Using the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) to measure sea surface temperature in West Scottish sea-lochs.

Harry Colvile

Department of Geography, Durham University & Scottish Association of Marine Science (SAMS)

August 2019

SAMS Internal Report number: 295

Edited Paul Tett, November 2019

Abstract

Understanding how environmental variables such as Sea Surface Temperatures (SST) affect biological primary production within Western Scottish sea-lochs (fjords) is important for fishing and tourism. A multi-year SST time-series has been lacking for sites in the Firth of Lorn. This report compares an ICOADS (International Comprehensive Ocean-Atmosphere Data Set), starting 1970, with observations of SST within sea lochs of this Firth. Two datasets were compiled and used to provide monthly SST data from sea-lochs: one from a temperature recorder deployed at 10 m depth since the mid-1990s; and the other from CTD observations made during inshore research cruises. The ICOADS was well but not perfectly correlated with other two datasets (r^2 0.86 to 0.90), and showed a greater annual range in temperature than the inshore data. Nevertheless, it can, with corrections, be used as a proxy for inshore SST.

1. Introduction

Western Scotland is host to a variety of different inland sea-lochs connected the NW Atlantic Ocean. How the waters off the coast of Britain are being affected by anthropogenic climate change (CC) is increasingly becoming the subject of interest for both fisheries and to an extent tourism (Wells, 2016). This is due in part due to how changes are likely affecting the abundance of marine plankton and thus the overall foodweb. Studies have found that some types of plankton within Firth of Lorn sea-lochs have decreased in biomass since the 1970s (Tett, 2014, 2019), exemplified in figure 1. Possible causes of the changes include increased water temperatures and decreased upper water salinities and transparencies due to greater precipitation and run-off. Both causes are associated with climate change and have the potential to influence primary biological productivity and thus the overall health of local marine ecosystems.

This report focuses on Sea Surface Temperature (SST), and attempts to remedy the lack of compiled SST time-series from sea-lochs. CTD observations have been made, intermittently, by SAMS (Scottish Association of Marine Science) research cruises commencing in 1970. These provides high quality data for specific time/location but lack long term coverage. In contrast data from ICOADS (International Comprehensive Ocean-Atmosphere Data Set) are less well spatial resolved but provide greater frequency of observation (Freeman et al. 2017), even if some uncertainty surrounds particular ICOADS data in terms of specific location and collection method(s).

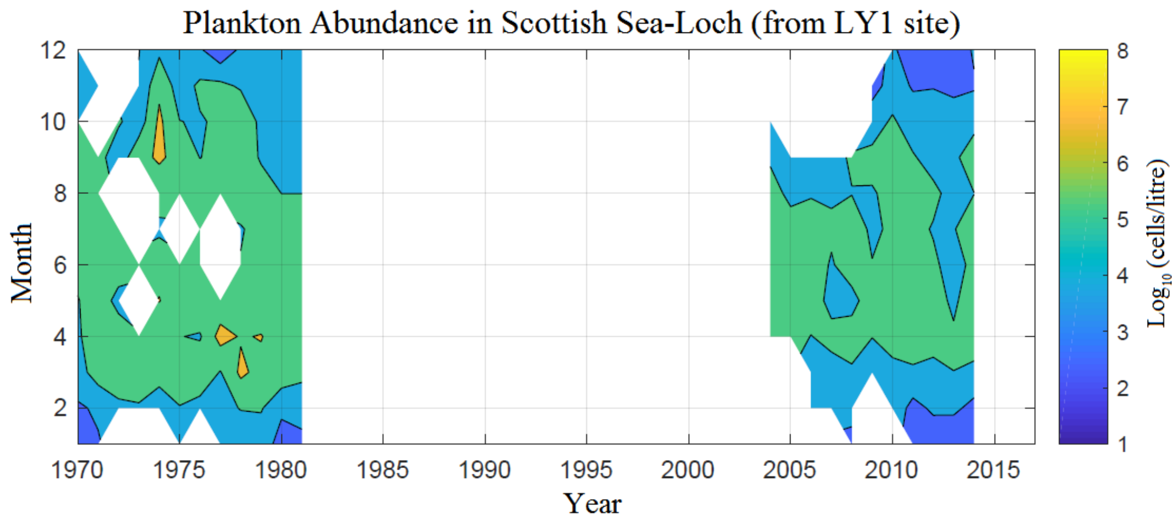


Figure 1, The Abundance of a common type of phytoplankton known as diatoms, at LY1 site. with bluer colours representing less abundance and yellower colours representing more abundance originally taken from Tett PC, (now found in Tett 2019).

The ICOADS global dataset draws from a variety of sources including 455 million unique reports from: the log-books or data records of naval and commercial ships; buoys (moored/drifting) and various research expeditions (Freeman et al. 2017). The data set does not use interpolation, being solely comprised of observational data making its usage primarily in calibrating meteorology models (Berry & Kent 2011). This may be why details of how and exactly when data were collected at a given grid cell in a given month, are difficult to find. This report uses monthly data from the ICOADS for the gridcell centered on 56.5 °N 5.5 °W and compares it with SAMS ovservations of SST at some times during the period 1970 to 2019. SAMS observational locations are shown infigure 2.

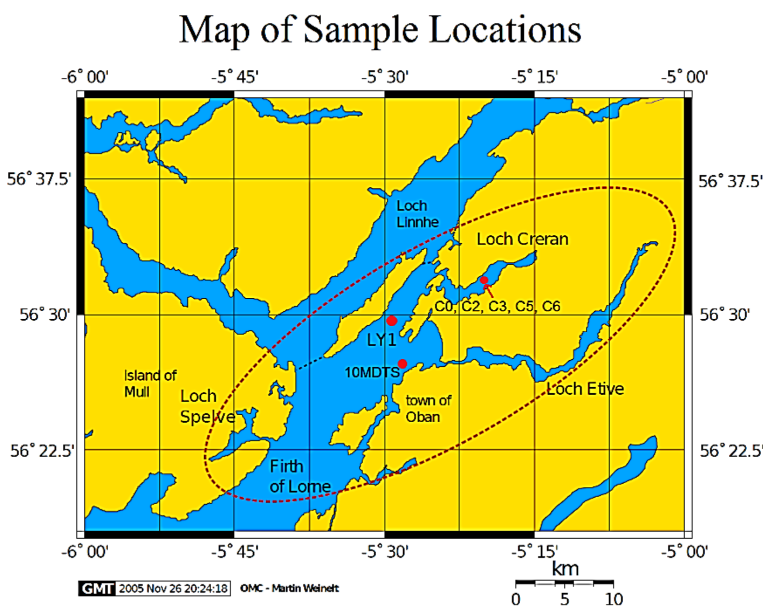


Figure 2, A map of typical West Coast Scottish sea-loch and sites where data was collected, site locations coordinates are in the appendix, map provided by Paul Tett using OMC Martin Weinalt amended by author to include additional points.

2. Methods

Acquiring data from ICOADS

ICOADS version 3 was used for this report, providing global coverage of various surface oceanic and near oceanic atmospheric variables on a 1°× 1° grid from 1960 onwards. (Freeman et al. 2017). The data were obtained by way of the National Oceanic Atmospheric Administration (NOAA), Boulder CO, Physical Science Division (PSD) website at http://www.esrl.noaa.gov/psd/. This linked to the ICOADS data download page under the 'gridded climate' section. The data were downloaded in netCDF format using anonymous FTP. 'Trimming' describes the extent to which apparent anomalies have been excluded: data trimmed at ±4.5σ has been used in this report. Python was used with the xarray module to extract the netCTD array data and output the SST time series in csv format for the specified grid square (centered on 56.5° N 5.5° W).

Compiling SAMS data:

Observed SST and salinities were compiled from various sources within SAMS. The ship-lowered CTD data derived from 1977 to 2017, with some years and months having better coverage than others There was a particular gap in the 1990s). Data were binned into 1 metre discrete depth intervals ranging from 0 m (i.e. near surface) to 10 m, and time assigned to the nearest hour. The observation site was also recorded (details in appendix), based either on shipboard GPS or a standardised triangulated point at sea).

The data itself came in a variety of forms. For older data (before 1980) hardcopies were manually digitized from binders/files or old reports. Obvious errors were either excluded or if possible corrected (e.g. salinity written as 312 interpreted as 31.2). Post-2000 data had higher temporal and depth resolutions and required further binning to match the older data. The mean of all values ±0.5 of the metre or hour were taken.

In addition to the ship data, a high-resolution continuous temperature record was available from a recorder on th sea-bed at a water depth of 10 metres. This was located at 56°27.280'N 005°24.815W from 1995 to 1997 and then at 56°27.313'N 005°26.814'W. This will be referred to as the 10 m Deep Temperature Series (10MDTS).

Processing data:

The compiled SAMS data was further processed so it could be compared to ICOADS monthly values. To do this the arithmetic mean of temperature data was taken and time intervals were approximated to the nearest hour. Stata 15.0 was used to write the scripts that extracted and binned the data to these intervals.

Once this was done all data (including salinities were available) were put onto a master spreadsheet along with metadata. The CTD data and 10MTDS were aggregated monthly and into various depth intervals so that they could be compared with ICOADs data.

Because the term 'SST' is ambigious, the CTD data were grouped into smaller depth intervals ('top' 0-2 m 'middle' 3-5 m and 'bottom' 6-10 m) as well as over the full 10 m range. Basic linear regressions were calculated for relationships between ICOADs , 10MTDS, and CTD data (both as a whole and grouped into depth intervals 0-2 m 3-5 m and 6-10 m), and significance assessed by means of . t-tests.

3. Results

CTD data from a total of 11 unique sources (spreadsheets, csv and text files, and hardcopies) were used. After grouping across depths and within hours there were 400 unique data points present within the CTD dataset. There was generally good coverage across the depth interval with a mean number of 8.09 depth observations (out of 11 possible depth intervals). The reason for lack of complete coverage was mostly due to coarse depth measurements (e.g. taking temperature at just 2 m and 8 m, or skipping every second metre), as shown in the depth histogram in the appendix. In contrast there were 88,000 unique hourly groupings for the 10MDTS data.

When the data were aggregated monthly, out of a total of 508 months between 1970 and 2019, the number of observations were as follows:

- 95 months contained observed SST data (from 1977 onwards),
- 435 contained ICOADs data ,
- 243 months had 10MDTS data (nearly continuous from 1995).

As shown in the appendix, the CTD data show biases towards oarticular times of day and months of the year. The 10MDTS data were more uniform. Since only the only values were available from ICOADS, it was not possible to check for temporal biases.

Figure 3 shows all data-series, with values from the entire series averaged to month. The figure shows the expected pattern of (pseudo) sinusoidal seasonal change in SST for all datasets. However, ICOADS reported comparatively warmer Summer, and cooler Spring, SST. The non-ICOADS datasets seemed more similar to each other than they were to ICOADS, although all datasets were in good agreement during Autumn and early Winter.

Figure 4 shows the deviation of annual means of the 10MDTS and ICOADS from the ICOADS long-term mean. These anomalies showed only moderate agreement.

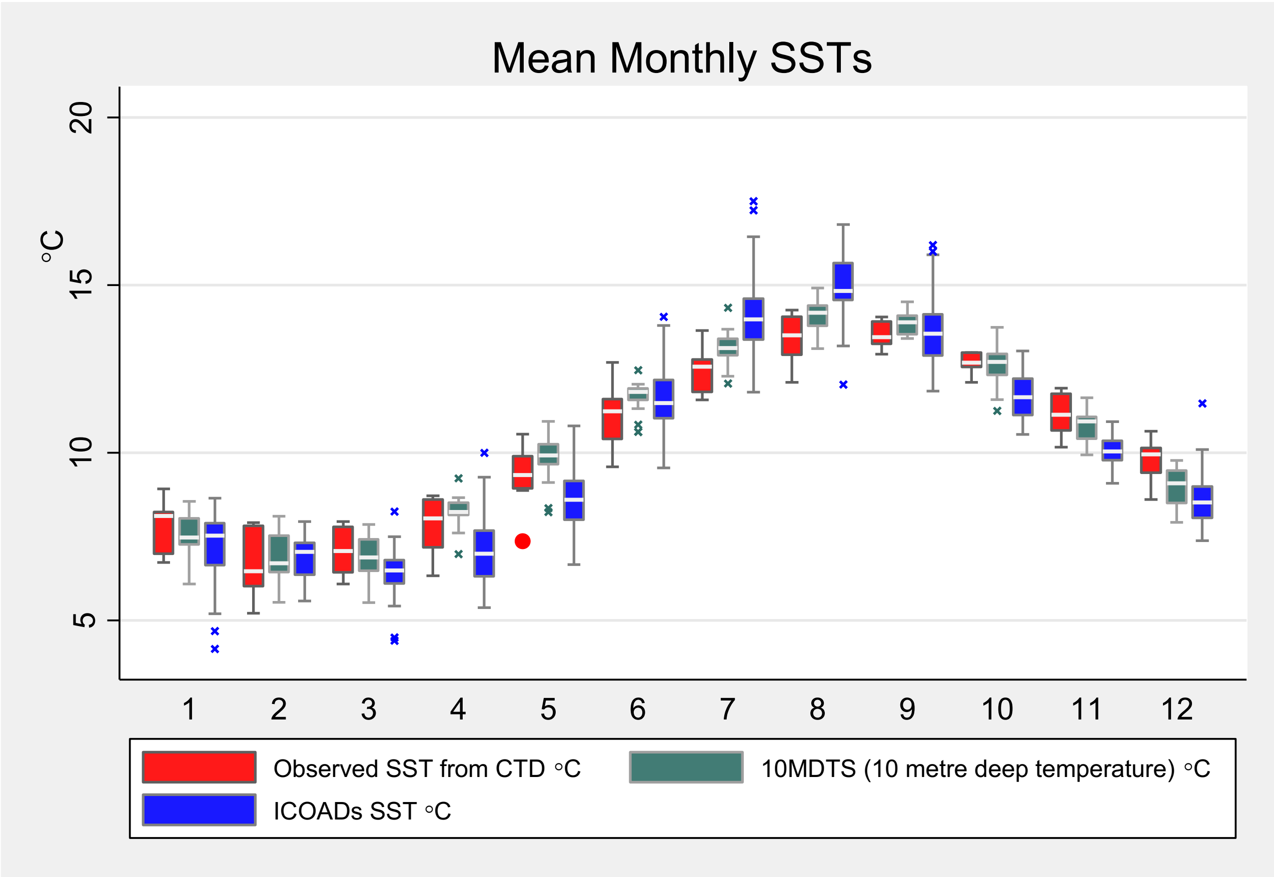


Figure 3, Box plot of mean monthly values as reported by each dataset. Greatest variability between datasets is in late springs While ICOADS is clearly warmer in summer. diagram plotted using Stata 15.0

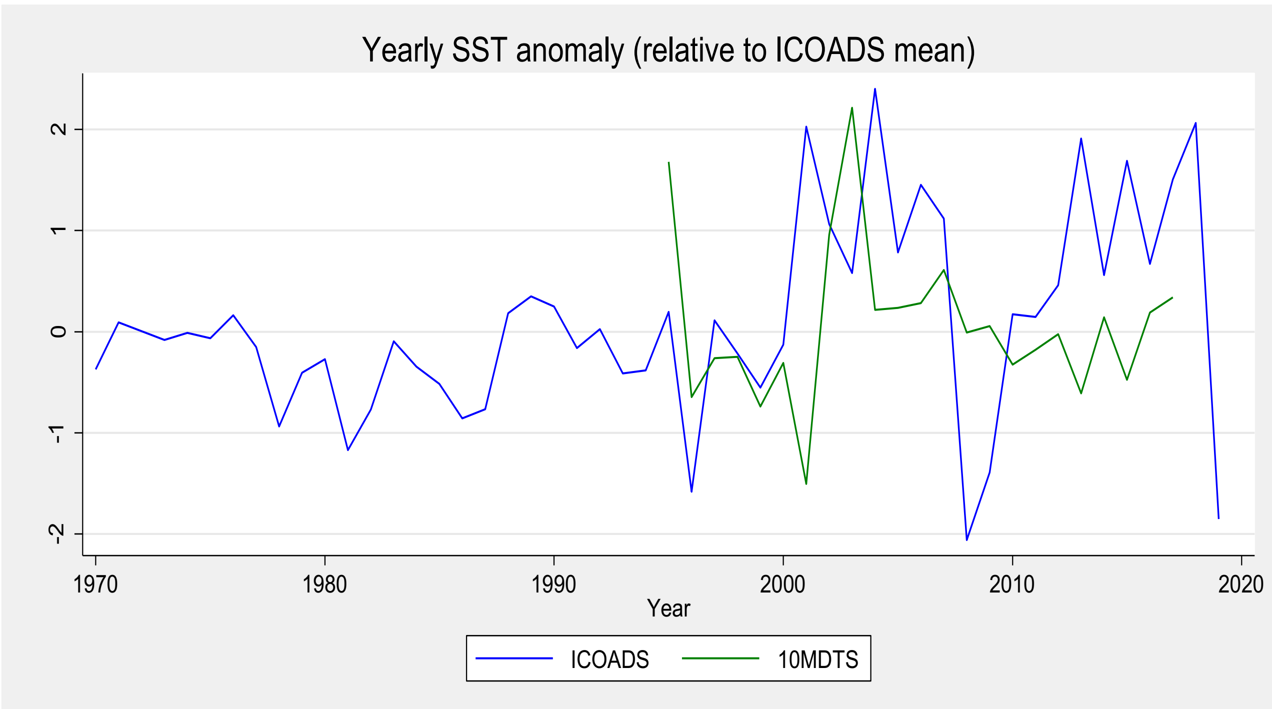


Figure 4, ICOADS and 10MDTS data compared in terms of yearly mean temperature with reference to ICOADS anomaly (i.e 0 would represent the mean average year across the study period). This is taken from existing monthly observations and no effort was made to exclude years where coverage is sparse

Figure 5 shows monthly values for all three datasets. The ICOADS data were more intermittent after 1995, but, fortunately, the 10MDTS series was available after this year to provide a complete picture. CTD data, when available, conformed with the ICOADS and 10MDTS data. The figure shows a trend to warmer Winters (i.e. higher minimum temperatures) over the neary 5 decades of the time-series.

Full details of an analysis of pair-wise regressions are given in the appendix. 10MDTS and ICOADS yielded both a relatively strong correlation R^2=0.90 and statistically significant P and t value. CTD data and ICOADS yielded a reasonable but weaker R^2 value of 0.86 when regressed as a whole. When analysed per depth groupings, correlation were weaker, although with 'top' CTD data correlating better than 'bottom' CTD data with ICOADS increments being generally stronger than the bottom increments.

The slope of both regressions (i.e. ICOADS on 10MDTS and ICOADS on CTD) was greater than one, confirming what had been seen in figure 3: that the ICOADS showed a larger temperature range than the two inshore data series. The regressions, simplified, were:

$$(ICOADS) = 1.10(MDTS) - 0.72$$
{equation 1}

$$(ICOADS) = 1.18(CTD) - 1.73$$
{equation 2}

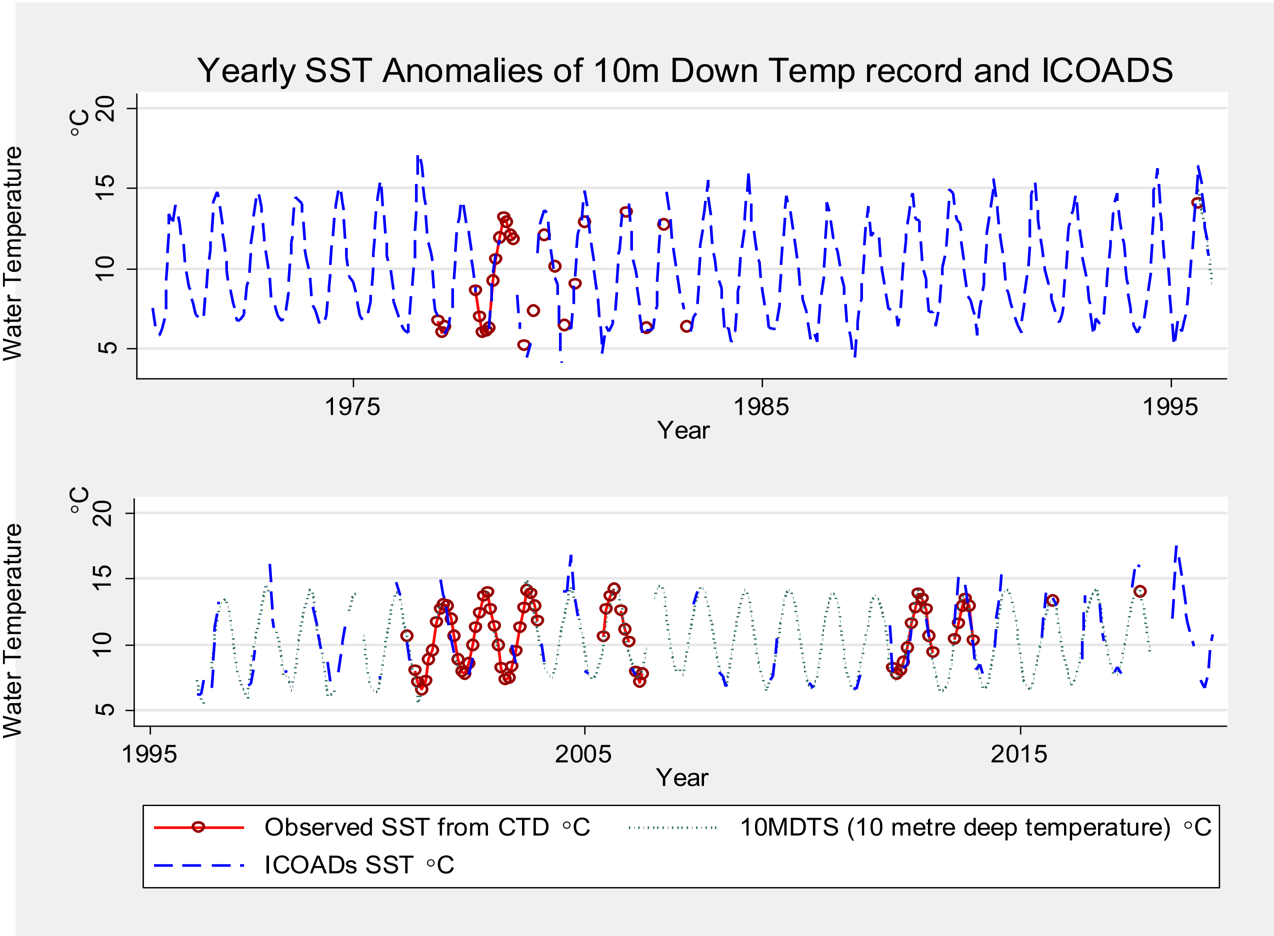


Figure 5, monthly values of all three time-series. (Heading incorrect and should be ignored) Diagram plotted using Stata 15.0

4. Discussion

Based on the results of this study, ICOADS and 10MDTS data agree best, even if the 10 metre data might not be properly called SST. The statistical significance of the correlation between ICOADS and CTD data was weaker, perhaps unsurprising given the lesser quantity of the CTD data. Furthermore, only 47 of the 91 months where there were CTD data coexisted with ICOADS observations, as a result of ICOADS data becoming scarcer after 2000 (figure 5). In some cases the CTD monthly values were based on a few observations that might have been unrepresentative on the month as a whole (given both seasonal trends and day-to-day variation in heating and runoff). Finally, there was also a time-of-day bias, which may have led a greater systematic error in CTD data compared to the 10MDTS data, even if CTD data is closer to a 'true' SST.

Another major factor to consider is the lack of information about the observations within the grid cell that is being used from the ICOADS dataset. The 1960-2014 time period saw ICOADS data going from predoinantly ship-based observations to more observations by data Buoy (NOAA); it isn't clear if this happened within the Firth of Lorn grid square. Efforts have been made in these internation data bases to adjust for the different biases in different types of measurement. Kent & Berry (2011) suggest an uncertainty of ±0.1 to ±0.15 C associated with bucket method of SST measurement by the Volunteer Observing Ships. Older data from ships' engine rooms shows a slight warm bias. It was also suggested that error was underestimated in 'high variability regions' such as those on the fringes of the North Atlantic. Finally, there is the unaswered question concerning where within the grid square each year's data came from. However, as the highest correlations between CTD and ICOADS data were obtained with CTD data either from the 'top' depth interval or from the whole of the top 10 metres, this suggests a similar pattern of collection of the raw data entered into ICOADS.

Alhtough there was good overall agreement amongst the data-sets, the results show the seasonal bias already mentioned, with ICOADS reporting higher Summer temperatures (figure 3). Better agreement between ICOADS and the other two data sets was found for Autumn data, despite Autumn being a time of seasonal change where factors mentioned in the first paragraph of this discussion might be expected to cause the CTD data to diverge from the other two. This may reflect the different environments in which the data were collected (assuming ICOADS data were observed in more oceanic waters). Terrestrial runoff might be cooling inshore waters in Summer in this mountainous region.

5. Conclusion

Although the correlations were not perfect, and although there are a number of unaswered questions about the ICOADS data, this study has shown a good agreement between the values as well as the patterns of SST recorded in all three data sets. Theere were persistent differences between the ICOADS and inshore data sets, but these can largely be corrected for using equations 1 and 2. What this means is that ICOADS data can be used as an estimator of the inshore SST time-series that are needed for relating to plankton data. Furthermore, it should be possible to augment ICOADS data with information from the 10 m sensor, as the ICOADS data series grows more sparse. Finally, the results reported here are not the end of analysis, as there is more inshore CTD data available to be digisitised. Thus it should be possible to improve the reliability of equation 2, in particular.

Acknowledgments

I would like to acknowledge and thank Martin Sayer for providing the 10 m deep continuous data (10MDTS) and associated funders that enabled him to carry out this research. I'm grateful to Paul Tett and Stuart Cuningham who advised and guided me through writing this report. At last this report would not have been possible without the ICOADS data provided by NOAA's PSD, accessed at: *http://www.esrl.noaa.gov/psd/* .

References

Freeman E., Woodruff S.D., Worley S.J., Lubker S.j., Kent E.C., Angel W.E., Berry D.I., Brohan P., Eastman B., Gates L., Gloeden W. ,Zaihu J., Lawrimore J., Rayner N.A., Rosenhagen G., Smith S.R. (2017) 'ICOADS Release 3.0: a major update to the historical marine climate *record' International Journal of climatology*, vol. 37 pp. 2211-2232.

Kent E.C., Berry D.I. (2011) 'Air–Sea fluxes from ICOADS: the construction of a new gridded dataset with uncertainty estimates' *Royal Meteorological Society*, vol. 37(7), pp. 987-1001.

Wells N.C. (2016) "The North Atlantic Ocean and climate change in the UK and Northern Europe" *Weather, RMetS* vol. 71(1) pp.3-6.

Tett, P. (2014). Net microplankton production in loch Creran and its approaches in September 2013. Internal report 286, Scottish Association for Marine Science, Oban, Scotland.

Anonymous, (accessed 2019-08-07) 'ICOADS Brochure' NOAA PSD, Boulder CO. https://icoads.noaa.gov/

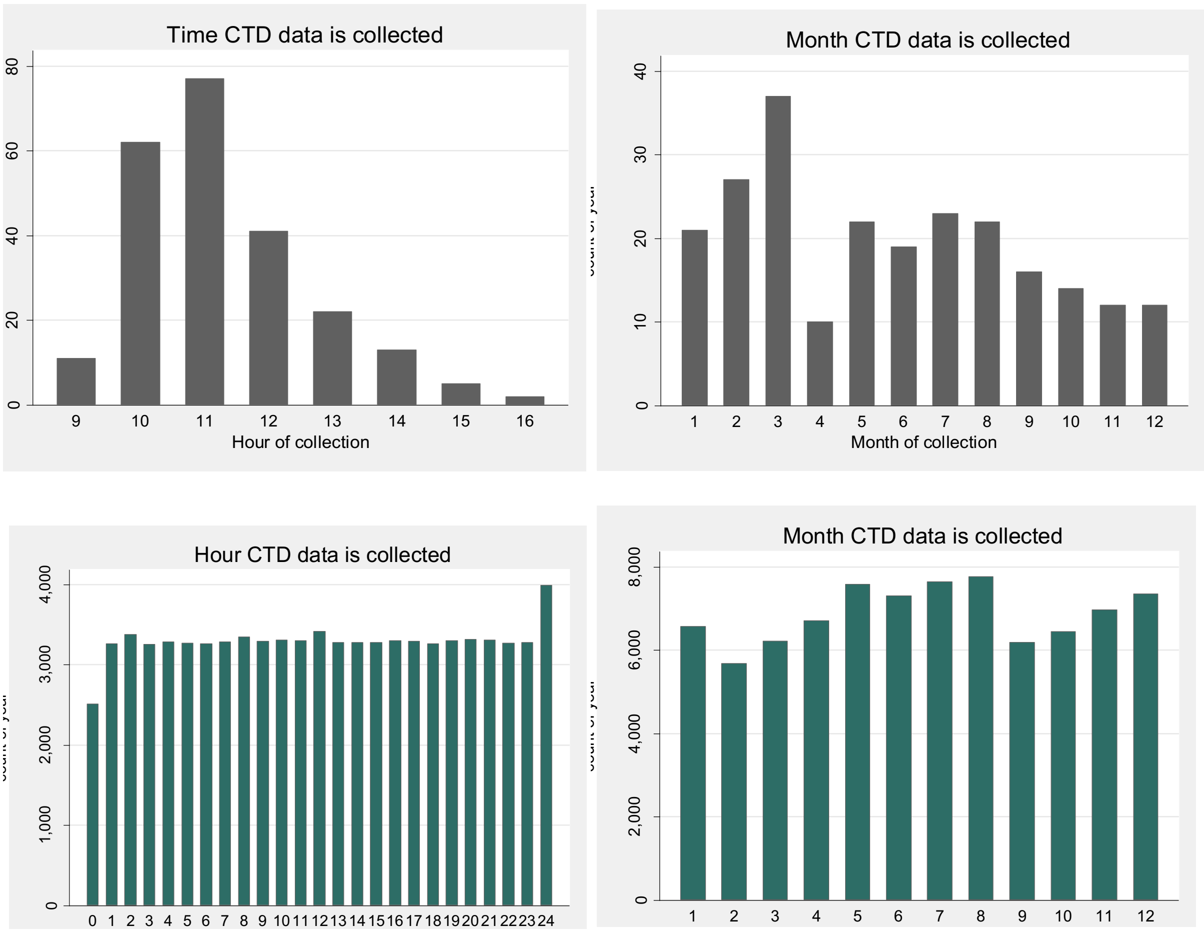
Tett, Paul (2019). Long-term changes in Phytoplankton Lifeforms and SST at the Lorn Pelagic Observatory, v.1. SAMS Internal Report 296, Scottish Association for Marine Science, Oban, Scotland.

Appendix

Sample site locations and naming convention

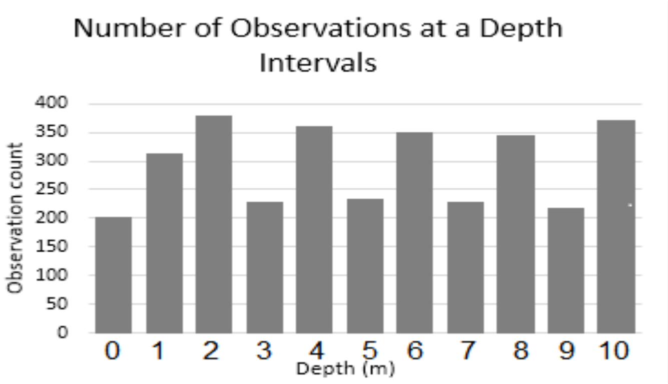
Sample site Name	Sample site Coordinate
LY1	56.479° N, 5.502° W
C0	56.532° N, 5.435° W
C2	56.522° N, 5.368° W
C3	56.518° N, -5.368° W
C5	56.532° N, 5.330° W
C6	56.521° N, -5.230° W
ICOADS	56.5±0.5° N, -5.5±0.5° W
10MDTS	56.453° N, -5.539° W

Time collection histograms for the datasets created for this report



Histograms showing the time data was collected for CTD compiled datasets

Counts per depth interval in observed dataset

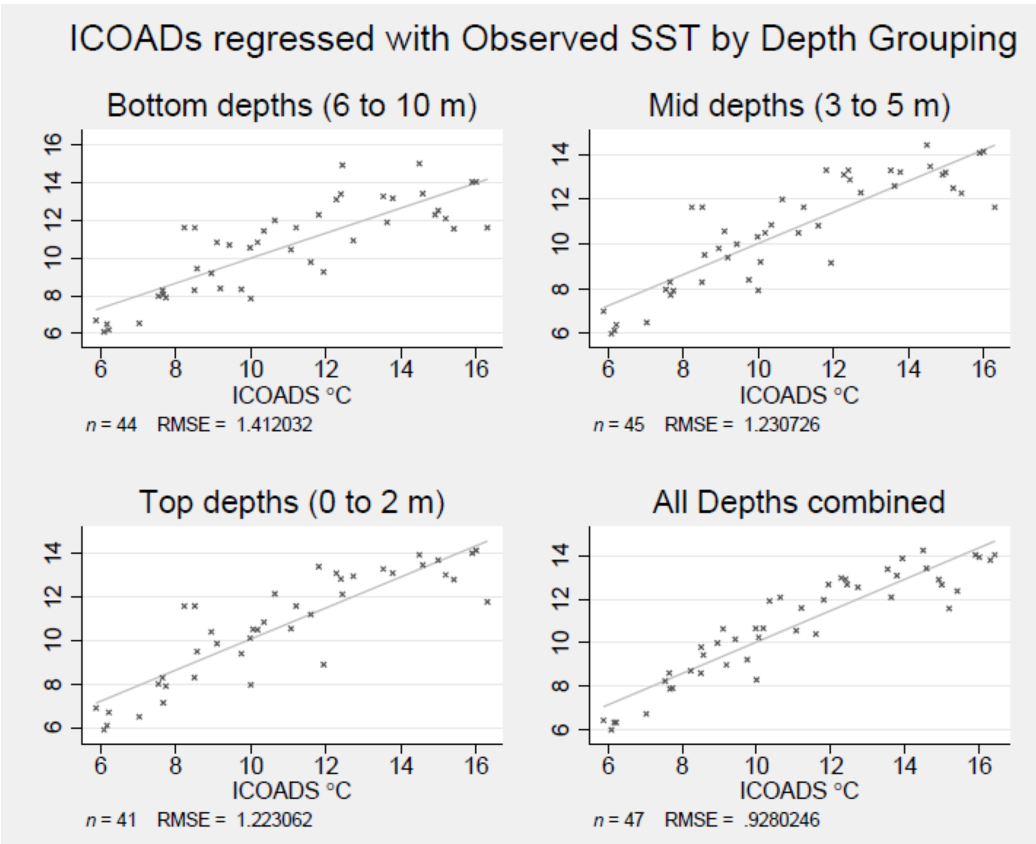


Names of data files used or created for this report

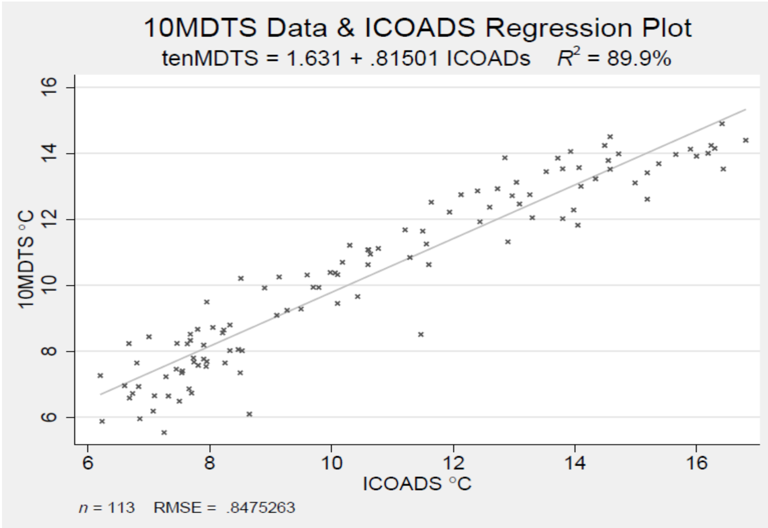
For All monthly aggregated data	ICOADs&obvs.csv
Just ICOADs data	ICOADs.csv
All none ICOADS data hourly	SAMSmasterspread.xlsx
Script for extracting/process10MDTS data (stata)	95-18sst.txt
script for extraction from netCTD to csv (python)	ICOAD_cell_extractor.py
Converts Mastersheet_hourly_to_monthly data (stata)	Mastersheet_hourly_to_monthly.txt
General script for processing newer data (stata)	ASCII_2013_2015_2017.txt & other txt files (general script used on all newer data)

*copies of files can be obtained from Harry Colvile or Paul Tett

Regression plots



*note all regressions are lower for certain depth intervals (bottom= 67.9% mid=74.5% top= 76.0% whereas combined it is 85.7% and thus the bottom chart shows only



10MDTS regression plot (just one detph itnerval)

Regression plots for CTD SST data and ICOADS

Source	SS	df	MS
Model	960.570925	1	960.570925
Residual	107.910155	111	.972163561
Total	1068.48108	112	9.54000964

Number of obs =	113
F(1, 111) =	988.08
Prob > F =	0.0000
R-squared =	0.8990
Adj R-squared =	0.8981
Root MSE =	.98598

ICOADs	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
tenMDTS	1.103057	.0350916	31.43	0.000	1.03352	1.172593
_cons	-.7216727	.3740195	-1.93	0.056	-1.462817	.0194719

Equation 1 regression data

Source	SS	df	MS
Model	378.785283	1	378.785283
Residual	63.4207124	45	1.40934916
Total	442.205995	46	9.61317381

Number of obs =	47
F(1, 45) =	268.77
Prob > F =	0.0000
R-squared =	0.8566
Adj R-squared =	0.8534
Root MSE =	1.1872

ICOADs	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
obvs_sst	1.183952	.0722182	16.39	0.000	1.038497	1.329407
_cons	-1.733385	.7995379	-2.17	0.035	-3.343737	-.1230331

Equation 2 regression data