the decimal separator is a dot, not a comma.

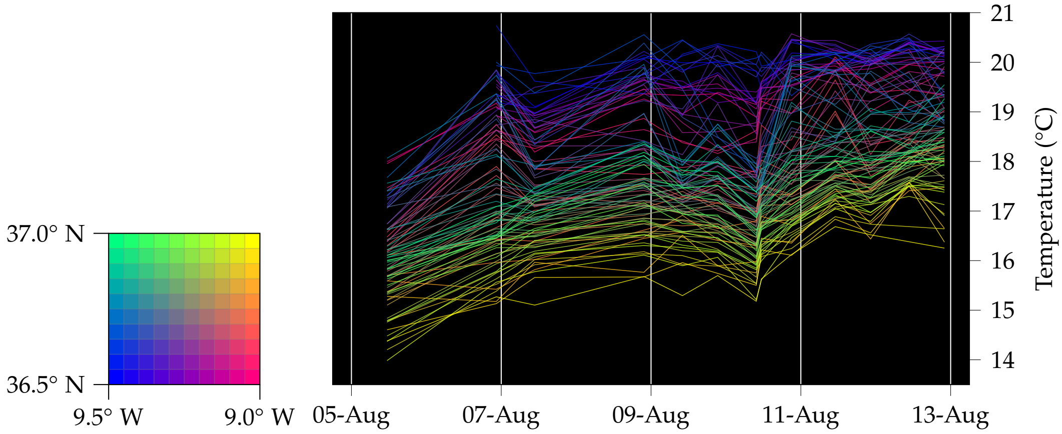


Figure 3: Example of a *spaghetti plot*. Each graph in the plot on the right describes the trend of the SST in the point with the same colour in the reference grid on the left. Here $\text{lat} = 36.5°$ N, $\text{LAT} = 37.0°$ N, $\text{lon} = 9.5°$ W, $\text{LON} = 9.0°$ W and $r = 0.05°$.

Both in "Single spot" and in "Area" modes, parameters are inserted into the script by the button "Update map" (7), which has to be pressed after any modification. Updates will appear in the map (18); in particular, in "Area" mode each point of the grid is coloured following a gradient, and the corresponding graph in the plot will use the same colour, so that it is easy to assign a single graph to a specific point.

"Temperature" frame. With the radiobutton (8), it is possible to let the plotting library decide the most suitable range for the temperatures in the plot, or set a minimal and maximal temperature for the y-axis of the plot (9) (see Figure 4). This second case is useful for direct comparison of different graphs.

(a) Default range. (b) Custom range.

Figure 4: Choosing the temperature range.

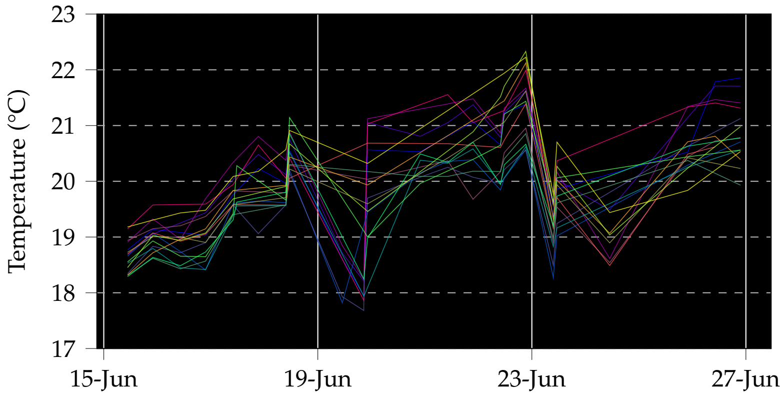
"Adjustments" frame. If (10) is checked, instead of plotting the function with equation $y = g(t)$ that gives the temperature at time t, the script plots the function with equation

$$\tilde{g}(t) = g(t) - \big(T(t) - T(t_0)\big)$$

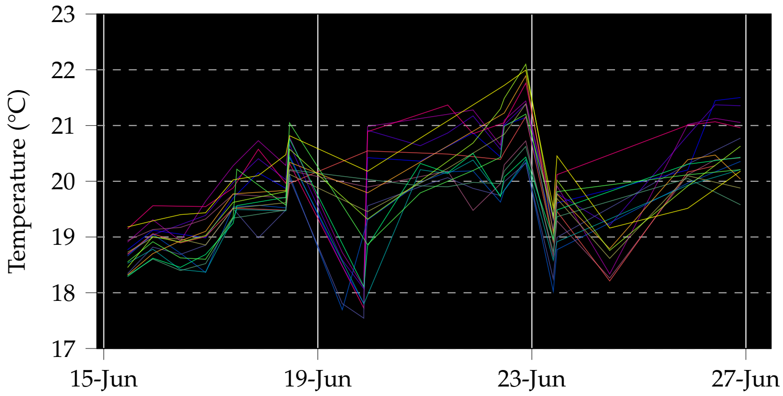
where t_0 is the starting time of the plot and

$$T(t) = A \cdot \sin(\omega t + \varphi) + \mu \qquad (\star)$$

with fixed ω and A, φ, μ specified parameters (11). The function $T(t)$ represents the annual trend of the mean temperature inside a specified window; parameters can be input directly or computed by clicking on the "Compute fitting curve..." button (12) — more on this in Section 2. Figure 5 shows the effect of this adjustment.



(a) Without correction.



(b) With correction.

Figure 5: Example of a plot corrected by the annual trend function $T(t)$ of Figure 7. The target window is $[36.5° \text{ N}, 37.0° \text{ N}] \times [9.5° \text{ W}, 9.0° \text{ W}]$ with $r = 0.125°$ between 15th and 26th June 2017.

If (13) is checked, the script masks the points with lower quality (see Appendix A for a description of the quality levels in the NetCDF files) so that their value is not used during the computation. If (14) is checked, the script computes a *weighted* average of the temperatures instead of a simple average. In particular, it gives the specified weight (15) to the lower quality data (same levels as above) and weight 1 to everything else.

In the files there are other types of flags assigned to each data point. If (16) is checked, the script masks the points marked as “cloud” in an Aqua file,[2] thus excluding them from the computation.

“Output” frame. It is possible to save the plot in a PNG file, a PDF file or a simple text file with extension `.tikz` that contains TikZ code[3] (17); the script will create a file for any selected extension — if nothing is selected here, the plot will be given on-screen. Each run of the script produces two images for each selected format: the plot itself and a reference map. In particular, the reference map in “Area” mode shows only the points of the grid for which a plot is actually computed.[4] File names use the following conventions (the first one is for “Single spot” mode, the second one is for “Area” mode):

```
<start>-<end>(<lat>,<lon>)R<res>_<type>
<start>-<end>[<lat>,<LAT>]x[<lon>,<LON>]R<res>_<type>
```

Here, <start> and <end> are the starting and end dates for the plot, in 8-digit format YYYYMMDD (2); <lat> and <lon> are the latitude/longitude (4) of the point (minimal ones in “Area” mode); <LAT> and <LON> are the maximal latitude/longitude (5) in “Area” mode; <res> is the resolution (6); and <type> is either `plot` for the plot or `map` for the reference map.
If (14) is active, file names contain the additional information of the weight:

```
<start>-<end>(<lat>,<lon>)R<res>W<weight>_<type>
<start>-<end>[<lat>,<LAT>]x[<lon>,<LON>]R<res>W<weight>_<type>
```

where <weight> is the weight (15) given to lower quality points.

Buttons. The “Plot” button (19) starts the plot and the “Quit” button (20) closes the window. During the plot it is possible to keep track of the process thanks to the progress bar (21).

2 Compute fitting curve window

This window (Figure 6) shows up when clicking on the “Compute fitting curve...” button (12) in the main window. It contains the settings used to compute the parameters of the $T(t)$ function ($\star$) introduced in the previous section. This computation is done by the `curve_fit` function of `scipy`’s `optimize` package.

[2] The “cloud” flag is not present in METOP-B files.

[3] This is achieved through the `tikzplotlib` Python module.

[4] At the moment, the graph of a point (i, j) of the grid, corresponding to a small square $Q = [i, i+r] \times [j, j+r]$, is present in the plot only if in at least 75% of the NetCDF files the ratio (valid data points in Q)/(total data points in Q) is at least 20%.

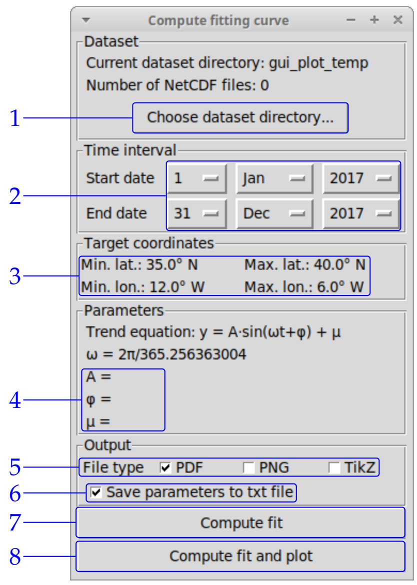


Figure 6: "Compute fitting curve" window.

"Dataset" frame. The "Choose dataset directory. . . " button (1) opens a *Select directory* window. Once chosen, the name of the directory and the number of NetCDF files in it appear in the frame. The default directory is the one that is currently selected in the main window.

"Time interval" frame. It is possible to select the start date and the end date for the fit (2). At the moment the script does not check coherence (e.g. 31st April is accepted), nor if the NetCDF files are compatible with the dates. The more the chosen period is wide, the better the fit will be.

"Target coordinates" frame. The fit is computed by taking, for each NetCDF file, the average over a selected spatial window.[5] *Please note that this computation takes into account any active modifier* (13)–(16) *from the "Adjustements" frame in the main window.* This frame shows the boundaries of the target window, which are taken from the fields (4)–(6) of the main GUI window. In particular, in "Single spot" mode the values are:

$$\text{Min. lat.} = \text{lat} - \frac{r}{2}, \qquad \text{Max. lat.} = \text{lat} + \frac{r}{2},$$
$$\text{Min. lon.} = \text{lon} - \frac{r}{2}, \qquad \text{Max. lon.} = \text{lon} + \frac{r}{2},$$

where lat and lon are taken from (4) and r is the resolution (6), while in "Area" mode the values are taken directly from the corresponding fields (4) and (5).

[5] At the moment, a NetCDF file is discarded from the plot if the ratio (valid data points in the window)/(total data points in the window) is less than 40%.

"Parameters" frame. Here (4) the parameters computed by the script are shown. The period ω is fixed and corresponds to the duration of a sidereal year.

"Output" frame. As in the main window, it is possible to save the plot in a PNG file, a PDF file or a simple text file with TikZ code (5); if nothing is selected here, the plot will be given on-screen. File name uses the "Area" mode convention described before, i.e.

```
<start>-<end>[<lat>,<LAT>]x[<lon>,<LON>]_fit
```

where <start> and <end> are the starting and end dates *for the fit* (2), in 8-digit format YYYYMMDD, and <lat>, <lon> (respectively <LAT>, <LON>) are the minimal (respectively maximal) latitude/longitude (3) of the selected window. It is possible to save the parameters to a TXT file (6) for future use; if the checkbutton is selected, a file named

```
<start>-<end>[<lat>,<LAT>]x[<lon>,<LON>]_parameters.txt
```

(using the same conventions as before) will be produced together with the plot.

Buttons. The "Compute fit" button (7) computes the parameters, updating both (4) and (11) (and removing previously computed parameters, if any), but without producing a plot of the fitting curve (thus ignoring the values of (5)); on the other hand, the "Compute fit and plot" button (8) produces also a scatterplot of the data in the NetCDF files with a plot of the fitting curve (see Figure 7 for an example).

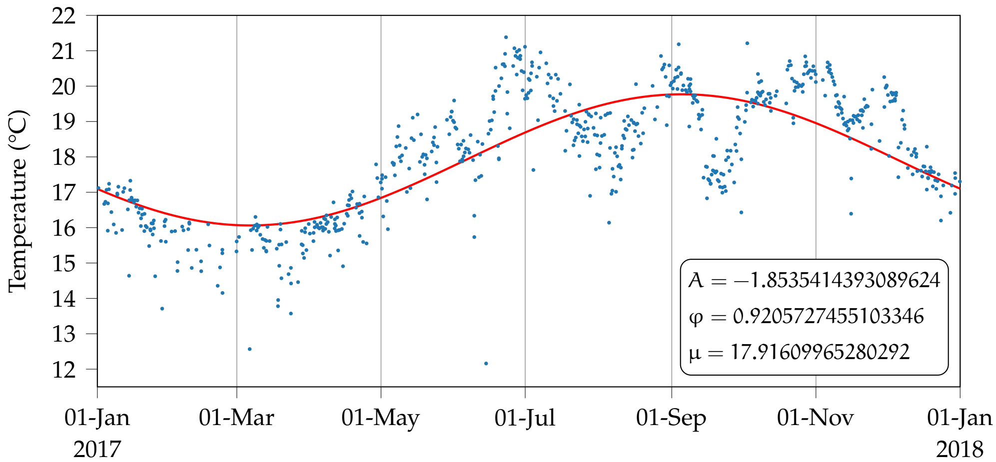


Figure 7: Example of annual trend function $T(t)$ for the year 2017 in the window $[36.5° \text{ N}, 37.0° \text{ N}] \times [9.5° \text{ W}, 9.0° \text{ W}]$. The figure shows also the corresponding parameters.

A Appendix: Quality levels in the NetCDF files

In the NetCDF files, each data point is marked with a flag that describes the quality level of the data in that point (see Table 1 for a list of the quality levels). The tool considers "bad quality" points the ones flagged with levels 0–2 for METOP-B files and levels 3–4 for Aqua files.

Flag	Quality level
0	No data
1	Bad data
2	Worst
3	Low
4	Acceptable
5	Best

(a) Data from METOP-B.

Flag	Quality level
0	Best
1	Good
2	Questionable
3	Bad
4	Not processed

(b) Data from Aqua.

Table 1: Quality levels in NetCDF files.

References

[1] *NetCDF format for Data Centre products*. Accessed on 30th August 2021. URL: `https://www.eumetsat.int/data-centre-netcdf-format`.

[2] *Ocean Level-2 Data Format Specification*. Accessed on 30th August 2021. URL: `https://oceancolor.gsfc.nasa.gov/docs/format/l2nc`.