The paper by Hemming and coauthors describes an application of the multi-decadal in-situ observational dataset of marine temperatures (AMDOT) for the detection of extremes (AMDOT-EXT).

We thank the reviewer for their insightful comments.

While I appreciate the importance of extended observational records for assessing long-term changes, I am not sure to understand how to interpret MHW/MCS numbers with less-than-daily frequency data. From the images included, it seems that high-impact events can be detected only over the last ~20 years (compare Table 3 and Figure 4), even if at l100 an event in 1944 is described.

It is true that MHW and MCS events can only be identified as per the Hobday (daily) definition when daily-binned data are available as they define MHW/MCSs as discrete prolonged events lasting 5 days or more. Hence, we can mostly derive MHW/MCS characteristics when we have daily mooring and satellite measurements since 2008/2009. In the absence of daily data, we can also consider temperature 'spikes', which are also interesting, especially considering the long-term time-series.

We included text stating this in the submitted manuscript. For example, in Section 3.2 (now edited):

"We identify HSs as temperatures that are higher than the daily-varying 90th percentiles and CSs as temperatures that are lower than the daily-varying 10th percentiles. When continuous daily fixed time series are available (after 2008 / 2009, see Table 1), we further consider MHWs and MCSs when the spikes last for 5 or more consecutive days (as per the widely adopted definition, where a MHW/MCS event needs to last more than 5 days)."

Please also refer to our response to your related comment below.

I find Table 1 and the duration of the MHW/MCS records not really clear. Please explain (with a table or figure) the periods for which high-frequency extreme statistics are available.

We incorrectly used the HS/CS time periods for MHW/MCSs and vice versa (see also response to R2). We have corrected the time periods in the updated Table 1 which will improve clarity.

We have changed the colour of the data points shown in Figure 2 (also copied below) that were used to derive MHW/MCS events, and we have added a corresponding label to the legend. These coloured data points highlight when high-frequency extreme statistics are available. Additionally, we have added '(high-frequency data)' underneath 'Mooring & Satellite' in the legend.
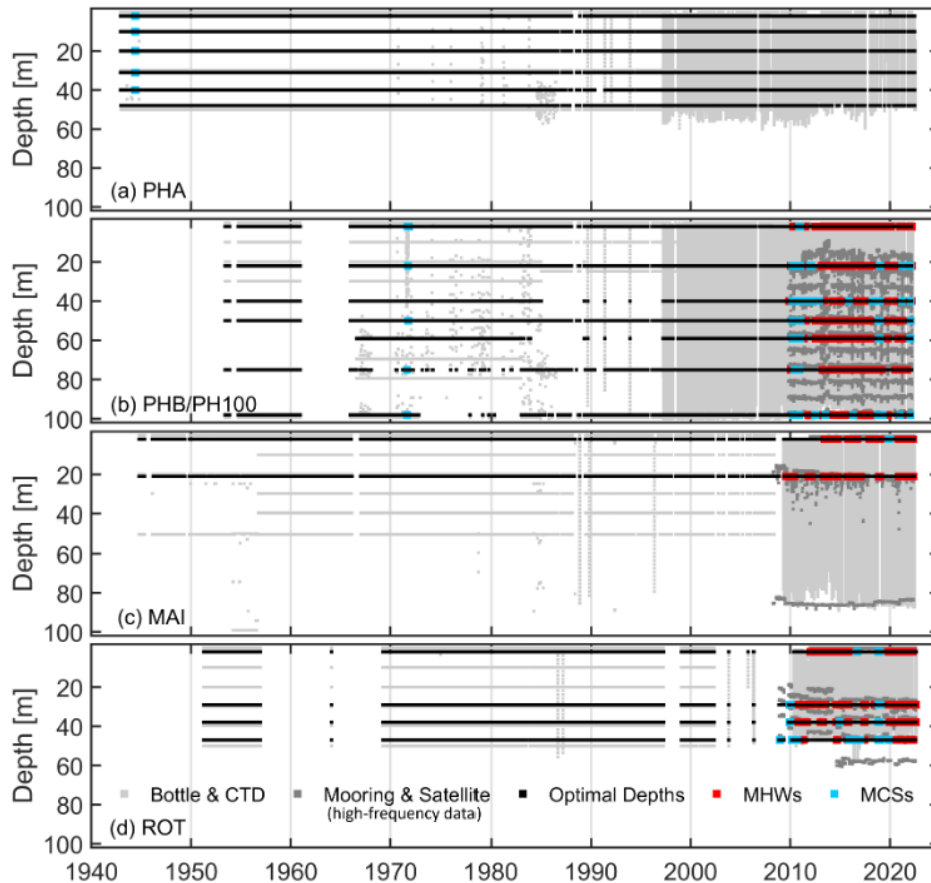
Figure 2. Depiction of available temperature data from bottle and CTD profiles (light grey), mooring and satellite (dark grey), and data binned at optimal depths (black) for a) Port Hacking 50 m (PHA), b) Port Hacking 100 m (PHB/PH100), c) Maria Island (MAI), and d) Rottnest Island (ROT). Data points associated with marine coldspells (MCSs, light blue) and marine heatwaves (MHWs, red) are also shown, and in most cases highlights the time periods when continuous data exists at the sites.

Please note that we have updated the AMDOT-EXT data products. Refer to our response to R2 for more details.

Moreover, some information on the changes in climatology should be provided, since using the whole record likely lead to more events in recent years.

This is indeed an area of ongoing research, and there is much debate in the community on which way is best (e.g. Amaya et al., 2023; Sen Gupta, 2023; Rosselló et al., 2023), and we believe the answer is dependent on the question being asked. In our case we have chosen to use the full record and have been very clear about the time period used, as stated in a dedicated section 3.4 'Effects of Data Availability and Choice of Baseline'. We do this as it would be challenging to compare events using a 30-year climatology below the surface due to a lack of data at our sites.

We have edited Section 3.4 for additional clarity:

"We acknowledge that using a climatology of approximately 30 years is recommended (Hobday et al.,2016; WMO, 2018), and is commonly used for surface *MHW studies (e.g. Frölicher et al. (2018); Oliver et al. (2018); Elzahaby and Schaeffer (2019)) where such data exists. Schlegel et al. (2019) showed that using a climatology period > 30 years may affect MHW detection in a similar way as using a climatology period < 30 years, dependent on environmental multi-decadal variability. Further, Rosello et al. (2023) highlight the impact of long-term trends on MHWs and argue that a moving*

baseline is more adequate for detecting consistently rare extremes, although presently there is much debate on this subject (Amaya et al. (2023); Sen Gupta (2023)). We use daily climatologies calculated using the entire temperature record at the long-term sites because 30 year daily sub-surface data sets do not yet exist. At this time, high-frequency sub-surface data at the 4 sites only extend back approximately 13 to 14 years. We acknowledge that the number of extreme temperature events and their metrics will differ in the future when a 30-year daily climatology becomes available, and that it will vary depending on which 30-year period is used (as suggested by Schlegel et al 2019)."

Amaya, Dillon J., Michael G. Jacox, Melanie R. Fewings, Vincent S. Saba, Malte F. Stuecker, Ryan R. Rykaczewski, Andrew C. Ross, et al. "Marine Heatwaves Need Clear Definitions so Coastal Communities Can Adapt." Nature 616, no. 7955 (April 2023): 29–32. https://doi.org/10.1038/d41586-023-00924-2.

Sen Gupta, Alex. "Marine Heatwaves: Definition Duel Heats Up." Nature 617, no. 7961 (May 16, 2023): 465–465. https://doi.org/10.1038/d41586-023-01619-4.

Rosselló, P., Pascual, A., & Combes, V. (2023). "Assessing marine heat waves in the Mediterranean Sea: a comparison of fixed and moving baseline methods," Frontiers in Marine Science, Volume 10, Page 1168368

Analysis quantifying the consequences of selecting different climatology periods is out of scope for this paper as our focus is to describe and explore the new AMDOT-EXT data products. This additional analysis would warrant a whole new paper, when considering work undertaken by e.g. Rosello *et al.*, (2023) and Schlegel *et al.*, (2019). These studies suggest that time series of different lengths will have different trends, and hence the climatologies will differ.

We have hence not included additional text on this in the manuscript.

Background on other possible data sources in the area should be given. For example, no long-term measurement is available for Northern Australia? Data for the last ~20 years are likely available for more sites.

The long-term sites along the southeastern coast of Australia are the focus of this paper. While there are some long-term data sites in northern Australia, we are not the custodians of it. However, to provide some additional context we have added the following sentence in Section 2:

"The ROT, MAI, and PHB/PH100 sites are the longest sub-surface ocean time-series in Australian coastal waters and are part of Australia's national reference station network; a network of long-term sampling sites (presently 7) around Australia (Lynch et al., 2014)."

The authors rely on Roughan 2022b for too many details. More information on the observational methods need to be provided. For example, no discussion on the long-term consistency of these records is provided.

We have provided more information on the observational methods in Section 3.1:

"Several different SeaBird Electronics CTD sensors (e.g. SBE17+, SBE19+, SBE25) have been used to measure temperature and pressure at the sites."

"At PHB/PH100 Thermistors (Aquatech AQUAloggers 520 T or 520PT) were mounted on a mooring line at 8 m intervals through the water column, while at MAI Wetlabs WQMs and SeaBird SBE37s were mounted on a mooring line at around 20 to 25 m and close to the bottom (approx. 90 m). At ROT various temperature sensors (e.g. Wetlabs WQMs, Seabird SBE37 and SBE39) were placed on a mooring line at approximately 22, 27, 35, 43, and 55 m (varying over time)."

"We use SST at drifting buoy depth (approximately 0.2 m) that has had the sensor specific error statistics (SSES) bias removed from each grid cell. Further, we only use SST associated with `quality level' flags >= 4 (less likely to be contaminated by clouds). A full description of the data QC, and how these aggregated data products were created is provided by Roughan et al. (2022c), and further information specifically on the SST data and methods is provided by Govekar et al. (2022)."

"The long-term consistency of the temperature data has been tested through comparisons of measurements overlapping in time and space. At PHB/PH100 and MAI satellite SST agrees well with surface buoy temperature (Hemming, 2023a). At PHB/PH100 there is also good agreement between satellite SST and the nearest subsurface mooring temperature (Roughan et al., 2022a; Malan et al., 2021), and between water sample and CTD sensor profiles (Hemming, 2023a)."

Familiarity with the Hobday's terminology is also assumed, please explain what technical terms (e.g., cumulative intensity) mean. Please carefully check the units.

We have added the following text to Section 3.3:

"The intensity is the temperature anomaly relative to the chosen statistic (e.g. mean climatology) (Hobday et al., 2016). Over the duration of a MHW/MCS, the cumulative intensity is the sum of the daily temperature intensities, while the variance of intensity is the variation (Hobday et al., 2016). We provide event intensity variables that are relative to both the mean and threshold (e.g. 90th percentile) climatology (see Table 2). The onset rate represents the temperature change rate from the start of a MHW/MCS to the maximum intensity, while the decline rate represents the temperature change rate from the maximum intensity to the end of the MHW/MCS (Hobday et al., 2016)."

We have also modified the units for cumulative intensity to be '°C days' throughout the manuscript.

The authors mention 'optimal depths', but it is unclear to me how these are defined.

We have added the following text to Section 2:

"These optimal depths were determined using binned data to establish the depths where data availability is highest over the full record (Hemming et al., 2020)."

The depth-resolved information seems an asset, please include some illustration of an event detected at multiple depths.

We have added a new figure showing a marine heatwave and a marine cold spell over depth at the PHB/PH100 site, which will be included in the updated manuscript, and which is also copied below.
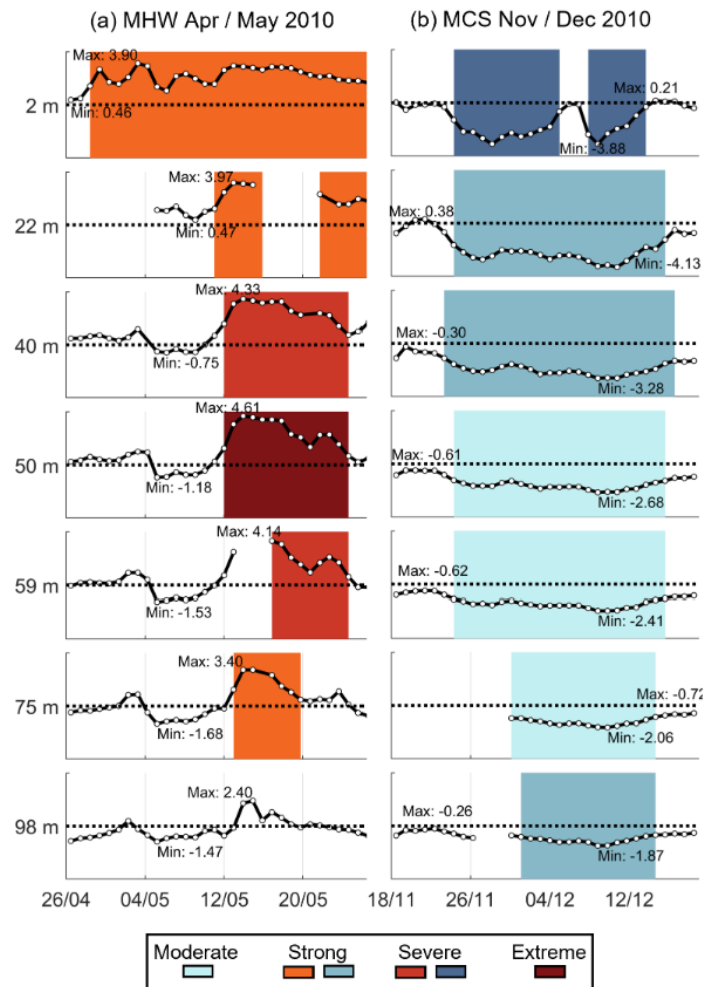
Figure 4. Examples of a MHW and MCS over depth at the Port Hacking (PHB/PH100) site: (a) MHW in April / May 2010, and (b) MCS in November / December 2010. The daily-varying temperature anomalies relative to the mean (black line with white dots) are shown for each depth. Time periods when MHWs or MCSs occur are shaded, with colors indicating the category ('moderate', 'strong', 'severe', and 'extreme') as shown in the legend. The maximum and minimum temperature anomaly relative to the mean for each time period and depth is also shown.

'Code and data availability' and A1 contents look too similar. Revise.

We have modified the first paragraph in 'Code and data availability' as follows:

"The AMDOT-EXT data products are available as NetCDF files (https://doi.org/10.26198/wbc7-8h24) (Hemming, 2023), and variables contained in the NetCDF files are listed in Table 2. Any and all use of the AMDOT-EXT data products provided here must include a citation (see Appendix A1). ~~to this paper, a reference to the data citation as written in the NetCDF file attributes, and the IMOS acknowledgements statement. Further information on these is provided in Sect A1.~~ Any updated data products, and or potential new products (e.g. at other sites or using other ocean variables) will be hosted at the same location. Therefore, it is advised that data users seek the latest data product version."

And we have removed the following from Appendix A1:

"~~Any updated data products, and / or potential new products (e.g. at other sites or using other ocean variables) will be hosted at the same location. Therefore it is advised that data users seek the latest data product version.~~"

The provided codes can be useful to know how to open the data but seems quite basic.

Thank you for looking at the code. In addition to opening the data sets they also demonstrate how to extract required information, save output in different formats (e.g. CSV), and how to slice and analyse the event characteristics (See Appendix A3 for a summary of this code). It is out of scope for this data paper to provide more complex code tutorials.

I do not understand how you deal with event statistics: categories can be defined at daily frequency, but statistics are for individual events (e.g., cumulative intensity). How do you save them as a function of TIME and DEPTH? The same values are repeated for all the days during a MHW/MCS?

We save event-constant statistics (e.g. cumulative intensity) for a MHW/MCS as the reviewer suggested. The same values are repeated for all the days of the event.

We have added the following text clarifying this in Section 3.3:

"*Some metrics for MHWs and MCSs, such as duration, maximum intensity, and cumulative intensity, for example, are constant over the duration of an event. In the NetCDF files these constant metrics are repeated for each day of an event. For example, a MHW lasting 6 days with a maximum intensity of 23.2 °C will have 6 repeated duration values (e.g. 6,6,6,6,6,6) and 6 repeated maximum intensity values (e.g. 23.2, 23.2, 23.2, 23.2, 23.2, 23.2) representing each date of the MHW.*"

l27 please give details on the depth they analysed

l30 the last part of the sentence is a repetition

In response to the two comments above, we have edited this sentence as follows:

"*… were amongst the first to analyse MHWs below the surface at multiple depths between the surface and the bottom at a 65 and 100 m coastal location site off Sydney, Australia. and showed that the sub-surface structure of MHWs is highly variable and different to the surface Using in situ data from these sites off Sydney …*"

l37 if from the early 40s, why not ~80 years?

It is true that we have data from the early 1940s at PHA and MAI, but at ROT and PHB we have data from the 1950s. Hence, to include all sites we stated 'more than 65 years'.

We have now modified this sentence to include:

"*… over periods of 70 to 81 year periods.*"

Fig 1 please explain what isolines are (bathymetry? Derived from?)

We have added the following to the caption for Fig 1:

"*… Isobaths (data from General Bathymetric Chart of the Oceans (GEBCO), Weatherall et al. (2023)) are also plotted in panels (b), (c) and (d).*"

Table 1 for 2m depth, I would show not the relative (i.e., always 100%) but the absolute available fraction (maybe on monthly or weekly basis). This would help interpreting other relative fractions.

We have changed the 2m depth relative % to the absolute available fraction as suggested. In other words, the total record available at 2 m depth as a percentage of the time period between the first and last daily-binned temperature.

We have also changed the relative % record available text as follows: e.g. '99%' -> '- 1%', '129.73%' -> '+ 29.73%', which we think is clearer. These percentages represent sub-surface data availability relative to 2 m depth. The column heading has also been changed from 'Relative % Record Available' to 'Record availability' and further text clarifying this column has been added to the table caption.

Moreover, I do not like the fact that the table is an image - please check the editorial guidelines.

The table is an image during the review period only as it is easier for us to modify it. We will include a LaTeX table in the final version of the manuscript.

l55 please provide information on the instrumentation, I guess it varied over time leading to varying uncertainty. How these levels are related to local conditions such as the mixed layer depth?

It is beyond the scope of this manuscript to explore how the temperature varies in response to local conditions such as MLD. At the reviewers request we have provided more information on the instrumentation in Section 3.1:

" Several different *SeaBird Electronics CTD sensors (e.g. SBE17+, SBE19+, SBE25) have been used to measure temperature and pressure at the sites.*"

"*At PHB/PH100 Thermistors (Aquatech AQUAloggers 520 T or 520PT) were mounted on a mooring line at 8 m intervals through the water column, while at MAI Wetlabs WQMs and SeaBird SBE37s were mounted on a mooring line at around 20 to 25 m and close to the bottom (approx. 90 m). At ROT various temperature sensors (e.g. Wetlabs WQMs, Seabird SBE37 and SBE39) were placed on a mooring line at approximately 22, 27, 35, 43, and 55 m depths (varying over time).*" More information on these configurations can be obtained from Lynch et al. (2014); Chen and Feng (2021); Roughan et al. (2022c).*"

"*The long-term consistency of the temperature data has been tested through comparisons of measurements overlapping in time and space. At PHB/PH100 and MAI satellite SST agrees well with surface buoy temperature (Hemming, 2023a). At PHB/PH100 there is also good agreement between satellite SST and the nearest subsurface mooring temperature (Roughan et al., 2022a; Malan et al., 2021), and between water sample and CTD sensor profiles (Hemming, 2023a).*"

l66 define CTD

We now define 'CTD' on this line.

l68 how L3S differs from L3? Define 'QC'

'L3' is used to identify gridded SST data, while the 'S' in 'L3S' represents gridded Super-Collated SST composites as per Govekar et al., (2022). We have added the following to this line:

" *… (SST) data from the IMOS Multi-sensor night-time gridded Super-Collated (L3S) SST data composites (Govekar et al., 2022)*"

l75 'free' => 'open source'

We have changed this.

l104 how do you identify the MLD?

We do not identify the MLD in this work as that is beyond the scope of this study. Instead on this line we are referring to the MLD in an example to demonstrate that event variability can be real, as well as due to data availability. We have therefore not added additional text to the submitted original manuscript related to this comment.

l105 missing citation

We have now added this citation to the manuscript.

l123 and elsewhere (e.g. Table 2) I would use italics or a typewriter font for variable names, to distinguish them from MHW, MCS,... - check journal guidelines

We agree that italics might be better for variable names, as the journal guidelines suggest using italics for words, phrases, and abbreviations that cannot be found in the English dictionary. We have therefore changed all variable names to italics in the manuscript.

l130 Excel is not really needed for csv files, they are simply text files (as they are not spreadsheets). Moreover, Excel is not open nor free - suggest providing open-source alternatives.

We have replaced 'spreadsheets' with 'files' throughout the manuscript. The reviewer is correct that CSV files can be opened by many different software. We have removed reference to any single software in the updated manuscript.

l132, 157, 221 what is 'onedrivelink'? Why not using the website you cite?

We agreed with Copernicus ESSD upon submission that for the review process we would supply the tutorials via a OneDrive link for the reviewers to access only. We will upload the tutorials to Zenodo with a permanent DOI after peer-review.

l149 I would suggest to compare the events detected in the last ~20-30 years when using your full-record climatology and a more standard 30-years baseline.

Please see our response to your comment above regarding the choice of baseline.

l161 explain 'cumulative intensity'

We have added the following text to Section 3.3:

"The intensity is the temperature anomaly relative to the chosen statistic (e.g. mean climatology) (Hobday et al., 2016). Over the duration of a MHW/MCS, the cumulative intensity is the sum of the daily temperature intensities, while the variance of intensity is the variation (Hobday et al., 2016). We provide event intensity variables that are relative to both the mean and threshold (e.g. 90th percentile) climatology (see Table 2). The onset rate represents the temperature change rate from the start of a MHW/MCS to the maximum intensity, while the decline rate represents the temperature change rate from the maximum intensity to the end of the MHW/MCS (Hobday et al., 2016)."

l165 remove space between the circle sign and C

We have removed the space between the degrees symbol and 'C' throughout the manuscript.

l170 without context, Table 3 is not very useful. Please add some references.

We have added the following paragraph to the introduction:

"*There are numerous examples of extremely warm or cold temperatures being recorded in Australian coastal waters. In 2011 off Western Australia, intense low pressure linked to a strong La Niña intensified the Leeuwin current leading to MHW conditions at the surface. This MHW lasted for months and maximum intensities peaked at 5∘C warmer than the 2000–2009 average climatology (Feng et al., 2013). After this event, coastal waters off Western Australia experienced further MHWs in the austral summer of following years before sea surface temperatures switched to a multi-year cold phase and when multiple MCSs occurred between 2016 and 2019 (Feng et al., 2021). In late 2015 / early 2016 off eastern Tasmania, the East Australian Current (EAC) Extension waters were warmer than average leading to MHW conditions over a large area of the Tasman Sea. The regionally-averaged MHW conditions at the surface persisted for 251 days between September 2015 and May 2016 and regionally-averaged sea surface temperature anomalies were between 1.5 and 3∘C warmer than average between November 2015 and February 2016 (Oliver et al., 2017). A similar event also occurred during austral summer 2016/2017 (Kajtar et al., 2022). MHW conditions persisted for 3 months, with mean Tasman sea surface temperatures 1 to 3∘C above the 1983 to 2012 seasonal climatology (Kajtar et al., 2022). Finally, in austral summer 2021/2022 the position of the EAC jet and its eddies off southeastern Australia had a crucial role in driving MHW/MCS conditions in coastal waters close to Sydney. A MHW occurred in December 2021 and February 2022, while a MCS occurred in January 2022 (Li et al., 2023).*"

We have also added some further discussion in Section 5:

"*The most intense MHWs were at approximately 22 m depth at sites PHB/PH100 and MAI in 2015 and 2015/16, respectively (Fig 4a,b). These coastal events coincide with the far-reaching MHW conditions in the Tasman Sea related to the East Australian Current (Oliver et al., 2017).*"

"*The longest MHW detected lasted 138 days at MAI (same as in Fig . 4) between 28/12/2015 and 13/05/2016 coinciding with the unprecedented Tasman Sea MHW (Oliver et al., 2017) while the longest MHWs at PHB/PH100 and ROT lasted 54 and 60 days, respectively. The longest ROT MHW occurred at the time of the `Ningaloo' MHW (Feng et al., 2013).*"

l172 and elsewhere, I think units should be 'C * day', not 'C / day'. These are the units for onset/decline rates instead. This seems correct in the netCDF files.

This was a mistake in the submitted manuscript, and we have now corrected the units for cumulative intensity to be '°C days'. The file is correct – thanks for checking.

l221 are you also using zenodo?

As explained above, we will upload the tutorials to Zenodo with a DOI at the time of publication.

l225 NetCDF is not necessarily a python format, it is actually aimed at portability.

We have removed '(python)' on L225

Please rephrase 'appendix section XX' => 'appendix XX'

We have changed the text as suggested.