Introducing pacea

An R package of Pacific ecosystem information to help facilitate an ecosystem approach to fisheries management

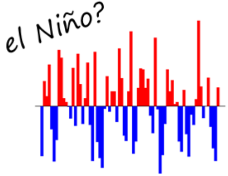
Andrew Edwards and Travis Tai

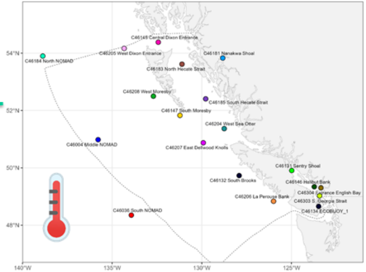
Introducing pacea: an R package of ecosystem information to help facilitate an ecosystem approach to fisheries management

Andrew Edwards[1,2] & Travis Tai[1]

[1]Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC.

[2]Dept. of Biology, University of Victoria, Victoria, BC.





Ecostats, UVic.
Tuesday 21st November 2023

Fisheries and Oceans Canada | Pêches et Océans Canada

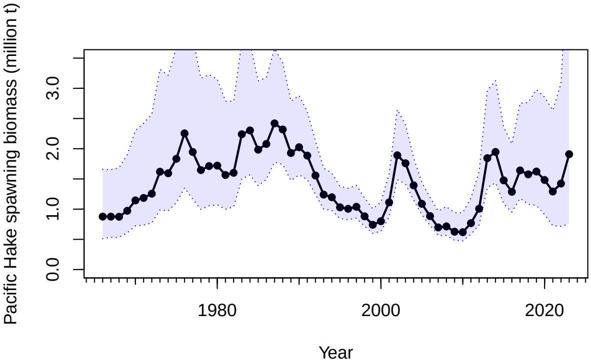
Acknowledgments

- Co-authors: Joe Watson, Angelica Peña, Andrea Hilborn, Charles Hannah, and Chris Rooper.
- Kelsey Flynn, Jessica Nephin, Lindsay Davidson, Strahan Tucker, Brianna Wright, Patrick Thompson, Matt Grinnell, Sean Anderson, Philina English, Chris Grandin, Jennifer Boldt, and others.
- Carley Colclough for designing the logo.
- DFO's Competitive Science Research Fund for funding (project 21-FS-03-13).

What is a stock assessment?

- Collect data from surveys and commercial fisheries, including ages of fish
- Data inform statistical population models to estimate historical biomass
- Models provide advice to fisheries managers

Pacific Hake assessment from 2023, estimate the historical population size:



What is a stock assessment?

Advice often in the form of ‘decision tables’ that indicate probabilities of future events given different catches:

Catch (t) in 2023	Probability spawning biomass declines from 2023 to 2024	Probability spawning biomass in 2024 falls below 40% of unfished biomass
0	50%	2%
180,000	72%	3%
225,000	75%	3%
320,000	78%	3%
430,000	85%	5%

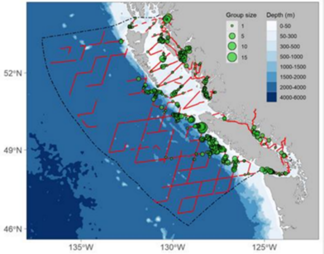
Motivation

- Revised Fisheries Act: "... the Minister shall take into account the environmental conditions affecting a fish stock."
- Yet <50% of DFO's stock assessments currently use environmental data.
- Only 28% of assessments in Pacific Region use environmental data.
- Leading cause of not using environmental data is availability of the data.

Kulka DW, Thompson S, Cogliati K, Olmstead M, Austin D, Pepin D. (2022). An Accounting of Integration of Environmental Variables in Fishery Stock Assessments in Canada. Can. Tech. Rep. Fish. Aquat. Sci. 3473: viii + 79 p.
https://publications.gc.ca/collections/collection_2022/mpo-dfo/Fs97-6-3473-eng.pdf

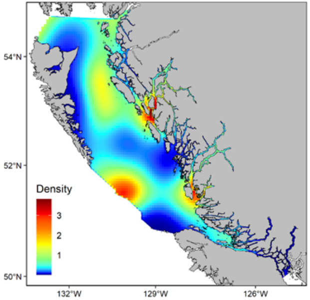
Motivation (based on a true story)

Survey sightings of Humpback Whales



Modelling

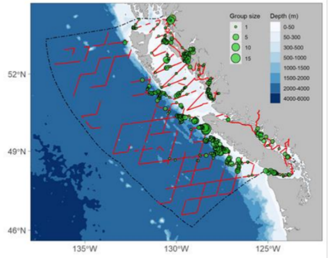
Estimated densities



See Doniol-Valcroze et al. (2022) in last year's SOPO report.
Density plot courtesy Brianna Wright.

Motivation (based on a true story)

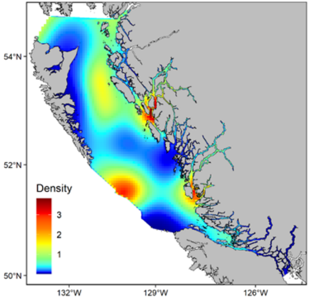
Survey sightings of Humpback Whales



Modelling

SST ?

Estimated densities



Motivation (based on a true story)

A search for sea surface temperature yields an overwhelming number (341) of choices:

ERDDAP
Easier access to scientific data
Brought to you by N

ERDDAP > Search

Do a Full Text Search for Datasets:

sea surface temperature Search

341 matching datasets, with the most relevant ones listed first.
(Or, refine this search with Advanced Search)

Grid DAP Data	Sub-set	Table DAP Data	Make A Graph	W M S	Source Data Files	Acces-sible	Title	Sum-mary	FGDC, ISO, Metadata	Back-ground Info	RSS	E mail	Institution
data			graph	M		public	Sea-Surface Temperature, NOAA ACSPO Daily Global 0.02° Gridded Super-collated SST and Thermal Fronts Reanalysis, 2012-present, Daily (L3S-LEO degrees C)		F I M	background	RSS		NOAA/NESDIS/STAR
data			graph	M		public	Sea-Surface Temperature, NOAA ACSPO NOAA-20 VIIRS CoastWatch Co-gridded 4km Daily (degrees C)		F I M	background	RSS		NOAA/NESDIS/OSPO
data			graph	M		public	Sea-Surface Temperature, NOAA ACSPO S-NPP VIIRS CoastWatch Co-gridded 4km Daily (degrees C)		F I M	background	RSS		NOAA/NESDIS/OSPO
data			graph	M	files	public	Sea-Surface Temperature, NOAA Geo-polar Blended Analysis Day+Night, GHRSST, Near Real-Time, Global 5km, 2019-Present, Daily		F I M	background	RSS		NOAA NESDIS Coast...
data			graph	M	files	public	Sea-Surface Temperature, NOAA Geo-polar Blended Analysis Diurnal Correction (Day+Night), GHRSST, Near Real-Time, Global 5km, 2019-Present, Daily		F I M	background	RSS		NOAA NESDIS Coast...

Likely requires extensive data wrangling to be usable, which usually takes way, way, longer than anticipated.
So the SST analysis did not happen.

Motivation

- Motivated by Dan Duplisea's `gslea` package for the Gulf of St. Lawrence.
- Outputs from a BC physical biogeochemical model have been shared by Angelica Peña.
- But the netCDF format may be unfamiliar to non-oceanographers – we take care of the wrangling into R
- Open Data is great, but can be hard to convert raw data into usable information.
- Primary audience is DFO stock assessment scientists, but usable by anyone (with a minimal working knowledge of R).

Duplisea et al. (2020). gslea: the Gulf of St Lawrence ecosystem approach data matrix R-package. R package version 0.1 https://github.com/duplisea/gslea

Why an R package?

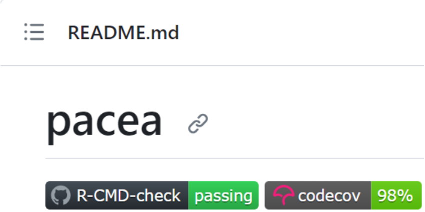
- R is the programming language most widely used by stock assessment scientists.
- An R package is the standard way to share code (and data).
- Easy installation:

```
remotes::install_github("pbs-assess/pacea")
```

- Ensures proper documentation of data objects and functions (helps users).
- Include vignettes which walk users through various features.
- All data (except one type) is saved within the package, no further downloading required.

Why an R package on GitHub?

- Can host freely on GitHub: https://github.com/pbs-assess/pacea
- Standard way to collaborate on code.
- Code is completely open source: transparent, traceable, and transferable.



- 'Passing' badge: ensures package can build in R.
- 'codecov' is the amount of the code that is covered by unit tests, which help weed out bugs.

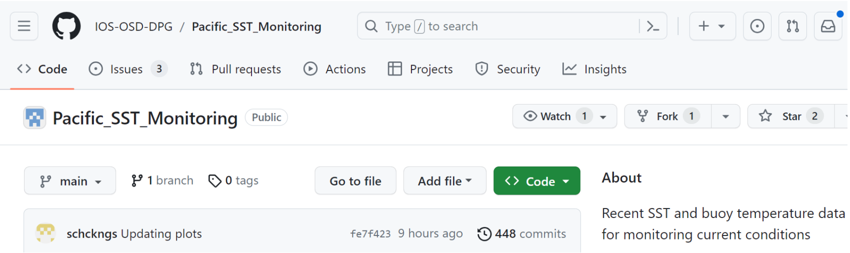
pacea objects

Currently, pacea contains:

- 203,092 calculations of daily sea surface temperature based on data from 19 buoys
- outputs from the spatial British Columbia continental margin (BCCM) model
- NOAA's spatial Optimum Interpolation Sea Surface Temperature (OISST) record
- 9 climatic and oceanographic indices, such as the Pacific Decadal Oscillation and those related to El Niño
- estimates of abundances for Harbour Seals and Pacific Hake

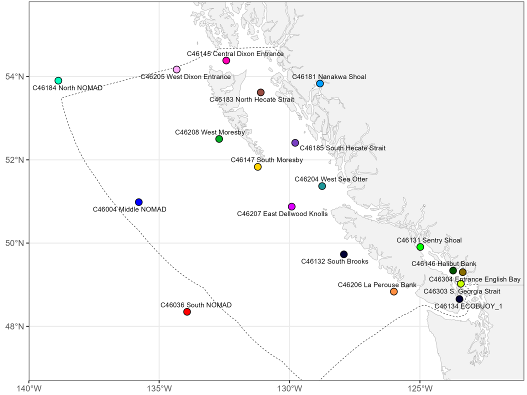
Sea surface temperature

Andrea Hilborn (IOS) maintains https://github.com/IOS-OSD-DPG/Pacific_SST_Monitoring which is updated weekly to show recent SST from various sources for the Northeast Pacific (and anomalies and marine heatwave information).



Sea surface temperature data from buoys

Data from 19 buoys maintained by DFO/ECCC (Environment and Climate Change Canada), hosted through CIOOS (Canadian Integrated Ocean Observing System).



Sea surface temperature data from buoys

Extensive data processing to calculate daily mean SST:

- Resolution is every 10 minutes or every hour
- Come with quality control flags (mostly)
- We decided that for a daily mean we require:
  - at least one measurement every 2 hours
  - $\geq$ 10 2-hour measurements every day
- Switch from DFO data stream to ECCC data stream when appropriate
- Still working on fixing a few outliers in the high-resolution data (hard to automate)

Sea surface temperature data from buoys

The data are saved as a `tibble` in `pacea`:

```
buoy_sst
# A tibble: 203,092 x 3
  date       stn_id   sst
  <date>     <fct>  <dbl>
1 1988-08-05 C46004  12.8
2 1988-08-06 C46004  12.7
3 1988-08-07 C46004  12.5
4 1988-08-08 C46004  12.5
5 1988-08-09 C46004  12.6
6 1988-08-10 C46004  12.6
# i 203,086 more rows
```

Sea surface temperature data from buoys

The data are saved as a `tibble` in `pacea`:

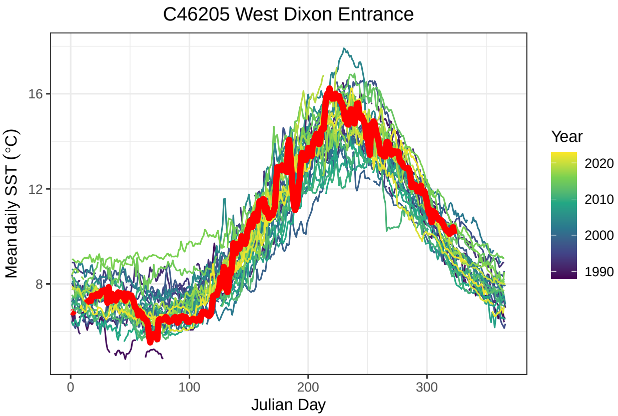
```
buoy_sst
# A tibble: 203,092 x 3
  date       stn_id   sst
  <date>     <fct>  <dbl>
1 1988-08-05 C46004  12.8
2 1988-08-06 C46004  12.7
3 1988-08-07 C46004  12.5
4 1988-08-08 C46004  12.5
5 1988-08-09 C46004  12.6
6 1988-08-10 C46004  12.6
# i 203,086 more rows
```

```
tail(buoy_sst)
# A tibble: 6 x 3
  date       stn_id   sst
  <date>     <fct>  <dbl>
1 2023-11-14 C46304  9.10
2 2023-11-15 C46304  8.90
3 2023-11-16 C46304  8.90
4 2023-11-17 C46304  9.03
5 2023-11-18 C46304  9.08
6 2023-11-19 C46304  9.24
```

Sea surface temperature data from buoys

Plot data from a single buoy for all years, the default is for buoy C46205:

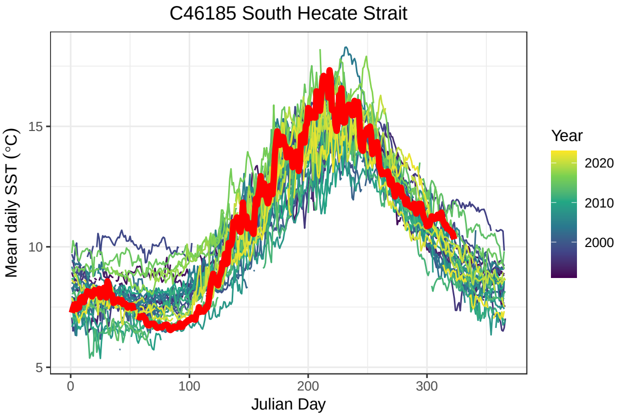
```
plot(buoy_sst)
```



Sea surface temperature data from buoys

Easy to look at values for any buoy

```
plot(buoy_sst,
     stn_id = "C46185")
```



Help file for every object

```
?buoy_sst

Daily average sea surface temperatures for 19 buoys (yielding over
200,000 values) calculated with data from Environment and Climate
Change Canada (ECCC) and DFO.

Description:

     A tibble of daily average calculations of sea surface temperature
     in coastal Canadian Pacific waters. The earliest data are from
     September 1987, and 14 buoys were still providing data as of May
     2023. See the example code below to see the start and end dates
     for each buoy.

Usage:

     buoy_sst

Format:

     A tibble with columns:

     date: date (of class `Date`) used to calculate the `sst`, based on
          UTC -8 hours (i.e. Pacific Standard Time, not changing due to
          daylight savings)

    ...
```

Vignettes for each type of data

We have written vignettes for each type of data. Viewable directly on GitHub site (see the README).

Buoy Sea Surface Temperature Data

Andrew Edwards

Daily average sea surface temperature from buoys

```
library(pacea)
library(dplyr)
library(tibble)  # Else prints all of a tibble
```

In pacea we include daily average sea surface temperature (SST) from 19 buoys in Canadian Pacific waters, yielding over 170,000 values. Data are from Environment and Climate Change Canada, and Fisheries and Oceans Canada. The earliest data are from September 1987, and 14 buoys were still providing data as of May 2023.

Metadata for the buoys is given by

```
buoy_metadata
#> # A tibble: 19 × 9
#>    wmo_id name     type  latitude longitude water_depth_m col_key stn_id name_key
#>    <fct>  <fct>    <fct>    <dbl>     <dbl>         <dbl> <fct>   <fct>  <fct>
#>  1 46004  Middle… NOMAD     51.0      -136.          3600 #0000FF C46004 C46004 …
#>  2 46036  South … NOMAD     48.4      -134.          3500 #FF0000 C46036 C46036 …
#>  3 46131  Sentry… 3 me…     49.9      -125.            18 #00FF00 C46131 C46131 …
#>  4 46132  South … 3 me…     49.7      -128.          2040 #000033 C46132 C46132 …
#>  5 46134  ECOBUO… 3 me…     48.7      -123.            65 #000033 C46134 C46134 …
#>  6 46145  Centra… 3 me…     54.4      -132.           257 #FF00B6 C46145 C46145 …
```

Climatic and oceanographic indices

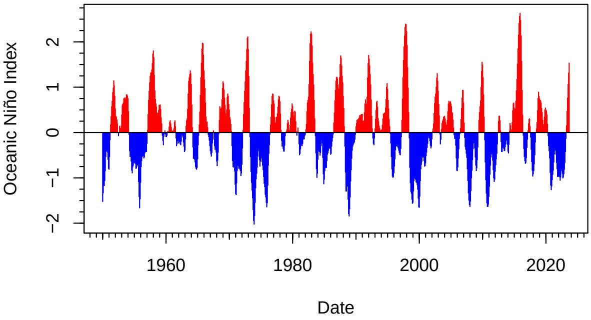
Various climate and oceanographic indices are currently included in pacea:

Object	Description	Resolution	Start year	End year
pdo	Pacific Decadal Oscillation	monthly	1854	2023
npi_monthly	North Pacific Index (monthly)	monthly	1899	2023
npi_annual	North Pacific Index (annual)	annual	1899	2023
alpi	Aleutian Low Pressure Index	annual	1900	2015
oni	Oceanic Niño Index	monthly	1950	2023
npgo	North Pacific Gyre Oscillation	monthly	1950	2023
ao	Arctic Oscillation	monthly	1950	2023
soi	Southern Oscillation Index	monthly	1951	2023
mei	Multivariate El Niño Southern Oscillation Index	monthly	1979	2023

Climatic and oceanographic indices

For example, to see a plot of the Oceanic Niño Index (ONI) anomaly, simply type

```
plot(oni)
```



This shows the onset of El Niño (positive index) conditions in April 2023.

Climatic and oceanographic indices

The values are readily available for doing your own analyses:

```
oni
# A tibble: 885 x 4
   year month value anomaly
  <dbl> <dbl> <dbl>   <dbl>
1  1950     1  24.7   -1.53
2  1950     2  25.2   -1.34
3  1950     3  25.8   -1.16
4  1950     4  26.1   -1.18
5  1950     5  26.3   -1.07
6  1950     6  26.3   -0.85
# i 879 more rows
tail(oni)
# A tibble: 6 x 4
   year month value anomaly
  <dbl> <dbl> <dbl>   <dbl>
1  2023     4  27.8    0.16
2  2023     5  28.3    0.48
3  2023     6  28.4    0.77
4  2023     7  28.4    1.07
5  2023     8  28.3    1.32
6  2023     9  28.3    1.54
```

Climatic and oceanographic indices

Each climatic and oceanographic index is saved as a data object, and properly documented, described, and referenced in its help file, for example:

```
?oni

ONI - Oceanographic Niño Index

Description:

     The Oceanic Niño Index is a monthly index which is one measure of
     the El Niño-Southern Oscillation.

Usage:

     oni

Format:

     A tibble also of class `pacea_index` with columns:

     year: year of value

     month: month (1 to 12) of value

     val: absolute values of three-month averages (preceding, current,
          and next month), deg C; note that recent values may change in
          subsequent updates - see details

     anom: anomalies based on 30-year base periods that are updated
```

Climatic and oceanographic indices

Continuing the help...

```
The Oceanic Niño Index (ONI) is a 3-month running mean of sea
surface temperature (SST) anomalies in the Niño 3.4 region (5 deg
N to 5 deg S, 120 deg W to 170 deg W) plotted on the center month.
The SST anomalies are calculated based on 30-year base periods
that are updated every 5 years, which accounts for global warming
and some of the decadal-scale SST variability (as seen in the
Pacific Decadal Oscillation index). The ONI is provided by the
NOAA's National Weather Service National Centers for Environmental
Prediction CPC:
http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml

...

    Because of the high frequency filter applied to the ERSSTv5 data,
`ONI values may change up to two months after the initial real
time value is posted`. Therefore, the most recent ONI values
should be considered an estimate.
```

Climatic and oceanographic indices

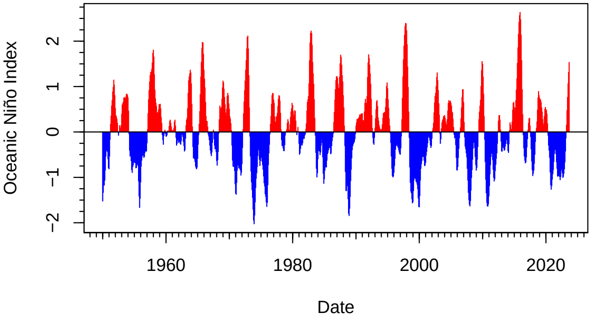
So for some analyses you may want to restrict values to a specified time period (for which values should not change in future pacea updates), for example

```
oni_fix <- dplyr::filter(oni,
                         year < 2023 | year == 2023 & month < 3)
tail(oni_fix)        # Will always end with February 2023
# A tibble: 6 x 4
   year month value anomaly
  <dbl> <dbl> <dbl>   <dbl>
1  2022     9  25.8   -1.01
2  2022    10  25.7   -0.99
3  2022    11  25.8   -0.92
4  2022    12  25.8   -0.82
5  2023     1  26.0   -0.67
6  2023     2  26.4   -0.43
```

Further plotting options and styles

Each index has a default plotting style (which you can override). For `oni` it shows the monthly anomalies as colour-code bars:

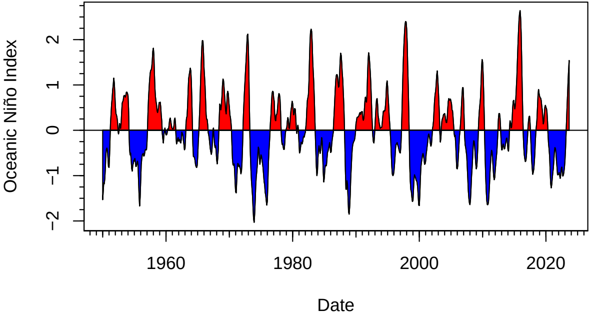
```
plot(oni)
```



Further plotting options and styles

Another option is as a black line with filled-in colouring:

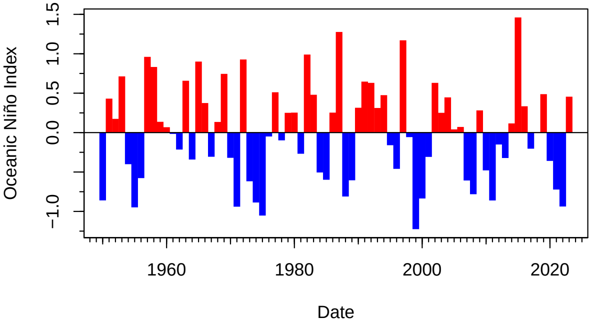
```
plot(oni,
     style = "red_blue")
```



Further plotting options and styles

To see `oni` as an annual (not monthly index):

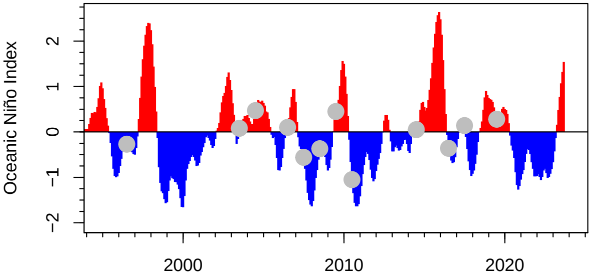
```
plot(oni,
     smooth_over_year = TRUE,
     lwd = 6)
```



Further plotting options and styles

And say you want to see if specific events coincide with El Niño (based on a true story, see help for details):

```
plot(oni,
     event_years = c(1996, 2003, 2004, 2006, 2007, 2008, 2009, 2010, 2014, 2016, 2017, 2019),
     xlim = c(lubridate::dmy(01011995),
              lubridate::dmy(01012024)),
     lwd = 2)
```



Technical aside regarding plot()

The object oni has the class pacea_index:

```
class(oni)
[1] "pacea_index" "tbl_df"      "tbl"         "data.frame"
```

which ensures that plot(oni) automatically uses our specialised function plot.pacea_index(oni), giving the red and blue colours, tickmarks, axis labelling, etc.

So see ?plot.pacea_index for further options. Similarly, check the class of other objects saved in pacea.

Estimates of animal populations

In pacea we include estimates of populations, with measures of uncertainty, as calculated from recent stock assessments.

Currently, we have estimates of:

- Pacific Hake (the coastwide stock off Canada and United States)
- Pacific Harbour Seals (most abundant pinniped species in the Northeast Pacific)
- Pacific Herring (coming soon)

Hake biomass

The hake biomass time series is saved as a `tibble` object in `pacea`:

```
hake_biomass
# A tibble: 58 x 4
   year   low median  high
  <dbl> <dbl>  <dbl> <dbl>
1  1966 0.510  0.877  1.66
2  1967 0.534  0.876  1.65
3  1968 0.534  0.874  1.68
4  1969 0.609  0.972  1.90
5  1970 0.723  1.15   2.30
6  1971 0.737  1.19   2.42
# i 52 more rows

tail(hake_biomass)
# A tibble: 6 x 4
   year   low median  high
  <dbl> <dbl>  <dbl> <dbl>
1  2018 1.08    1.58  2.77
2  2019 1.06    1.62  2.98
3  2020 0.910   1.48  2.85
4  2021 0.724   1.29  2.63
5  2022 0.716   1.42  3.08
6  2023 0.757   1.91  5.61
```

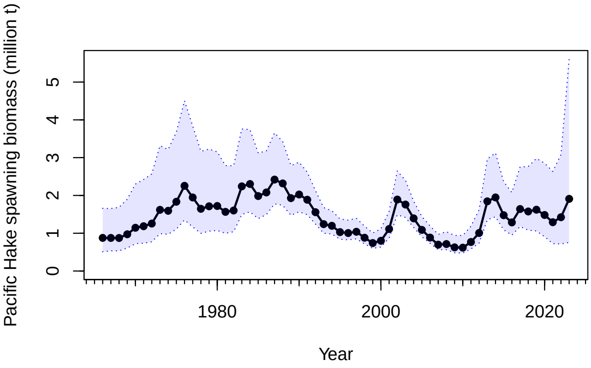
Hake biomass

```
?hake_biomass

Pacific Hake annual spawning stock biomass (females only) as estimated
by the 2023 stock assessment.

Description:

     The Pacific Hake stock is managed and assessed through an
     Agreement between Canada and the United States. The spawning stock
     biomass (mature females) estimates come from the most recent joint
     stock assessment, and, importantly, are for the coastwide stock
     from California to British Columbia.

Usage:

     hake_biomass

Format:

     A tibble also of class `pacea_biomass` with columns:

     year: year of the estimate of spawning stock biomass (mature
          females), where the estimate is for the start of the year

     low: low end (2.5th percentile) of the 95% credible interval for
          biomass, units as for `median`

     median: median estimate of biomass, in millions of tonnes of
          female spawning biomass at the start of the year
```

Hake biomass

Time series of spawning biomass:

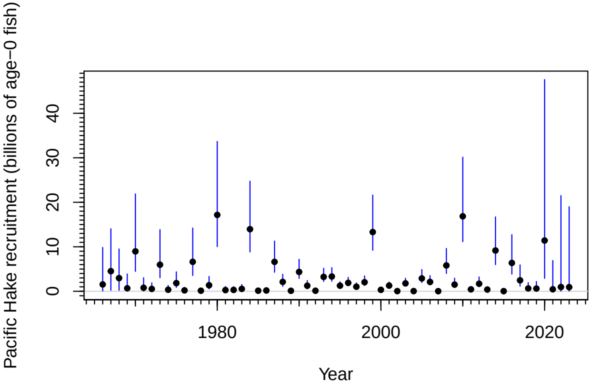
```
plot(hake_biomass)
```



Hake recruitment

Estimates of annual recruitment (exclude recent years for any analyses since not greatly informed by data; see `?hake_recruitment`):

```
plot(hake_recruitment)
```



Pacific Harbour Seals

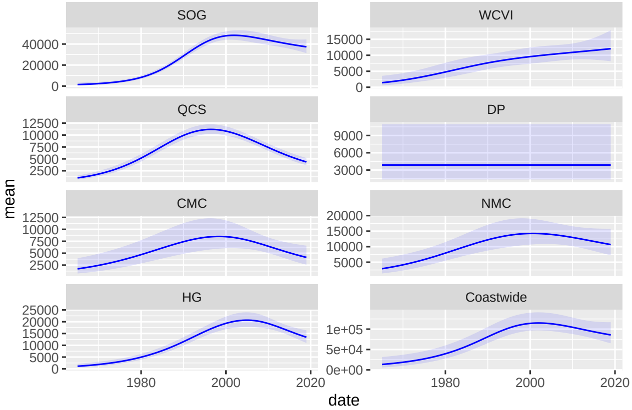
Estimated abundances for seven regions (and coastwide) were calculated by DFO (2022; SAR 2022/034), and included here:

```
harbour_seals
# A tibble: 640 x 5
  date       region    low  mean  high
  <date>     <fct>   <dbl> <dbl> <dbl>
1 1965-01-01 SOG      593. 1384. 3231.
2 1965-09-07 SOG      669. 1488. 3307.
3 1966-05-14 SOG      755. 1599. 3387.
4 1967-01-19 SOG      851. 1719. 3471.
5 1967-09-25 SOG      959. 1848. 3562.
6 1968-06-01 SOG     1080. 1988. 3661.
# i 634 more rows

tail(harbour_seals)
# A tibble: 6 x 5
  date       region       low   mean    high
  <date>     <fct>      <dbl>  <dbl>   <dbl>
1 2015-08-01 Coastwide 75072. 92401. 118166.
2 2016-04-07 Coastwide 73146. 91001. 117663.
3 2016-12-13 Coastwide 71189. 89638. 117317.
4 2017-08-20 Coastwide 69222. 88313. 117113.
5 2018-04-26 Coastwide 67259. 87026. 117031.
6 2019-01-01 Coastwide 65316. 85774. 117053.
```

Pacific Harbour Seals

Estimates for each region and coastwide, reproducing Figure 3 of DFO (2022):

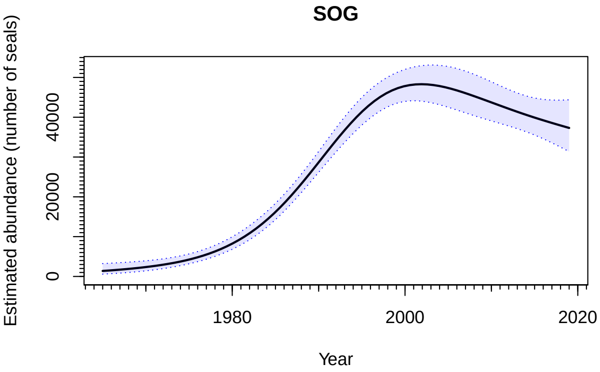
```
plot(harbour_seals)
```



Pacific Harbour Seals

Or just plot estimates for a single region:

```
plot(harbour_seals,
     region = "SOG")
```



NOAA's Optimal Interpolation Sea Surface Temperature

NOAA's Optimal Interpolation Sea Surface Temperature

OISST - long term gridded climate data record

- $1/4^\circ$ longitude x latitude
- Interpolated to fill gaps
- Canada's Pacific EEZ
- 925 gridded cells
- September, 1981 to (relatively) present
  - Updated ~monthly

2 data products: weekly and monthly mean SST

- `oisst_7day` and `oisst_month`

R `erddap` - https://www.ncei.noaa.gov/products/optimum-interpolation-sst

Huang, B., C. Liu, V. Banzon, E. Freeman, G. Graham, B. Hankins, T. Smith, and H. Zhang, 2021: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. J. Climate, 34, 2923–2939, https://doi.org/10.1175/JCLI-D-20-0166.1.

OISST – monthly mean

The data are saved as a `sf` (simple feature) `tibble` object in `pacea`:

```
oisst_month
```

```
Simple feature collection with 467125 features and 7 fields
Geometry type: POINT
Dimension:     XY
Bounding box:  xmin: -138.625 ymin: 46.625 xmax: -123.125 ymax: 54.625
Geodetic CRS:  WGS 84
# A tibble: 467,125 x 8
   year month   sst sst_sd sst_n start_date end_date
* <dbl> <dbl> <dbl>  <dbl> <int> <date>     <date>
1  1981     9  16.4  0.830    30 1981-09-01 1981-09-30
2  1981    10  14.3  0.838    31 1981-10-01 1981-10-31
3  1981    11  12.7  1.38     30 1981-11-01 1981-11-30
4  1981    12  10.2  0.549    31 1981-12-01 1981-12-31
5  1981     9  16.4  1.05     30 1981-09-01 1981-09-30
6  1981    10  14.2  0.697    31 1981-10-01 1981-10-31
# i 467,119 more rows
# i 1 more variable: geometry <POINT [B0]>
# i Use `print(n = ...)` to see more rows
```

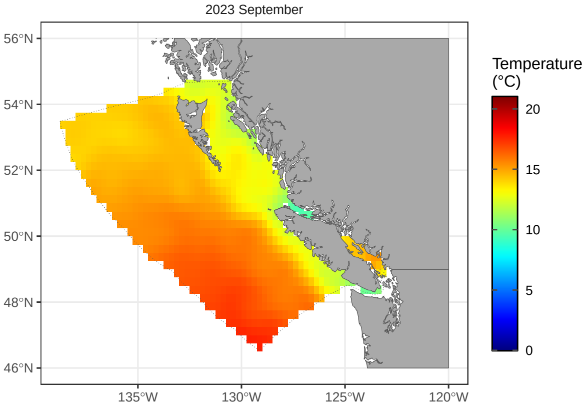
Help file

```
?oisst_month
```

```
Description
  Simple features objects of NOAAs Optimal Interpolation SST

oisst_7day
  7-day mean sea surface temperature (deg C)

oisst_month
  Monthly mean sea surface temperature in (deg C)

Usage
  oisst_7day
  oisst_month

Format
  year    Year

  week    Calendar week

  month   Calendar month

  sst     Mean sea surface temperature, SST (deg C)
...
```

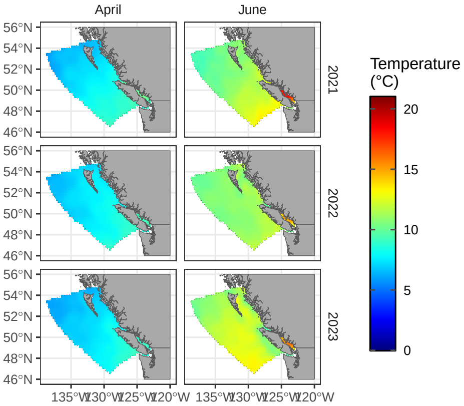
Built in plots – default

```
plot(oisst_month)
```



Built in plots – select years and months

```
plot(oisst_month, months.plot = c("April", "June"), years.plot = c(2021:2023))
```



British Columbia Continental Margin (BCCM) model

British Columbia Continental Margin (BCCM) model

Physical biogeochemical oceanographic model

- Regional Ocean Modelling Systen (ROMS)
- Curvilinear grid at 3km x 3km resolution
- Interpolated to regular grid
- Clipped to Canada's Pacific EEZ

Model output provided by Angelica Peña at the Institute of Ocean Sciences (Fisheries and Oceans Canada).

Peña, M.A., Fine, I. and Callendar, W. 2019. Interannual variability in primary production and shelf-offshore transport of nutrients along the northeast Pacific Ocean margin. Deep-Sea Research II, doi:10.1016/j.dsr2.2019.104637.

Nearest neighbour linear interpolation

Curvilinear –> regular grid

- Inshore (2 x 2km) and offshore (6 x 6km)
- Nearest neighbour interpolation
    - Performace = RMSE error
- 40,480 cells
- Monthly from 1993-2019

Available variables

The variables are:

- dissolved oxygen concentration
- pH
- salinity
- temperature
- depth-integrated phytoplankton
- depth-integrated primary production.

For applicable variables these are given for

- sea surface
- 0-40 m integration
- 40-100 m integration
- 100 m to the sea bottom
- sea bottom.

List of available variables

bccm_data for available variables

```
bccm_data
                             data_name
1             bccm_surface_oxygen
2                 bccm_surface_ph
3           bccm_surface_salinity
4        bccm_surface_temperature
5           bccm_avg0to40m_oxygen
6               bccm_avg0to40m_ph
7         bccm_avg0to40m_salinity
8      bccm_avg0to40m_temperature
9         bccm_avg40to100m_oxygen
10            bccm_avg40to100m_ph
11      bccm_avg40to100m_salinity
12   bccm_avg40to100m_temperature
13       bccm_avg100mtoBot_oxygen
14           bccm_avg100mtoBot_ph
15     bccm_avg100mtoBot_salinity
16  bccm_avg100mtoBot_temperature
17             bccm_bottom_oxygen
18                 bccm_bottom_ph
19           bccm_bottom_salinity
20        bccm_bottom_temperature
21             bccm_phytoplankton
22         bccm_primaryproduction
```

Downloading data locally

The data are saved as a `sf` (simple feature) in a separate GitHub repository `pacea-data` (https://github.com/pbs-assess/pacea-data).

There are built in functions to download specific variables:

```
bccm_avg100mtoBot_ph()

pacea would like to download and store these data in the directory:
C:\Users\TAIT\AppData\Local/pacea/Cache
Is that okay?
1: Yes
2: No

Selection:
```

```
Simple feature collection with 33642 features and 324 fields
Geometry type: POLYGON
Dimension:     XY
Bounding box:  xmin: 164671.6 ymin: 163875 xmax: 1097612 ymax: 1103653
Projected CRS: NAD83 / BC Albers
# A tibble: 33,642 × 325
  `1993_1` `1993_2` `1993_3` `1993_4` ...                      geometry
     <dbl>    <dbl>    <dbl>    <dbl> ...                 <POLYGON [m]>
1     7.83     7.87     7.88     7.90 ... ((575612.5 1103653, 577612.5 ~
2     7.83     7.87     7.88     7.90 ... ((577612.5 1103653, 579612.5 ~
3     7.83     7.87     7.88     7.90 ... ((579612.5 1103653, 581612.5 ~
```

Downloading data locally

Or all variables (will take several minutes):

```
bccm_all_variables()

Downloading all BCCM data may take many minutes. Files will be downloaded to pacea_cache directory:
C:\Users\TAIT\AppData\Local/pacea/Cache
Would you like to continue?
1: Yes
2: No

Selection:
```

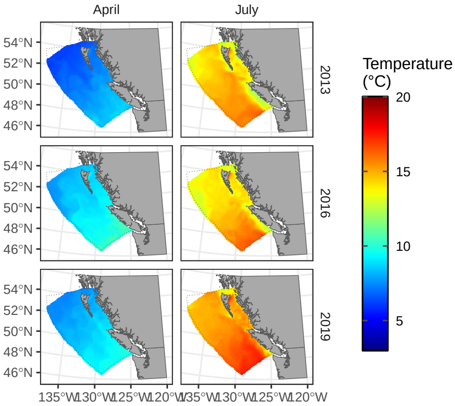
Help file

```
?bccm_surface_oxygen
```

```
BCCM environmental variable data

Description
  Loading and/or downloading British Columbia continental margin (BCCM) model data for the BC Pacific region to local drive.

  To download all BCCM data, use the function 'bccm_all_variables()'. See help page for details ('?bccm_all_variables').

Usage
  bccm_bottom_oxygen(update = FALSE, ask = interactive(), force = FALSE)

  bccm_bottom_ph(update = FALSE, ask = interactive(), force = FALSE)

  bccm_bottom_salinity(update = FALSE, ask = interactive(), force = FALSE)
  ...

Arguments
  update      Logical. Would you like to check for a newer version of the layer?

  ask         Logical. Should the user be asked before downloading the data to local cache? Defaults to the value of interactive().

  force     Logical. Should download of data be forced? Overrides 'ask' argument if TRUE.

Format
  A simple features dataframe.
...
```

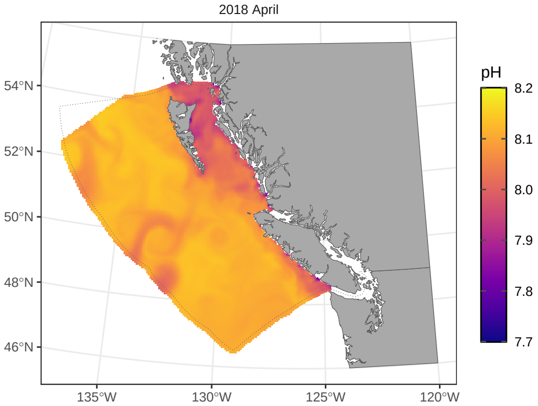
Built in plotting

```
plot(bccm_surface_temperature(), months.plot = c("April", "july"), years.plot = c(2013, 2016, 2019))
```



Built in plotting

```
plot(bccm_surface_ph())
```



Extras – climatology function

Built in functions to calculate 30-year climatology

```
?calc_clim
```

```
Description
  Function for calculating climatology of a pacea data object.

Usage
calc_clim(
  data,
  clim_years = c(1991:2020),
  clim_time = "month",
  time_period_return = "all"
)

Arguments
  data                  pacea data object: BCCM, OISST, buoy_sst data only

  clim_years              climatology period years

  clim_time             time units (e.g. month) to summarize climatologies

  time_period_return      vector of value(s) for the specific time units to estimate climatologies (e.g. '4' for week 4 or April). Set
...
```

Extras – climatology function

Output with each climatology time unit (e.g. week, month)

```
calc_clim(bccm_surface_temperature())
```

```
Simple feature collection with 486960 features and 4 fields
Geometry type: POLYGON
Dimension:     XY
Bounding box:  xmin: 164671.6 ymin: 163875 xmax: 1097612 ymax: 1103653
Projected CRS: NAD83 / BC Albers
# A tibble: 486,960 x 5
  month clim_value clim_sd clim_n                  geometry
  <dbl>      <dbl>   <dbl>  <int>             <POLYGON [m]>
1     1       5.80   0.721     27 ((575612.5 1103653, 5776~
2     2       5.47   0.610     27 ((575612.5 1103653, 5776~
3     3       5.64   0.627     27 ((575612.5 1103653, 5776~
4     4       6.61   0.623     27 ((575612.5 1103653, 5776~
5     5       8.49   0.732     27 ((575612.5 1103653, 5776~
6     6      10.5    0.662     27 ((575612.5 1103653, 5776~
# i 486,954 more rows
# i Use `print(n = ...)` to see more rows
```

And we can also use this with the OISST and buoy data

```
calc_clim(oisst_7day)
calc_clim(buoy_sst)
```

Extras – anomaly function

Built in functions to calculate anomaly to the 30 year climatology

```
?calc_anom
```

```
Description
  Function for calculating climatology of a pacea data object.

Usage
calc_anom(
  data,
  clim_years = c(1991:2020),
  clim_time = "month",
  time_period_return = "all",
  years_return
)

Arguments
  data                pacea data object: BCCM, OISST, buoy_sst data only

  clim_year           climatology period years

  clim_time           time units (e.g. month) to summarize climatologies

  time_period_return    vector of value(s) for the specific time units to estimate climatologies (e.g. '4' for week 4 or April). Set t

  years_return        vector of value(s) to return the years of interest. Defaults to all years in input data
...
```

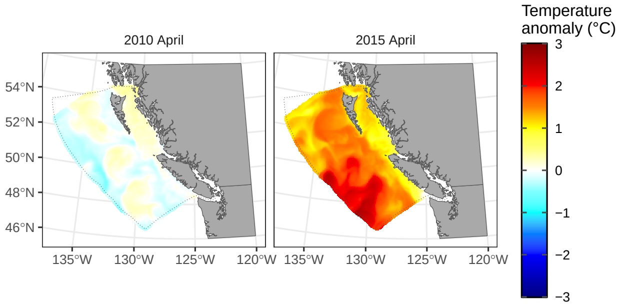
Extras – anomaly function

Specify year(s) and time period (e.g. week, month) to return – otherwise all.

```
anom_dat <- calc_anom(bccm_surface_temperature(), time_period_return = "April", years_return = c(2010, 2015))

Simple feature collection with 40580 features and 2 fields
Geometry type: POLYGON
Dimension:     XY
Bounding box:  xmin: 164671.6 ymin: 163875 xmax: 1097612 ymax: 1103653
Projected CRS: NAD83 / BC Albers
# A tibble: 40,580 x 3
  `2010_4` `2015_4`                                                                    geometry
     <dbl>    <dbl>                                                               <POLYGON [m]>
1    0.114     1.09 ((575612.5 1103653, 577612.5 1103653, 577612.5 1101653, 575612.5 1101653, 575~
2    0.124     1.08 ((577612.5 1103653, 579612.5 1103653, 579612.5 1101653, 577612.5 1101653, 577~
3    0.150     1.07 ((579612.5 1103653, 581612.5 1103653, 581612.5 1101653, 579612.5 1101653, 579~
4    0.204     1.08 ((581612.5 1103653, 583612.5 1103653, 583612.5 1101653, 581612.5 1101653, 581~
5    0.204     1.07 ((583612.5 1103653, 585612.5 1103653, 585612.5 1101653, 583612.5 1101653, 583~
6    0.244     1.08 ((585612.5 1103653, 587612.5 1103653, 587612.5 1101653, 585612.5 1101653, 585~
# i 40,574 more rows
# i Use `print(n = ...)` to see more rows
```

Extras – anomaly plot

```
plot(anom_dat, months.plot = "April", years.plot = c(2010, 2015))
```



Extras – other tidbits in vignettes

Geospatial functions in vignettes

- `area_mean()` estimates inshore and offshore spatial mean (200m isobath)
- Masking/clipping specific areas (e.g. fishing regions of interest)
- Various other data visualisation examples

Commitments from package authors

We will:

- (try to) not break the package, by working on branches on GitHub before merging.
- ensure any improvements are back compatible.
- update the NEWS as we add/update data and make improvements.
- ideally update the following on the 20th of every month (or soon after) to get the latest raw data:
  - climatic and oceanographic indices
  - buoy sst
  - OISST

Commitments from package authors

- update other data as appropriate:
  - hake assessment output: February each year
  - harbour seals: next time assessment is updated
  - herring assessment output: when SAR published each year
  - BCCM model output: in early 2024 the model may be run into 2020 (boundary conditions come from a global model currently only available to Jan 2021).
    - note that BCCM is a research model (not an operational model) so updates depend on projects and access to High Performance Computing.

Recommended guidelines for users

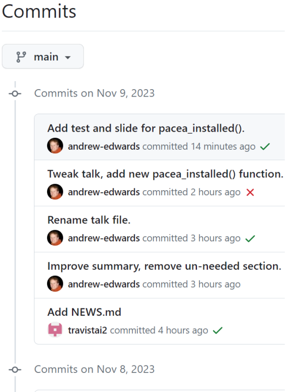
Please:

- read the README!!
- check the help files and vignettes
- check the NEWS (for updates to data, useful additions, new data, etc.)
- understand the limitations of the data

Recommended guidelines for users

To check when you last installed pacea

```
pacea_installed()

You last locally installed pacea on 2023-11-01 which is 8 days ago.

Compare that with the updates in the NEWS file (and you can also check dates of the latest commits at
https://github.com/pbs-assess/pacea/commits/main)
```



Recommended guidelines for users

- any problems, make a minimum working example (usually makes you figure it out yourself)
  - check the `Issues` on GitHub in case someone else has asked or we are already working on it
  - simple quick question just ask us (email, or Andy is on EcoStats Slack channel)
  - more detailed questions or bug reports post a new `Issue` (allows others to see, and maybe answer, your question)

Recommended guidelines for users

- cite pacea and the original sources for any data you use
- let us know of any applications (so we understand the uptake and how much effort to devote to maintaining it)
  - you could be out on a survey and wondering how warm the water is compared to previous years
  - Chris Rooper: looking at Pacific Saury process errors from a stock assessment using the indices (PDO etc.). "It made my life easier today."

We wrangle the data so you don't have to

```
citation("pacea")

Edwards AM, Tai TC, Watson J, Peña MA, Hilborn A, Hannah CG, Rooper CN
(2023). "pacea: An R package of Pacific ecosystem information to help facilitate an ecosystem approach to fisheries management."
  https://github.com/pbs-assess/pacea.
```