OH14 v1.2

Downscaled Australian continental shelf sea surface temperature data

User's manual

Eric C. J. Oliver and Neil J. Holbrook

Institute for Marine and Antarctic Studies
University of Tasmania, Hobart TAS, Australia
contact: eric.oliver@utas.edu.au

March 20, 2014

1 Introduction

The Oliver and Holbrook (2014) [1], or OH14, data set provides spatially and temporally homogeneous measurements of sea surface temperature (SST) variability at high resolution (1/10° in space and daily in time) across the western, southern and eastern continental shelves around Australia between 14 October 1992 and 13 May 2008. The data set has been produced using a hybrid statistical-physical downscaling technique designed to more accurately and robustly represent SST on the continental shelf, applied to a combination of large-scale satellite observations and reanalysis data [1]. The downscaled shelf SST is modeled using: (i) offshore SST from Bluelink ReANalysis (BRAN) versions 2.1 and 2.2 [2, 3, 4], (ii) the statistical relationship between inshore and offshore SST in observations from Advanced Very High Resolution Radiometer (AVHRR) Pathfinder Version 5.2 (PFV5.2) data [5], and (iii) the mean circulation which provides connectivity information between the shelf and the offshore regions. The SST time series were separated into the mean, seasonal cycle, and the residual variability, and separate models were developed for each component. Since the analysis technique applied here is linear, the three component parts are readily reconstructed, providing complete SST coverage in space and time, including error estimates.

2 OH14 data file

The data are provided as a single NetCDF file, `OH14.nc`. A description of its contents follows, interspersed with the NetCDF metadata (e.g., accessible via `ncdump -h` in UNIX, `ncdisp` or `ncdump` in MATLAB).

```
Global Attributes:
           title         = 'OH14 downscaled sea surface temperature
               data on the Australian continental shelf'
```

```
          summary       = 'OH14 (Oliver and Holbrook, 2014) downscaled
              continental shelf sea surface temperature (SST) dataset.
              Provided are estimates of SST, mean SST, SST seasonal
             cycle and SST residual (with error estimates) at 1/10
             degree horizontal resolution over the Australian
             continental shelf (south of 20S).'
          version       = 'v1.2'
          creator_name  = 'Eric C. J. Oliver'
          creator_email = 'eric.oliver@utas.edu.au'
          creator_url   = 'http://passage.phys.ocean.dal.ca/~olivere/'
          date          = '2014-Mar-19 17:33:48'
          File_Name     = 'OH14.nc'
          references    = 'Oliver, E. C. J. and N. J. Holbrook (2014),
              A statistical method for improving continental shelf and
              near-shore marine climate predictions, Journal of
             Atmospheric and Oceanic Technology, 31, pp. 216-232, doi:
              10.1175/JTECH-D-13-00052.1'
Dimensions:
          time     = 5691
          lon_map  = 902
          lat_map  = 552
          location = 6064
          ymd      = 3
```

The `time` vector is simply the time vector from BRAN (daily values from 14 Oct 1992 to 13 May 2008), in MATLAB datenum format. The `dates` vector is the time vector expressed as `[year month day]`.

```
    time
          Size:        5691x1
          Dimensions:  time
          Datatype:    double
          Attributes:
                       title = 'Time'
                       units = 'Number of days since Jan-0-0000 (matlab
                           datenum format). For example, January 2,
                          year 0 gives time = 2.'
    dates
          Size:        3x5691
          Dimensions:  ymd,time
          Datatype:    double
          Attributes:
                       title = 'Date in [year month day] format.'
```

The `lat` and `lon` vectors are the latitudes and longitudes (in decimal degrees) of each of the locations at which downscaled shelf data are available.

```
    lon
          Size:        6064x1
          Dimensions:  location
          Datatype:    double
          Attributes:
                       title     = 'longitude of each location'
                       units     = 'decimal degrees longitude'
                       valid_min = 112
                       valid_max = 154
```

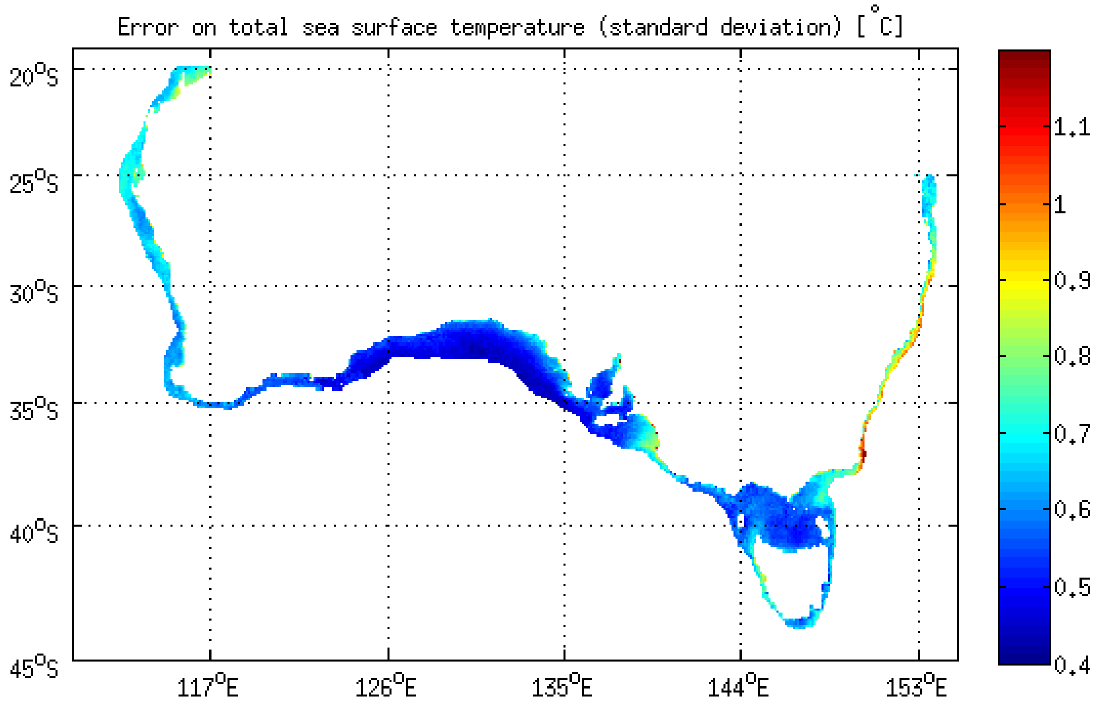

Figure 1: Error (°C) on total sea surface temperature of OH14 data set. Error is provided as a standard deviation (σ) on the error time series at each grid point in space.

```
lat
          Size:         6064x1
          Dimensions:   location
          Datatype:     double
          Attributes:
                        title     = 'latitude of each location'
                        units     = 'decimal degrees latitude'
                        valid_min = -44
                        valid_max = -20.1
```

The following variables provide total SST (`T`), mean SST (`T_mean`), SST seasonal cycle (`T_seas`), residual SST (`T_res`) and error estimates (`*_sige`). Total SST is defined as the sum of the mean, seasonal cycle and residual. Note that mean SST and the error estimates are provided as a single value at each of the shelf locations, while the total SST, SST seasonal cycle, and residual SST are provided as daily time series (14 Oct 1992 to 13 May 2008) at each of the shelf locations. All values are provided in degrees Celcius.

The error estimates are provided as a standard deviation (σ), therefore 67% confidence limits can be calculated from $\pm\sigma$ about the estimated SST and 95% confidence by $\pm 2\sigma$ about the estimated SST. A map of the error distribution for total SST (`T_sige`) across the region of coverage is shown in Figure 1.

These variables are not provided as matrix (i.e., lat$\times$lon$\times$time), since the shape of the domain (the Australian continental shelf) is very irregular, and are simply provided at a list of locations on the shelf (i.e., location$\times$time), with location coordinates identified by the `lat` and `lon` vectors (defined above). A script is provided for mapping these data into matrix form (see Section 3).

These variables are provided as integers, in order to save storage space. To convert back to decimal precision you must first scale by the `scale_factor` and then add the `add_offset`. These values are provided as attributes on the temperature and error variables. Your local NetCDF implementation may or may not automatically apply this conversion at loading time, so it is best to check.

```
T
       Size:        6064x5691
       Dimensions:  location,time
       Datatype:    int16
       Attributes:
                    title        = 'Total sea surface temperature, i
                       .e., the sum of the mean, seasonal cycle, and
                        the residual variability (T_mean + T_Seas +
                       T_res)'
                    units        = 'degrees Celcius'
T_sige
       Size:        6064x1
       Dimensions:  location
       Datatype:    int16
       Attributes:
                    title        = 'Error on total sea surface
                       temperature, given as a standard deviation'
                    units        = 'degrees Celcius'
T_mean
       Size:        6064x1
       Dimensions:  location
       Datatype:    int16
       Attributes:
                    title        = 'Mean sea surface temperature'
                    units        = 'degrees Celcius'
T_seas
       Size:        6064x5691
       Dimensions:  location,time
       Datatype:    int16
       Attributes:
                    title        = 'Seasonal cycle of sea surface
                       temperature'
                    units        = 'degrees Celcius'
T_seas_sige
       Size:        6064x1
       Dimensions:  location
       Datatype:    int16
       Attributes:
                    title        = 'Error on seasonal cycle, given
                       as a standard deviation'
                    units        = 'degrees Celcius'
T_res
       Size:        6064x5691
       Dimensions:  location,time
       Datatype:    int16
       Attributes:
                    title        = 'Residual sea surface temperature
                       '
                    units        = 'degrees Celcius'
```

```
T_res_sige
        Size:        6064x1
        Dimensions:  location
        Datatype:    int16
        Attributes:
                     title        = 'Error on residual, given as a
                        standard deviation'
                     units        = 'degrees Celcius'
```

The vectors `lat_map` and `lon_map` provide latitude and longitude vectors on a 1/10° horizontal resolution grid (the BRAN grid). The variable `shelfmask`, defined on this grid, is a map of the locations at which downscaled shelf data are provided.

```
lon_map
        Size:        902x1
        Dimensions:  lon_map
        Datatype:    double
        Attributes:
                     title     = 'longitude (on 1/10 degree grid, for
                         mapping)'
                     units     = 'decimal degrees longitude'
                     valid_min = 90
                     valid_max = 180
lat_map
        Size:        552x1
        Dimensions:  lat_map
        Datatype:    double
        Attributes:
                     title     = 'latitude (on 1/10 degree grid, for
                        mapping)'
                     units     = 'decimal degrees latitude'
                     valid_min = -55.1
                     valid_max = 0.0494
shelfmask
        Size:        902x552
        Dimensions:  lon_map,lat_map
        Datatype:    int16
        Attributes:
                     title     = 'Map of the shelf mask. shelfmask is
                         equal to 1 where downscaled data on the
                        shelf is provided and 0 where no data is
                        available (i.e., land points or ocean points
                        off the shelf)'
                     units     = 'none (logical 1/0)'
                     valid_min = 0
                     valid_max = 1
```

The indices of `lat_map` and `lon_map` at which the downscaled shelf data are provided are given by `js` and `is` respectively. In other words the variables `lon` and `lat` can be constructed from `lat=lat_map(js)` and `lon=lon_map(is)`.

```
is
        Size:        6064x1
        Dimensions:  location
        Datatype:    int16
        Attributes:
```

```
                        title     = 'index of all location longitudes (i
                           .e., lon = lon_map(is))'
                        units     = 'none (index)'
                        valid_min = 225
                        valid_max = 639
    js
           Size:        6064x1
           Dimensions:  location
           Datatype:    int16
           Attributes:
                        title     = 'index of all location latitudes (i.
                           e., lat = lat_map(js))'
                        units     = 'none (index)'
                        valid_min = 112
                        valid_max = 351
```

3 Loading data with MATLAB

The MATLAB script `OH14_loadmap.m` is provided which loads a subset of the OH14 data into memory. The user simply identifies the domain of interest by specifying the longitudes of the western and eastern boundaries (`lon_W` and `lon_E`), the latitudes of the southern and northern boundaries (`lat_S` and `lat_N`), and the start and end dates (`date_start` and `date_end`, in row-vector format `[year month day]`). The user must also provide the path to the OH14 NetCDF file (`path_to_OH14_nc`; e.g., `path_to_OH14_nc = 'OH14.nc'` or `path_to_OH14_nc = 'data/OH14.nc'` or `path_to_OH14_nc = '/home/ecoliver/data/OH14.nc'`). Temperatures and error estimates at any locations for which there is no downscaled shelf data (i.e., outside the domain specified by `shelfmask`) are returned as NaN.

For example, if the user wishes to load shelf SST data within the domain bounded longitudinally by (143°E,150°E), latitudinally by (44.5°S,37.5°S) and temporally between 1 April 1993 and 31 November 2001 (inclusive) then the following example script will do so:

```
> lon_W = 143; % 148 degrees East
> lon_E = 150; % 149 degrees East
> lat_S = -44.5; % 44 degrees South
> lat_N = -37.5; % 43 degrees South
> date_start = [1993 4 1]; % 1 April 1993
> date_end = [2001 11 31]; % 31 November 2001
> path_to_OH14_nc = '../data/OH14.nc'; % path to OH14 NetCDF file, in
   this case it lies in the directory data, sitting one level below the
    current directory
> [time, lon, lat, T, T_sige, T_mean, T_seas, T_seas_sige, T_res,
   T_res_sige] = OH14_loadmap(date_start, date_end, lon_W, lon_E, lat_S
   , lat_N, path_to_OH14_nc);
```

Note: NetCDF on older versions of MATLAB. Older versions of MATLAB (certainly version 2010a and earlier, possibly 2012a and earlier) lack the built-in MATLAB NetCDF API. You can check if you have this by typing `help ncread` at the MATLAB prompt – if no function help information is returned then you are lacking the built-in MATLAB NetCDF API. In this case, we have provided an alternative method of calling `OH14_loadmap.m` using the old `'netcdf_toolbox'` and `mexcdf` libraries[1]. To do this you must call `OH14_loadmap.m` with two

[1] http://mexcdf.sourceforge.net/

additional inputs: (i) specifying `'netcdf_toolbox'` as a character string and (ii) the path to a folder containing the `netcdf_toolbox` and `mexcdf` libraries (these are provided with the OH14 release, in the `lib/netcdf_toolbox/` folder):

```
> [time, lon, lat, T, T_sige, T_mean, T_seas, T_seas_sige, T_res,
   T_res_sige] = OH14_loadmap(date_start, date_end, lon_W, lon_E, lat_S
   , lat_N, path_to_OH14_nc, 'netcdf_toolbox', '../lib/netcdf_toolbox
   /');
```

Once the data are loaded into memory the user can plot, for example, a map of the mean SST (Figure 2):

```
> pcolor(lon, lat, T_mean);
> shading flat;
> colorbar;
> grid;
```

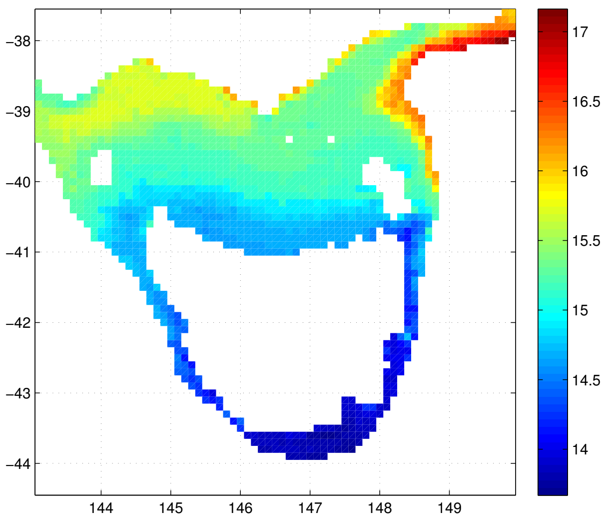

Figure 2: Example plot of Mean SST.

or a map of the total SST on 10 April 1993 (i.e., 9 days after `date_start`, or the 10th time step; Figure 3):

```
> pcolor(lon, lat, squeeze(T(10,:,:)));
> shading flat;
> colorbar;
> grid;
```

or a time series of SST at the nearest grid location to the East and North of (146°E,40°S), along with a ±1 standard deviation error estimate (67% confidence limits; Figure 4):

```
> i = find(lon>146, 1, 'first');
> j = find(lat>-40, 1, 'first');
> plot(time, T(:,j,i), 'b-', 'linewidth', 2);
> hold on;
> plot(time, T(:,j,i) + T_sige(j,i), 'k-');
> plot(time, T(:,j,i) - T_sige(j,i), 'k-');
> datetick('x', 10)
```

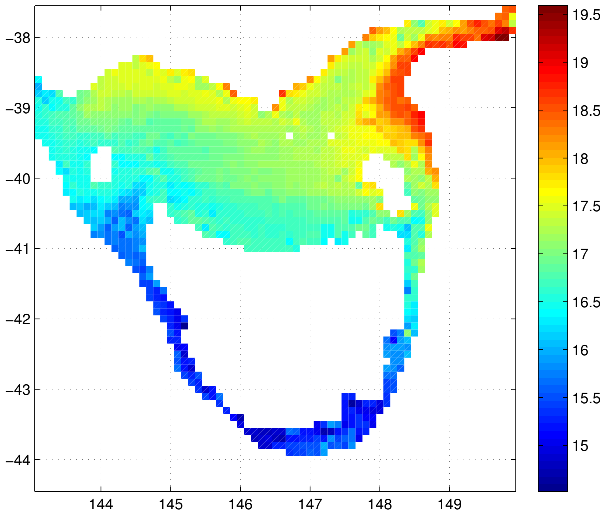

Figure 3: Example plot of SST on 10 April 1993.

```
> grid;
```

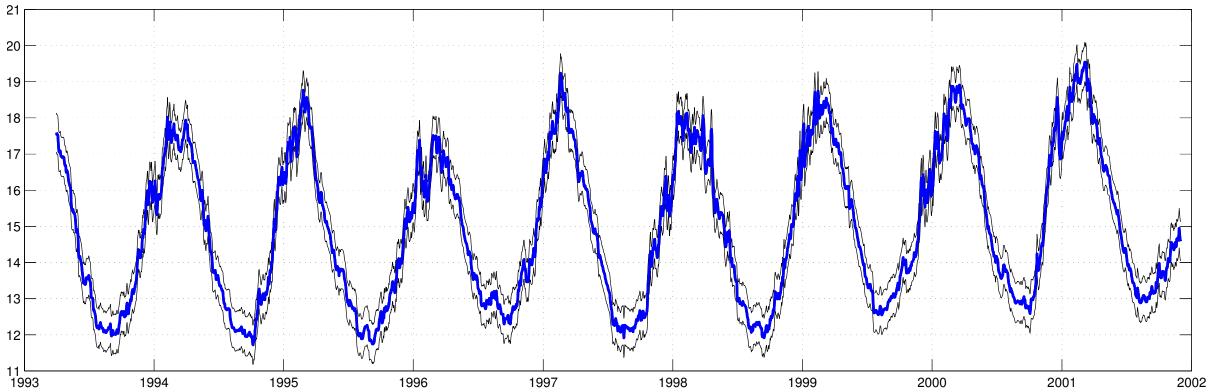

Figure 4: Example plot of SST at the nearest grid location to (146°E,40°S).

No scripts are provided in other languages (i.e., R or python) to load the OH14 data but it should be straightforward to create such a script using `OH14_loadmap.m` as a template.

4 Sample time series

Several comparisons of sea surface temperature time series from the OH14 data set against observations (AVHRR) and BRAN is shown in Figures 5–8 at selected locations on the southern and eastern Australian continental shelf. For a complete description of the technique used to generate the OH14 data set, and for additional comparisons of OH14 over the Australian continental shelf against observations, we refer the user to Oliver and Holbrook (2014).

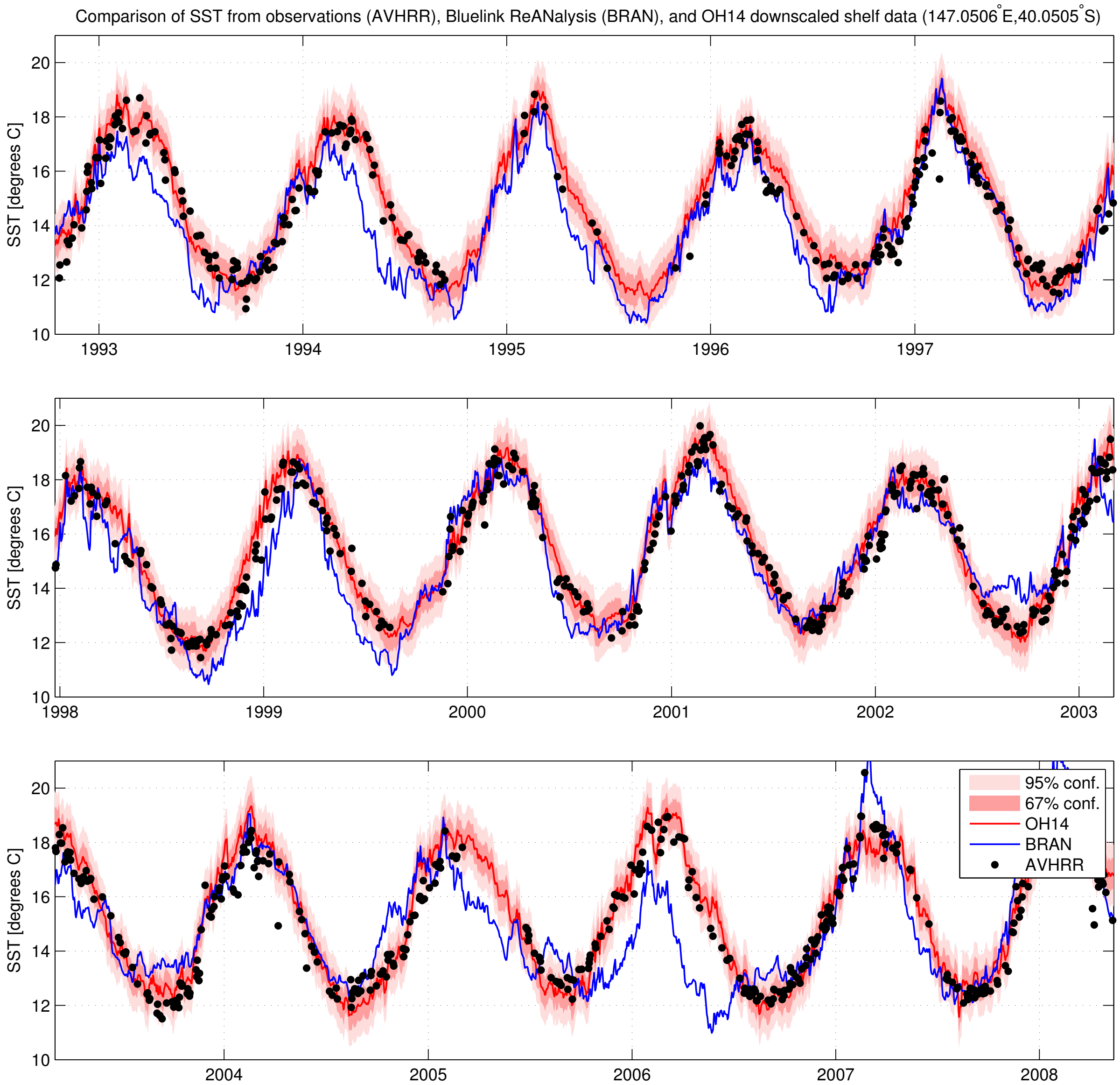

Figure 5: A comparison of sea surface temperature from observations (AVHRR, black dots), Bluelink ReANalysis (BRAN, blue line), and OH14 downscaled data (red line, with 67% and 95% confidence limits shown by shaded areas) for a location in Bass Strait.

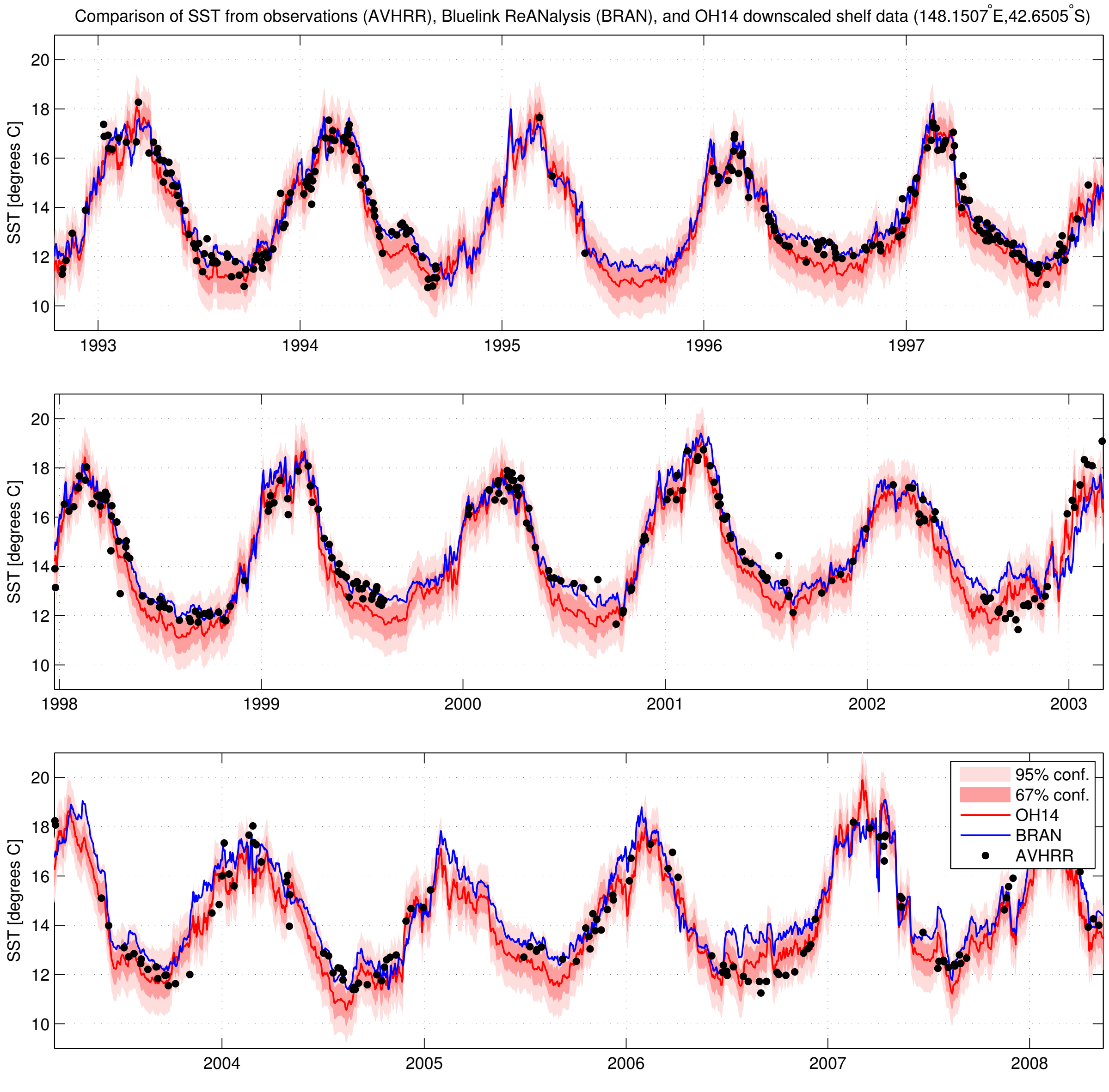

Figure 6: As in Figure 5 but for the nearest location to Maria Island, Tasmania.

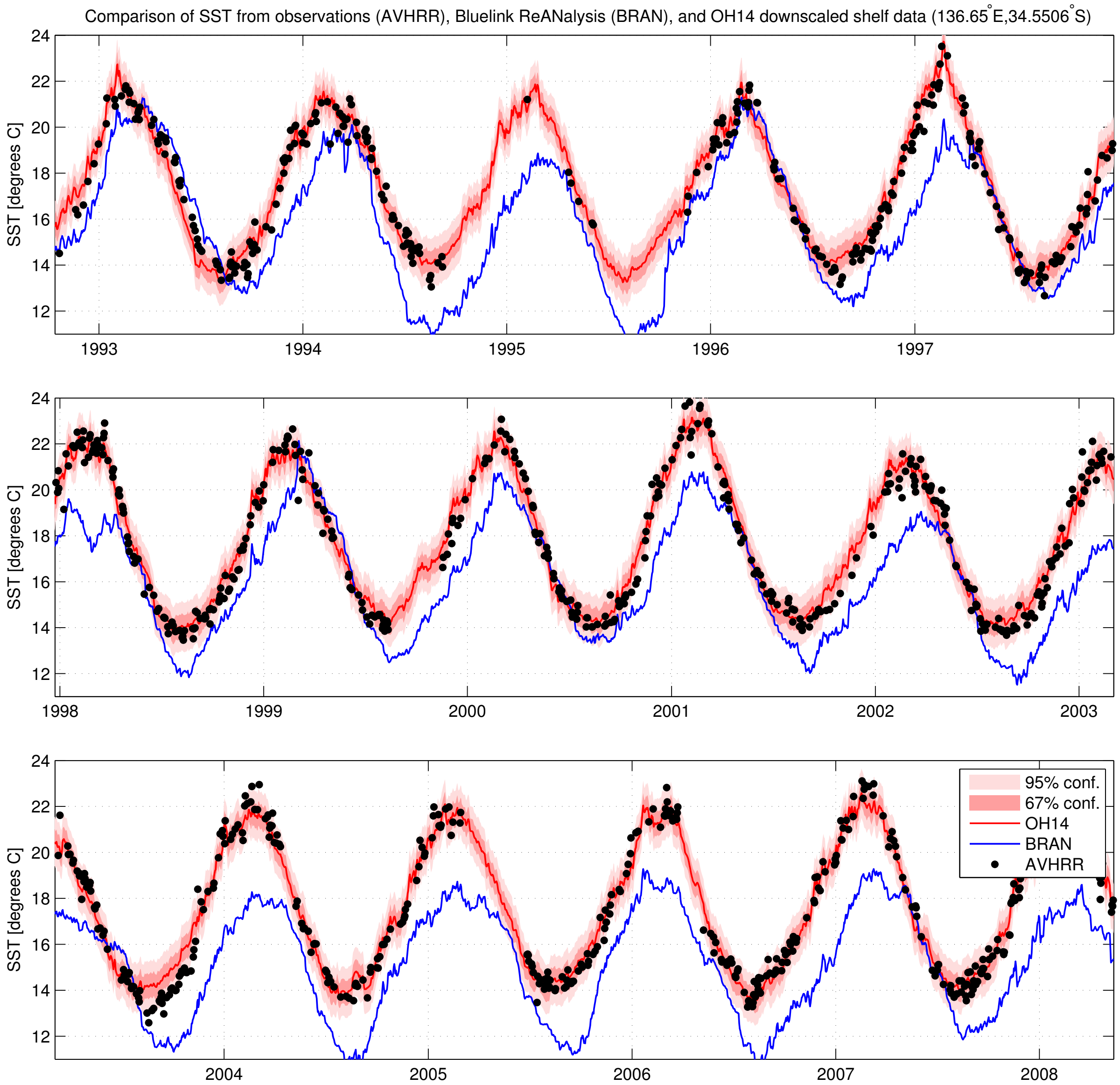

Figure 7: As in Figure 5 but for a location in Spencer Gulf, South Australia.

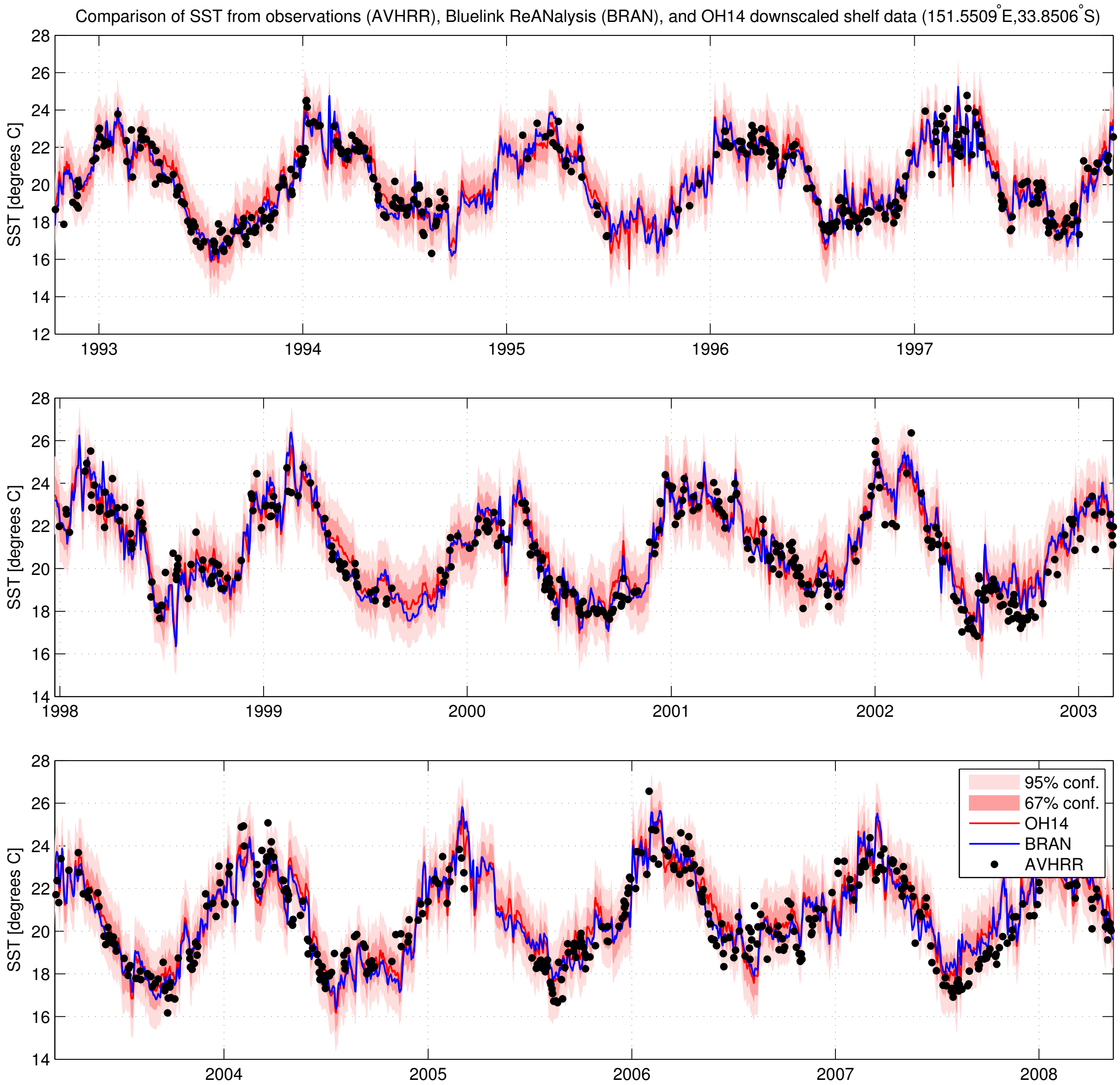

Figure 8: As in Figure 5 but for a location just East of Sydney.

Disclaimer

This data set is an experimental product and is provided “as is” without any warranty, including implied warranties of merchantability and fitness for a particular purpose. The use is forewarned that the methods and SST estimates are new and subject to further change and improvement. The names of the authors may not be used to imply and kind of endorsement.

Version history

- **Version 1.0** Initial “quiet” release, made available to Maria Beger and Amanda Bates.
- **Version 1.1** Updated “User’s Manual” text and MATLAB code surrounding `netcdf_toolbox` option in `OH14_loadmap.m` script to make them more user friendly. Added `dates` variable to `OH14.nc`, a vector of year-month-day dates. Made NetCDF float precision double instead of single.
- **Version 1.2** Updated “User’s Manual”. Made temperature and error variables in NetCDF as short integers, with a scale factor and an add offset to convert back to real numbers.

References

[1] E. C. J Oliver and N. J. Holbrook. A statistical method for improving continental shelf and near-shore marine climate predictions. *Journal of Atmospheric and Oceanic Technology*, 31:216–232, 2014.

[2] P. R. Oke and A. Schiller. Impact of Argo, SST, and altimeter data on an eddy-resolving ocean reanalysis. *Geophysical Research Letters*, 34(19), 2007.

[3] A. Schiller, P R Oke, G. Brassington, M. Entel, R. Fiedler, D A Griffin, and J V Mansbridge. Eddy-resolving ocean circulation in the Asian–Australian region inferred from an ocean reanalysis effort. *Progress in Oceanography*, 76(3):334–365, 2008.

[4] P. R. Oke, G. B. Brassington, D. A. Griffin, and A. Schiller. The Bluelink ocean data assimilation system (BODAS). *Ocean Modelling*, 21(1-2):46–70, 2008.

[5] Kenneth S Casey, Tess B Brandon, Peter Cornillon, and Robert Evans. The past, present, and future of the AVHRR Pathfinder SST program. In *Oceanography from Space*, pages 273–287. Springer, 2010.