

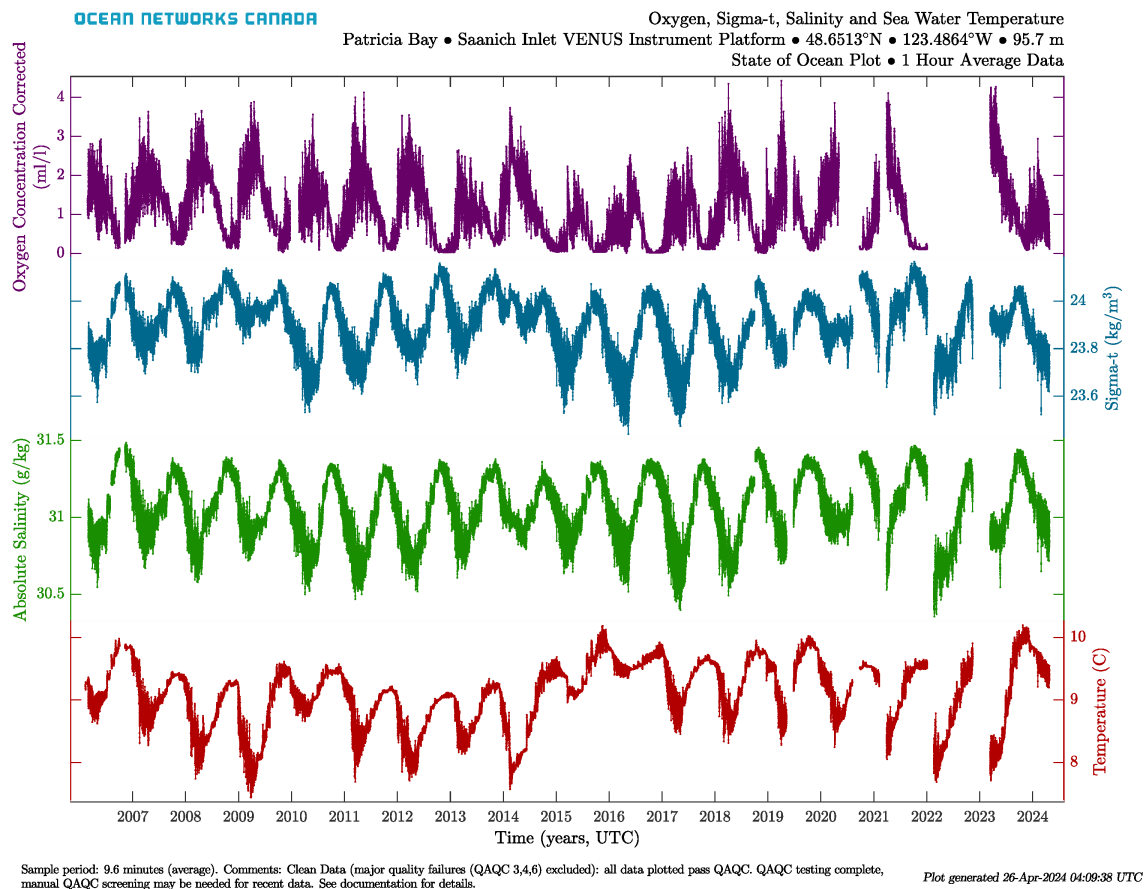
State of Ocean and Environment Data Products

State of Ocean and Environment Plots

State of Ocean and Environment plots are designed to give a long-term overview of state variables for either the Ocean or Atmosphere. These plots appear on the website and on the Data Preview tool. They are equivalent to a 'Variables by Location' search with 'All Available' time range, 1 hour averaging for the four or five variables put together into one search request. These searches are run on a schedule, usually once per day. State of Ocean /Environment plots are essentially device-level time series scalar plots, but have a specific colour scheme, as follows:

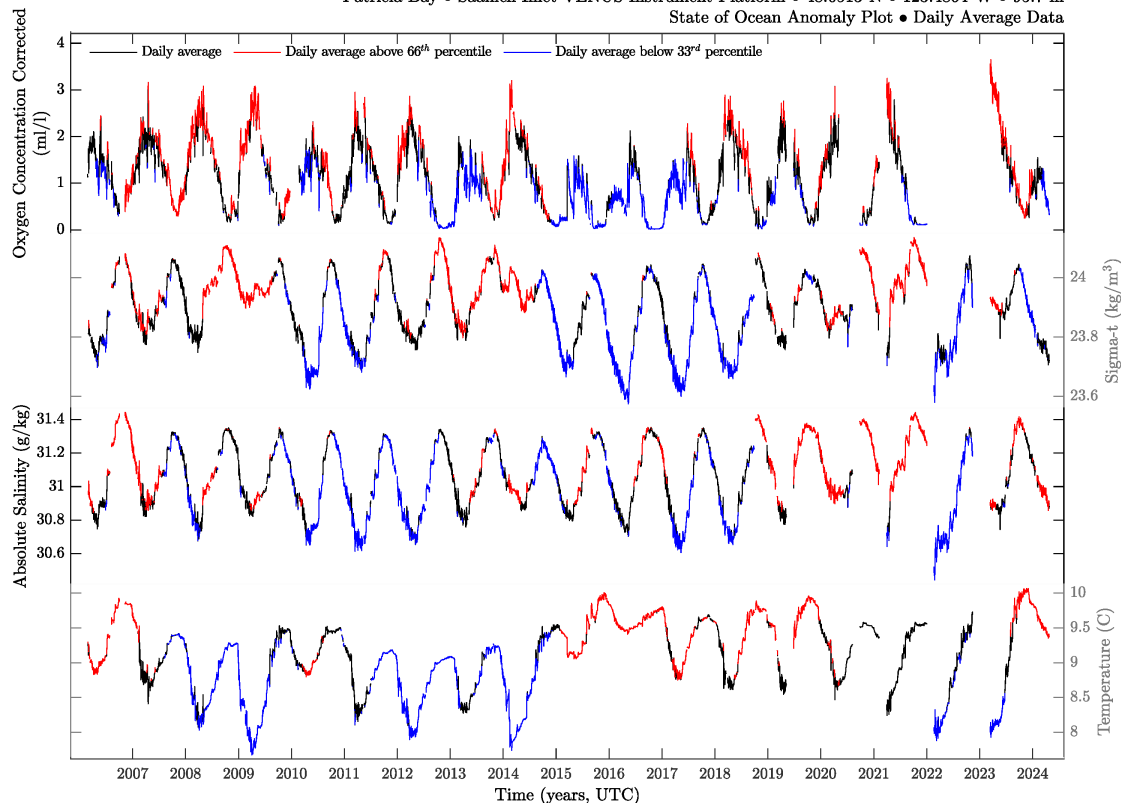
Sensor Type	Colour	State of:	Sensor Type	Colour	State of:
Sea Temperature	Red	Ocean	Air Temperature	Red	Atmosphere
Salinity	Green	Ocean	Barometric Pressure	Green	Atmosphere
Sigma-Theta (0 dbar)	Blue	Ocean	Solar Radiation	Blue	Atmosphere
Oxygen Concentration	Magenta	Ocean	% Humidity	Magenta	Atmosphere
Ice Draft	Orange	Ocean			

Here is an example State of Ocean plot for the Saanich Inlet VENUS Instrument Platform:



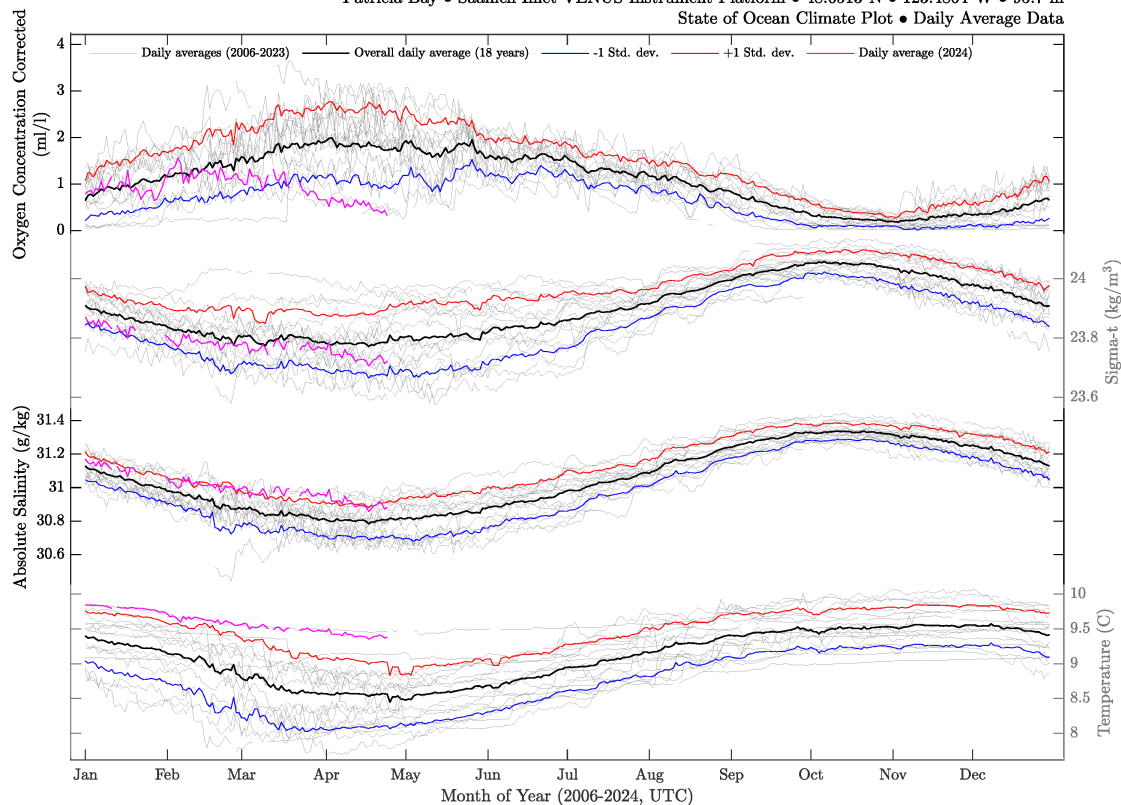
State of Ocean and Environment Climate and Anomaly Plots

These plots are based on the same information/configuration as the original State of Ocean/Environment plots, except that they use daily min/max + average data. These plots are available in Data Preview alongside the originals. Here are examples from Saanich Inlet:



Sample period: 9.5 minutes (average). Comments: Clean Data (major quality failures (QAQC 3,4,6) excluded): all data plotted pass QAQC. QAQC testing complete, manual QAQC screening may be needed for recent data. See documentation for details. Percentiles calculated from daily averages.

Plot generated 26-Apr-2024 04:19:49 UTC



Sample period: 9.5 minutes (average). Comments: Clean Data (major quality failures (QAQC 3,4,6) excluded): all data plotted pass QAQC. QAQC testing complete, manual QAQC screening may be needed for recent data. See documentation for details. Standard deviations calculated from daily averages. Leap days excluded.

Plot generated 26-Apr-2024 04:19:54 UTC

The anomaly plot (left) shows a time series of daily averages coloured on whether the value for that date is above the 66 percentile (red) or is below the 33 percentile (blue), where the percentiles are calculated from daily average values for that day of the year. The climatology plot (right) shows the day of year data for each sensor, showing the seasonality of the data. The grey lines in this plot are daily averages from every year, while the black lines are the overall day of year averages computed from the daily averages. The red and blue lines represent the day of year average plus and minus 1 standard deviation, respectively. Day of year standard deviation is also calculated from the daily averages from each day of the year. The magenta line is the current years' daily averages. Leap days are excluded from the climate plot and the day of year averages. The percentiles and standard deviations aren't showed if there is less than 3 years of data on each date, in that case, the anomaly plot will only show daily averages as black lines, the climate plot will omit the red and blue lines and comments will be added to the plots to indicate that this has occurred (it's common for new deployments).

State of Ocean and Environment Climate and Anomaly Data

The data behind these plots are available via links on their display in Data Preview. The file data products (CSV, MAT, ODV) are as documented in [time series scalar data products](#), with the exception of the MAT file, which contains an additional structure called 'dataClimate'. 'dataClimate' is a copy of the 'data' structure, excluding non-climate sensors and replacing the min/max+avg values that exist in the 'data' structure with additional calculated values:

dataClimate														
1x4 struct with 40 fields														
Fields	dailyTime	dailyAverage	dailyCount	dailyAverageByYear	dailyCountByYear	iYear	dayOfYearAverage	dayOfYearCount	dayOfYearCountDays	dayOfYear33Percentile	dayOfYear66Percentile	dayOfYearStdDev	dailyAverageAbove66Percentile	dailyAverageBelow33Percentile
1	3947x1 double	3947x1 double	3947x1 double	12x365 double	12x365 double	12x1 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	3947x1 double	3947x1 double
2	3947x1 double	3947x1 double	3947x1 double	12x365 double	12x365 double	12x1 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	3947x1 double	3947x1 double
3	3947x1 double	3947x1 double	3947x1 double	12x365 double	12x365 double	12x1 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	3947x1 double	3947x1 double
4	3947x1 double	3947x1 double	3947x1 double	12x365 double	12x365 double	12x1 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	1x365 double	3947x1 double	3947x1 double

dailyTime, *dailyAverage*, *dailyCount* are copies of time, average, count in the normal 'data' structure, with data gaps filled. The code is capable of calculating these variables from hourly or other averaged data (currently we use daily data, so this step is essentially a pass-through). The *dailyAverageByYear*, *dailyCountByYear* are a reshape of the *dailyAverage* and *dailyCount* data into a matrix that is year by date, where the year is specified by *iYear* and the date is 1 to 365 (leap days excluded), so the matrix is n years by 365 days. *dayOfYearAverage*, *dayOfYearCount*, *dayOfYearStdDev*, *dayOfYear33Percentile*, *dayOfYear66Percentile* are the day of year statistics for each of the 365 years of the year, calculated using the *dailyAverageByYear* data only, without weighting by *dailyCount* or recalculation from the raw data. This average of an average approach is normal metrological procedure, as many metrological data sets only report daily min/max average values. The *dayOfYearCount* is the count of the raw readings on that day of the year, whereas *dayOfYearCountDays* is the number of days that have data. So for the 15 years of Saanich Inlet, if only 5 years have data on a specific day, that *dayOfYearCountDays* value is 5, but *dayOfYearCount* will be ~400000 readings. *dayOfYearCountDays* is used to determine if there is enough data (minimum 3) for a valid percentile calculation. *dailyAverageAbove66Percentile*, *dailyAverageBelow33Percentile* are the values used the anomaly plot, so that they are daily time series of the same size as the daily average data, and their values are either the *dailyAverage* or NaN if they are not above/below the percentile. The anomaly plot is then simply:

```
hold on;
plot(dailyTime, dailyAverage)
plot(dailyTime, dailyAverageAbove66Percentile, 'r')
plot(dailyTime, dailyAverageBelow33Percentile, 'b')
```

so that the red/blue colouring goes over top of the black daily average.

One more addition to the State of Ocean / Environment file products is a special "externally-derived" sensor. The MAT, CSV and ODV file have a "levelled-pressure" sensor in addition to the regular scalar pressure sensor. This purpose of this version of the pressure is to remove the variation between deployments so that a continuous time series is available. The algorithm to derive it is as follows: for each deployment remove the mean from the time series (i.e. zero mean), then concatenate the time series back together and add a fixed offset, calculated from the site depth (from the metadata) converted to pressure using the GSW toolbox. The effect is that the jumps in value at each deployment are removed and the time series is much more suitable for calculating tides, storm surge, etc. Users can do this calculation themselves with the deployment data that exists in the headers of the ODV and CSV files or the metadata structure in the MAT file.