SOURCE	COMMENT	RESPONSE	AMENDMENT TO M/S
R1			
C1	There are no bias estimates between the individual data sources. Is there a plan to address that? For long-term data sets small biases between instruments can cause problems, since over long periods the climate-change signal is also small. Therefore, relative instrument bias should be estimated and if possible, corrected.	There are many different sampling methods used and a very large number of data points collected in different manners. We do not believe that there is bias in the majority of this data. There is a plan to follow up this paper with a statistical analysis of combinations of multiple, different, data sources across the time and spatial scales presented. One area where bias may be a significant factor is Source 1. The heat island effect (more apparent in atmospheric temperature records) is not relevant to the majority of the data. However, there are stations of the , the Coastal Temperature Network that are associated with power station developments and some of these may, potentially, be affected by the thermal plume of the discharge water. The section of the paper on the Cefas Coastal Network (Source 1) and the associated metadata and references makes the sources clear. Numerically, these data only make up a small section of the paper.	Section 3.9 *Data Ingestion Quality Control* will be renamed 3.9 *Data Ingestion, Quality Control and Bias Estimation* with the current text under 3.9.1 Data Ingestion and Quality Control plus a new section on Bias estimation added. *3.9.2 Bias Estimation* *The provision of these raw data is 'as measured' with appropriate metadata to allow subsequent scientific trend analysis to be performed which would usually include additional scrutiny for systematic bias. The main exercise here is to identify and facilitate access to a large source of hither-to unavailable data that is, as yet, unseen and unscrutinised by the broader community.* *An assessment of accuracy and bias has been conducted by the data creators for some of the sources included here. For example, for Source 10, we referenced Wright et al. (2016) who examined whether the temperature data derived from 100s of recreational SCUBA divers, and many different models of dive-computer, were consistent with global sea temperature datasets. Similarly, temperature sensors on Cefas SmartBuoys and WaveNet platforms (Sources 6 and 7) are calibrated annually at Cefas against certified platinum resistance thermometers. Data are subject to a full quality assurance procedure which assigns flags to poor quality data (e.g. for sensor malfunction, drift) [see https://www.cefas.co.uk/cefas-data-hub/dois/cefas-smartbuoy-monitoring-network/].* *We note that ICOADs and other collated datasets (e.g. HadSST) tend to carry out their own systematic bias correction routines whenever new data are uploaded/admitted. Our intention is to make our data available so that it can be easily included (by other authors) in platforms such as the ones listed (ICOADS, COBE-SST, ERSST and HadSST3). Within the text of the manuscript we already include reference to papers that discuss bias correction (e.g. Mathews 2013; Kennedy et al. 2011a, b; Karl et al. 2015; Hausfather et al. 2017), but we leave it to those who might make use of the data, to judge what procedures might be necessary for their own purposes.*
C2	The North Atlantic is a densely-sampled region, so it is not clear how much additional information these new data bring to the region. It would be useful to compare a field such as SST from ICOADS and from a combination of ICOADS and the new data to see how resolution of the field is improved for a few decades. It may also be useful to evaluate a subsurface field compared to an existing subsurface data set. If clear improvements can be shown, that would better justify the use of these data and also justify a project to incorporate the new data into existing larger data sets such as ICOADS.	Whilst the suggested activities are under consideration at Cefas, the aim of this Data Paper is to provide the data to enable other people to do novel and useful things, not to do them in the paper. We note the Journals explicit aim of "*furthering the reuse of high-quality data of benefit to Earth system sciences*". The gridded data sets are certainly at their strongest in the wider North Atlantic but, given the standard gridded 1 degree scale compared with the scale of shelf and coastal oceanographic processes, the value of our data may be at its greatest in analysis of change in particular locations. We cite, for example "*The full dataset also facilitates the construction of long-term temperature time series and an examination of changes in the phenology (seasonal timing) of ecosystem processes*".	None

C3	Are the authors in contact with any organizations producing long-term ocean data sets, such as ICOADS or the Met Office? Much of the new data could be incorporated and has the potential to improve resolution of the North Atlantic.	Cefas has completed and is engaged in a number of historic data recovery processes involving data archaeology and the recovery of information from often long-dormant systems. Over the period of our data collections, links have been in place to the broader community, particularly via the International Council for the Exploration of the Sea (ICES) and, where relevant, to the UK Met Office Hadley Centre and the British Oceanographic Data Centre. We have contacts into these programmes and can explore whether they would be interested in adding our data. Scientists working on ICOADS, COBE-SST, ERSST and HadSST3 may be unaware of the wealth of data existing within government agencies such as Cefas. Unlike in the US, where coordinated programmes such as ICOADS are run from within government (in this case by NOAA), such programmes tend to be much more disparate in Europe, with a clear and historical separation between institutes responsible for fisheries management and those that deal with meteorology and oceanography (e.g. Cefas and the Met Office in the UK). The contacts exist and are being utilised, facilitated by this assembly of Cefas seawater temperature data.	None
R2			
C1	Major points: 1. Unfortunately the MS merely makes a very limited basic statistical analysis and, although it is not the purpose of doing an oceanographic analysis, lacks a minimum link with the oceanographic features of the area. It is very important for the reader to understand the changes in temperature, especially for those not familiar with the study area. I suggest including a small chapter to explain the basic oceanographic characteristics.	The physical oceanography of the wider North Atlantic is now referenced in the Introduction (ICES Report) as is a Feeder Report to *UK Charting Progress 2* to which Cefas and several authors actively contribute. Sections on The Greater North Sea and The Celtic Seas are now included, as are minor edits to the original m/s to ensure consistency.	Added to Introduction: The International Council for the Exploration of the Sea (ICES) produces an annual report on the marine climate of the North Atlantic (the ICES Report on Ocean Climate). This gives a broad description of the oceanography of this region and documents the year-by-year variations using a set of hydrographic stations collected by the international community (Larsen et al., 2016). They describe the variation in the northern North Atlantic and sub-Arctic Seas where the North Atlantic current provides a source of heat and salt along the eastern margin into the Barents Sea and entry to the Arctic Ocean. Along the western margin the Arctic influence of cold and fresh conditions extends from the Fram Strait to Cape Farewell. At the southern part of the region covered by Cefas from the western channel down to Iberia the influence of subtropical waters is more evident. The combination of gyres and the North Atlantic current places the UK shelf waters at the boundary between temperate and subpolar waters exerting a heavy influence on the variability of conditions in the Greater North Sea and Celtic Seas. **The Greater North Sea** The temperature of the Greater North Sea is controlled by the seasonal cycle of heat exchange with the atmosphere, the vertical mixing in the water column and the circulation of waters from the North Atlantic. The annual mean temperature generally increases from the south (in the English Channel) to the north (near Shetland), this pattern is not representative of all seasons. During the winter the shallow waters in the southern North Sea that are furthest from the influence of the inflowing North Atlantic waters tend to be the coolest in the entire Greater North Sea.

			Northern North Sea: Modified Atlantic water flows into the region via the Fair Isle current maintaining relatively warm winter temperatures, typically 6 to 9 °C minimum, decreasing to the south as water from the Atlantic is cooled by atmosphere and depth shallows. Summer temperatures are typically 12 to 14 °C near the surface with a cooling influence evident from the North Atlantic inflow and it generally stratifies. *Southern North Sea*: The Southern North Sea is shallow, mostly less than 50 m in depth, and furthest from inflows and influence of Atlantic water. Temperature minima in winter are typically 4 to 8 °C; they depend strongly on the weather in any one year, and on depth (shallower->cooler). Likewise, the typical summer maxima 16 to 19 °C depend on the weather and strongly on depth (shallower->warmer) *English Channel*: From depths of less than 50 m near the coast and Dover Strait the Channel deepens westwards to 100 m. Influence of Atlantic water also increases towards the west and only some parts in the very west stratify in the summer. Thus minimum winter temperatures, typically 5 to 8 °C, are strongly dependent on the weather in any one year and on depth. Summer maximum temperatures are typically 16 to 19 °C. The Greater North Sea near bottom temperatures differ from SST due to stratification which takes place only during the summer. Where the region does stratify (in the northern North Sea and at the very western part of the English Channel), summer temperatures near the bottom remain cool until the breakdown of stratification in the autumn. **The Celtic Seas** The various temperature and salinity characteristics Celtic Seas are reflective of the inhomogeneity of the region, from enclosed shallow shelf sea with large river catchments all the way through to deep oceanic waters and across a wide range of latitude. Surface temperature is controlled by a balance of seasonal heating, vertical mixing and the circulation of Atlantic water, the relative importance depending on local depth, tides. wind and exposure to the ocean. *Celtic Sea*: - Sea temperatures are strongly related to the weather in any one year and to water depth. The climate being strongly maritime, typical winter minima are 8 to 11 °C and summer maxima 14 to 18 °C. The seasonal cycle of near bed temperature in this part of the region is controlled by the vertical mixing. When well mixed vertically in the winter its temperature is similar to that at the surface typical winter minima are 8 to 11 °C . During the summer the area stratifies and near bed temperatures do not reach the temperature maxima of the surface, the maximum annual temperature here is reached typically in October when the heat of surface waters is fully mixed down. *Irish Sea*: - Temperatures depend strongly on the weather in any one year and on water depth. Typical winter minima are 4 to 8 °C and summer maxima 14 to 18 °C. As elsewhere, temperatures depend strongly on the

			weather in any one year and on water depth and on whether the area stratifies. The area is well mixed vertically in winter and typical winter minima match the SST at 4 to 8 °C. In the areas that stay well mixed throughout the year summer maxima 14 to 18 °C are typical while areas that stratify in the summer reach their annual maximum of 13 to 15 °C in autumn when the heat of surface waters is fully mixed down. *Minches and Western Scotland*: - There is some influence of (modified) Atlantic water arriving from the west. Resulting typical winter minimum temperatures are 6 to 8 °C and summer maxima 13 to 15 °C. Typically, there is summer stratification in deep waters away from islands and north of the Islay front (west of Islay to Ireland). There is some influence of (modified) Atlantic water arriving from the west. Resulting typical winter minimum temperatures are 6 to 8 °C and summer maxima 13 to 15 °C in well mixed areas or 11 to 13 °C where stratified. *Scottish Continental Shelf*: - Except for shallow areas near coasts, there is summer stratification. Temperature minima in winter are typically 9 to 10 °C at the shelf edge but 6 to 9 °C elsewhere; they depend on the weather in any one year, on depth and on travel time for any Atlantic water arriving from the shelf edge. Summer maxima are typically 12 to 14 °C for surface water. Except for shallow areas near coasts, there is summer stratification. Temperature minima in winter are typically 9 to 10 °C at the shelf edge but 6 to 9 °C elsewhere; they depend on the weather in any one year, on depth on travel time for any Atlantic water arriving from the shelf edge. Summer maxima are lower than the 12 to 14 °C maxima seen in the surface water.
C2	I suggest including a table with the climatology of the area. Since the area considered is broad and certainly has different climatological values (min, avg, max) between subareas, I recommend to divide it in squares as done in the Medar/Medatlas database and in: Manca, B., Burca, M., Giorgetti, A., Coatanoan, C., Garcia, M.J., and A. Iona, 2004: Physical and biochemical averaged vertical profiles in the Mediterranean regions: an important tool to trace the climatology of water masses and to validate incoming data from operational oceanography. J. Mar. Syst, 48, 83-116. It could be interesting to see what is the variance of temperature records respect to climatology in different periods	This is an interesting suggestion for our future work (see Response to R1C1 where we plan to investigate how multiple sources and varying amounts of data across time and space are best taken into account). The combined data set, by its very nature, has varying temporal, vertical and horizontal resolution, including some very high-resolution data at 1m vertical resolution. These raise a number of questions that could be asked of such a data set, such as the optimal resolution required to sample certain features, or the ones that the referee proposes. However, they go beyond the scope of this data paper. We note our response to C1 above which provides some information on subareas. We feel that the nature of the data provided precluded any attempt to produce a gridded climatology from this dataset as a single analysis. Some of the data sources included here have produced and published seasonal cycle and long-term climatologies (see e.g. the Cefas CTN dataset in Jones and Jeffs 1991 or Joyce 2006) or contributed to wider assessment reports such as Charting Progress 2 (UKMMAS, 2010) or the ICES Report on Ocean Climate (Larsen et al 2016). A gridded climatological analysis of these areas including the data made available in this paper would be useful next step, done in conjunction with other data sources. This and the full analysis needed is beyond the scope of this current data paper. There are papers and products of this type already	None

		available for parts of the area that do this job well. They also demonstrate the need for specialist analysis where making the climatology or gridded product the specific the purpose of the paper. For instance see the Berx and Hughes (2009) climatology or the Adjusted Hydrography Optimal Interpolation (AHOI) Dataset which is a gridded product based on collected *in-situ* data in the North Sea (Núñez-Riboni and Akimova, 2015) and is available from https://www.thuenen.de/en/sf/projects/ahoi-a-physical-statistical-model-of-hydrography-for-fishery-and-ecology-studies/ References Berx, B. and Hughes, S.L. (2009) Climatology of Surface and Near-bed Temperature and Salinity on the North-West European Continental Shelf for 1971-2000. Continental Shelf Research, 29, 2286-2292. Jones, S.R. and Jeffs, T.M. Near-surface sea temperatures in coastal waters of the North Sea, English Channel and Irish Sea. Data Report, MAFF Directorate of Fisheries Research, Lowestoft 24: 70 pp, 1991. Joyce, A.E. The coastal temperature network and ferry route programme: long-term temperature and salinity observations. Science Series Data Report, Cefas, Lowestoft 43: 129 pp, 2006. Núñez-Riboni, I., and Akimova. A. (2015) Monthly maps of optimally interpolated in situ hydrography in the North Sea from 1948 to 2013. Journal of Marine Systems, 151, 15-34.	
C3	I'm a little bit concern about the QC method applied, since you considered only part of the standards QC flags for temperature. Please justify why and see: SeaDataNet, 2007. Data quality control procedures IOC/IODE, 1993. IOC Manual and guides No26, Manual of quality control procedures for validation of oceanographic data among others.	Given the wide variety of sources and, in a lot of cases, a non-physical oceanographic focus for the data generating activities, a formal and rigid retrospective application of oceanographic data quality control procedures was not applied across the board. However, where appropriate at source, e.g. the CTD and ScanFish data (Sources 03 and 16), the standard IOC methodology has been applied. Therefore, in the manuscript, the descriptive terminology is more general. However, the areas covered by the SeaDataNet standard QC were covered as part of the ingestion of the data. Specifically, basic checks for all data types: • Date and time • Latitude and longitude • Position must not be on land as were other relevant checks, e.g.: • Impossible speed • Spike • Global range • Regional range • Check for duplicates along with other sense checks as described in the manuscript (which cover, for example, reasonable domain, logical edits, loose background	Clarified in the text of the appropriate section on QC (3.9) inserting: Given the wide variety of sources and, in a lot of cases, the non-physical oceanographic focus for the data generating activities, a formal and rigid retrospective application of oceanographic data quality control procedures was not applied across the board. However, where appropriate they were applied at source, e.g. the CTD and ScanFish data (Sources 03 and 16). In both cases the relevant standard IOC methodology was applied. For the remaining sources, the descriptions above cover the intent of such standards, specifically, basic checks for all data types, e.g. *Date and time*, *Latitude and longitude*, *Position must not be on land* as well as other relevant checks, e.g. *Impossible speed*, *Spike*, *Global range*, *Regional range* and *Check for duplicates*.

		checks on reported values, constant value, spike & step checks, track checks, exclusions and final multi-level checks).	
C4 Minor	Minor corrections: I. Page 4-5 there are two introductions in the ms, II. please correct the numbering Page 27: Figure 14 Average (green), minimum (blue) and maximum (red) annual temperatures for the Southern North Sea including all sources and all depths	I. Accepted II. Accepted	I. Amended II. Amended
R3	The authors are to be congratulated for their efforts to make these previously inaccessible observations available to the public. For the type of long-term record considered the quality of individual observations is not the key issue: what is important is that the observations are made accessible, and as much information as possible is made available to enable users to interpret the observations. The temperature observations in this collection should find wide application. As noted by Reviewer 1 the data can be compared with existing compilations such as ICOADS and the data products derived from it, and some of the data sources could provide valuable quasi-homogeneous reference for evaluation of bias adjustments applied to historical SST (http://journals.ametsoc.org/doi/pdf/10.1175/BAMS-D-15-00251.1)	Thank you for the statement and for the detailed comments. Addressed in responses to Reviewer 1.	N/A See Reviewer 1 changes
C1	Firstly, the decision to consider only temperature observations. Those working with observations over land are now trying to integrate observations of different parameters that have been historically separated (http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-16-0165.1) The observations in this collection come from a wide range of sources, and will have many different associated parameters recorded and it is clearly a major task to have extracted any information from the sources. However, much of the effort has gone into gathering information, like locations, times, and platform metadata, that is common across all observations. Co-located observations of additional parameters (e.g. salinity) can unlock new applications (and are often sparser than temperature observations so could be considered even more valuable), and information on ambient environmental conditions, whilst useful in its own right, can help to interpret the temperature observations and improve uncertainty estimation. The temperature collection is valuable already, and whilst not a barrier to issue and publication, the authors should consider whether future releases should include a wider range of the parameters recorded.	A key scientific driving force behind this exercise is the requirement from Cefas scientists for ready access to physical data for their ecological and assessment work. Seawater temperature is the most common physical parameter measured and deemed highly relevant to a lot of Cefas research. Temperature is the most widely and easily collected parameter we have. Physical oceanographic studies, where, e.g. salinity is highly relevant, have collected and utilised the co-located data at source. There are plans to continue the publication of the publicly funded scientific datasets held by Cefas through the Cefas Data Hub. Currently, datasets are made available in their original file formats with accompanying metadata but efforts are underway to provide data in similar form to that provided in this paper, recognising the increasing need for machine-to-machine interoperability through an increased use of, for example, Open Data, FAIRdata Principles, international parameter codes and Linked Data. We note that, in the main, seawater temperature data are inherently accurate enough to be useful in most non-specialist applications and reflect, seasonal and depth related changes. More general sensors for, say, salinity are, unless specifically designed and calibrated, less able to reflect changes. For example, the salinity range at the Gabbard SmartBuoy site is 32.5 – 34.5 psu and the sensors are not calibrated to the same level as accompanying water bottle samples (0.002 psu). The cross calibrations deployed, or not, plus the relative paucity of widespread salinity data encouraged our focus on temperature. Extraction of our co-located data can now be driven by end-user priorities led by the more widely available seawater temperature data.	The following paragraph will be added at the end of Section 1: *This paper focuses on seawater temperature data but we recognise the value of assembling and publishing co-located data such as salinity and, for example, in the case of the plankton dataset, the presence of species, amongst other parameters. The Cefas Data Hub currently holds published data in source formats with the intention of making the these and other datasets more accessible by transformations similar to those executed here.*
C2	Secondly, the approach to suspect observations is problematic. The default position is that data should be retained and flagged, rather than	We agree with the theoretical position that all data should be retained and flagged and, for example, in the case of the SmartBuoy and WaveNet	None

<table>
<tr>
<td></td>
<td>removed. This enables other researchers to apply their own quality assurance procedures, to understand the uncertainty they might expect to see in the data, and perhaps to apply corrections and adjustments. Any quality assurance process has some subjective element, and the removal of observations is a barrier to improved quality assessments in the future. The authors should re-instate, and flag, observations regarded as erroneous or suspect. This could be done by releasing a second dataset version aimed at the expert user (as ICOADS does with total and final versions).</td>
<td>data (Sources 6 & 7) this is the case, along with relevant QC flags covering, for example, drift and optical sensor degradation with time (fouling). However, a significant number of the data sources are not part of our operational database systems and hence are subjected to less formalised and more individual quality control and data handling processes.

The operational database systems are being connected to the Cefas Data Hub with a view to allowing user interrogation. However, current efforts are focused upon the provision of access to quality controlled data not raw data and associated flags. Access for the mooted expert user is probably better obtained by direct contact with the Cefas Data Manager and the relevant database managers, enlightened by the data published here.

As recognised in the Comment, "Any quality assurance process has some subjective element". In our case we choose a balanced approach taking into account the nature of the sources, the available metadata & data (including numbers and formats of files) and available resources.

"In any QC system there is a balance between trying to reject all "bad" observations and retail all "good" ones – different users might require a different balance" – see Ingleby & Huddleston, 2007, Quality control of ocean temperature and salinity profiles — Historical and real-time data (Journal of Marine Systems, Volume 65, Issue 1-4, p. 158-175. DOI: 10.1016/j.jmarsys.2005.11.019).

The above provides the context for our balance. Cefas recognises the potential for different user needs and is addressing those (see above). In practice, detecting "impossible" data (see R2C3 above) occurred at source, during original activity post processing, at raw file ingestion and inside the collated dataset. As described in the manuscript, suspect data were also eliminated at station level if other quality control checks indicated, e.g. data on land.

We note that there are other, probably more significant, features of this type of "archaeological data reconstruction", specifically the time and place and number and accuracy of the data from individual sources had and has meaning, the combination does not and is a source of potential doubt about the value of any interpretation that exceeds that of missing "impossible" data (see responses above).</td>
<td></td>
</tr>
<tr>
<td>C3 Minor</td>
<td>1. Introduction, line 3. It seems unlikely that "most, if not all" marine researchers have been measuring sea temperatures
2. Page 5, line 13. A better reference for GCOS and ECVs would be: <u>http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-13-00047.1</u>
3. Page 6, paragraph starting on line 10: ICOADS, HadSST3, ERSST and COBE all require appropriate citation in the reference list.
4. Page 6, line 10. assimilate is probably not a good word to use here as none of these products use the technique "data assimilation". Also ICOADS is a data archive (unless the authors</td>
<td>1. Will change "most, if not all" to "a significant proportion of"
2. Accepted
3. Accepted
4. Accepted
5. Dealt with in responses to Comments above.
6. These data arose from an early part of a wider programme. They were not formally written up. The following is extracted from relevant project documentation and dataset metadata (see link below).</td>
<td>1. Change "most, if not all" to "a significant proportion of"
2. Reference added
3. Links provided
4. Assimilate amended to ingest and "data syntheses" added to recognise the last point
5. None
6. From Modelled depths are <2 metre when not inundated. To Modelled depths are <2 metre when not inundated (see <u>http://data.cefas.co.uk/#/View/3236</u> for a fuller description of the data and required modelling of depth).</td>
</tr>
</table>

	are referring to the gridded products) and the others are analyses of various types. 5. Page 7, line 16. it would be better practice to include the data and flag it as likely erroneous. 6. Page 7, line 34. what is a modelled depth? 7. Pate[sic] 9, line 16. link broken 8. Page 10, line 25. unclear what a "conservative approach" means 9. Page 10, line 30. have the data been excluded or flagged? the latter would be better. 10. Page 12, line1. improving rather than falling? 11. Page 13, line 36. were the manual adjustments flagged? 12. Page 18, line 11. HadSST3 is a gridded data product, derived from observations in ICOADS. 13. Page 19, line 13. Other sources such as the World Ocean Database and the Met Office EN4 contain subsurface observations. 14. Figure 15. add information on meaning of red and black to caption 15. Page 30, line 8. Figure 16 does not go back to 1962/3. also 1989 and 2003 don't look very different to adjacent years.	Three point locations in the Loughor Estuary (Great Pill, Bennett's Pill), Oxwich Bay (Nicholstan Pill) and Swansea Bay (Black Pill) in South Wales. Positions were Loughor (51.6262° N, 4.2257° W), Oxwich Bay (51.5701° N, 4.1458° W) and Swansea Bay (51.5987 N, 3.9937° W) (WGS84, Plate carrée). The Pills are only inundated for part of the tidal cycle. Tidal heights for open water locations adjacent to each Pill were estimated with a tidal model (see Additional Information) at the following locations: Loughor (51.651° N, 4.226° W), Oxwich Bay (51.554° N, 4.132° W) and Swansea Bay (51.594° N, 3.976° W) (WGS84, Plate carrée). Inundation depths are strongly dependent on sea conditions, but inundation occurs at tidal height of approximately 6.0 m, 8.8 m and 8.0 m at Loughor, Oxwich Bay and Swansea Bay respectively. 7. http://www.ferrybox.org/ link was tested and found to be active. 8. Given the wide range of documentation of accuracies across the sources and the anticipated uses (i.e. not physical oceanography to high degrees of accuracy – source data for these are available in CTD formats Including e.g. conductivity/salinity – see comments and responses above), data were provided to decimal places that we were certain would not mislead. 9. Covered above. Specifically, data were excluded where values were physically impossible or obviously sensor and/or recording errors. 10. Accepted 11. No. The requirements for the manual adjustments arose from two main areas. Firstly, the early PC systems and software used in the 1980s had issues dealing with date and time changes across midnight, in the middle of the usual night-time deployment period for plankton tows. For reasons that are unclear the software issues were not addressed. The raw data from plankton tows was processed separately with duration calculations clearly identifying where these errors occurred. These were rectified manually during the analysis but, as was frequently the case, the raw data were not "flagged" and/or "corrected". This disconnect between quality controlled data and raw data was not uncommon. In the case of Cefas and its evolving IT systems, raw data were archived and preserved in reasonably consistent function/data acquisition based processes. The, frequently spreadsheet based, quality controlled data used for reports and publications, was also preserved but in project based file assemblages that were not preserved in accessible, documented ways. Secondly, there is a similar source of complication in some of the early physical measurement systems, specifically the use and evolution of different file formats associated with changes in sensors and systems. These file formats include header and data information and were driven by then limitations of data acquisition and storage limitations. Flagging the adjustments	7. None 8. Sentence replaced with *The data provided reflect our best estimates of accuracy when transforming the data from a wide variety of bespoke measuring, recording and use systems (some data were presented with decimal places beyond those implied by statements regarding accuracy of measurement or, in the case of position, than is known to have be feasible at the time of collection)*. 9. None 10. Implemented 11. None – covered elsewhere 12. Clarified 13. The World Ocean Database (https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html) and the Met Office EN4 database (https://www.metoffice.gov.uk/hadobs/en4/) do contain subsurface data and ICES (www.ices.dk/marine-data/data-portals/Pages/ocean.aspx) attempts to provide insights into near-seabed temperature conditions in certain geographical areas, but data are generally sparser than for the surface. 14. Implemented 15. Text changed to: *Figure 16 clearly shows the annual cycle of seawater temperatures around the British Isles, as well as interesting features such as the run of 3 cold winters (1985-87) followed by 3 warm winters (1988-90), plus warm summers (1995, 2006).*

		needed to deal with the format of the raw data files was not deemed beneficial. 12. Accepted 13. Accepted 14. Agreed 15. Thank you. The original text was written for a different version of this Figure. Figure 16 evolved from an illustration of annual cycles and variation to one that also showed the availability of data and cycles in the subsurface data. This entailed change in the time scales used and data subsets plotted but changes in the text did not reflect this.	
Supplem-ental Material	I use R so was pleased to see that the authors had provided R code as a supplement to the main paper. I tried (albeit briefly) to use this to check out the data (which I downloaded successfully as per the instructions). I failed to do this, but make some comments on my progress that may help the authors streamline this aspect of their documentation. a) The filename in the command to read the csv file was not the same as the file I downloaded. b) The conversion of time does not require the *lubridate* package c) I downloaded marmap_0.9.6 but was unable to use it. Following the instructions (from both home and at work several days apart): > map.ukcs <-getNOAA.bathy(-15, 10, 48, 63, res = 1, keep=TRUE) Querying NOAA database ... This may take seconds to minutes, depending on grid size Error in getNOAA.bathy(-15, 10, 48, 63, res = 1, keep = TRUE) : The NOAA server cannot be reached d) Many of the pages in the supplement seemed to give information required to plot the different maps and regions, which I couldn't get to work. It would have been much more helpful to me if simple R scripts to read in the data, save to .Rda, then do some simple diagnostic plots were provided. The *marmap* package does give high resolution coastlines and provides bathymetry (which I couldn't actually see on the printouts, perhaps because of the blue background), but the visualisations in Figures 3 and 4 could have been done without this package, and perhaps provided clearer information about the density of observations (in Figure 3) and without unhelpful contouring (in Figure 4).	As indicated in the File Description section of the code, my (D Morris) R coding skills are newly acquired and limited. The comments below are addressed and the code document amended where appropriate. In addition, the RMarkdown formatting has been made consistent, e.g. Heading Level 1 now all uppercase, Level 2 are lowercase. The use of RMarkdown headings is also made consistent to assist in separating data and plot related code. Where appropriate, the descriptive text has been clarified based on a general review in the light of the comments. These changes are described at a high level. Any changes to the code are described in detail. a) This comment is correct, reflecting an omission in the code document. This is because there are 18 data files, one for each source and one for all data. Suitable suffixes were added to the filenames but this was not reflected in the code document. b) In assembling the data from the 17 sources, many different date formats were encountered and *lubridate* was part of the solution to standardising a number of them (dates in any programming language are not the easiest of things to deal with and R is no exception). The citation of the *lubridate* package in this section is not needed but *lubridate* functions are used elsewhere. c) During the development of the code I too encountered this error message and the server was down for an extended period in early 2017. However, since then it has been up, working and reliable for me at home and at work. A test access of the NOAA site on 13 September 2017 at around 11:00 from the Cefas Lowestoft office was successful. d) The requirements for code provision were for Figures used in the manuscripts. Text was provided giving the R code to facilitate the import of the data from .csv files. Simple statistics that illustrated the distributions of data numbers with time and depth were provided in Figures 5 & 6 and the code for these is included. The Data Paper addresses issues around combining data sets from different sources and types of source and this will be followed up in a separate paper. Saving to .Rda files is routine in the RStudio/R combination I use but recommendations for file sharing focus on common formats such as .csv. I also note the existence of a significant number of Python programmers in Cefas alongside R programmers and my inability to readily provide Python scripts and .Rda equivalents.	Changes to Code document Formatting of RMarkdown headings. File description section edited to make it clearer why the R code is not elegant or efficient. a) Amended. b) Text description of the import code suitably amended. c) None d) None

		In the absence of specific formal requirements, I followed the Google style guide which provided an overall structure and a few basic principles. The code section on BASE MAPS describes the selection and use of *marmap*. The mapping efforts started with *marmap* as a simple, marine package with useable examples, for example, for colour schemes. Figure 3 shows the locations of data, essentially in terms of cruise tracks, stations and point sources, whilst Fig 4 was provided at the specific request of Cefas scientists to illustrate the density of observations; this required contouring. Requests were also made for map styling that resulted in the coding becoming complicated as a result of issues arising from projection and mapping issues around the coasts and the occurrence of horizontal banding arising from unclosed polygons in alternate map sources. Time pressures and skill levels led to a decision to continue using *marmap* which, combined with internal style requirements led to country overlays which in turn led to fixes associated with the fixes. Thus, the code evolved under a natural selection pressure of is it quick, can I get it to work, does it do what is needed no matter how inelegant? "*It's ugly but it works*" seems to be a frequently invoked sentiment.	