**KongCTD : CTD data upload portal**

Draft version 2023-03-24

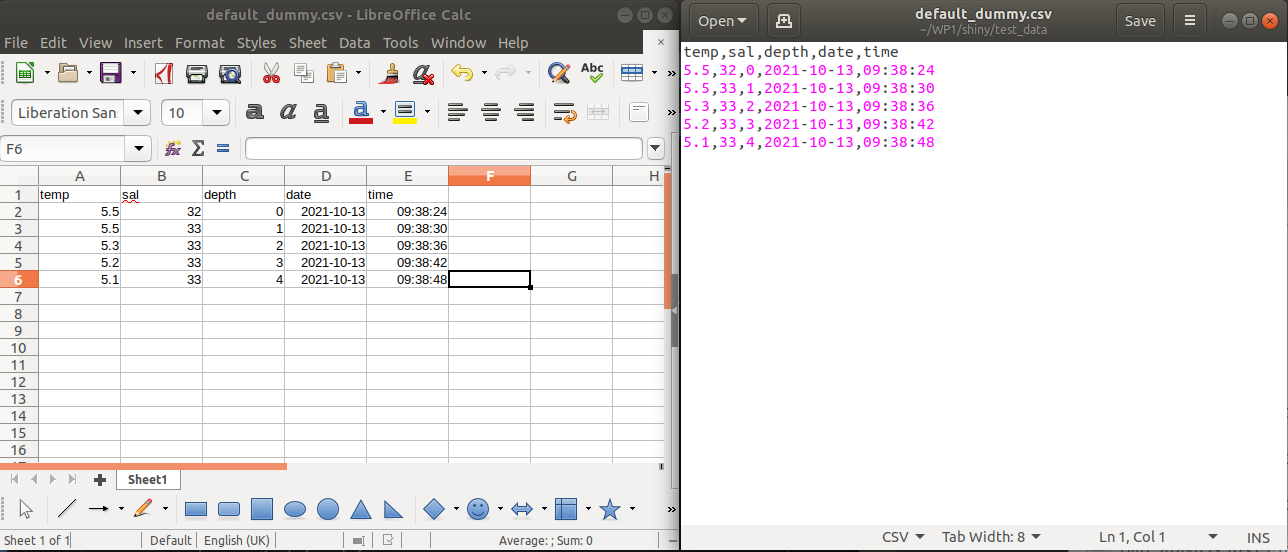
**The aim of developing this portal was to create a data repository to ensure the archiving of all Kongsfjorden CTD data for the good of the entire scientific community and to track long term trends in an important Arctic environment undergoing rapid change.**

**Have data you want to upload? Follow these steps:**

Choose who will upload

* + The Kings Bay Marine Lab manager can upload your CTD data for you if the data were collected on the Kings Bay CTD, as the marine lab manager will have your data. Contacts for questions: Marine Lab manager: [**engineer@kingsbay.no**](mailto:engineer@kingsbay.no), Flagship leads: Clara Hoppe (clara.hoppe@awi.de) and Allison Bailey ([allison.bailey@npolar.no](mailto:allison.bailey@npolar.no)).
  + You can upload your own data with the following steps:
    1. Create user account:
       1. Contact the portal’s administrator Robert Schlegel ([**robert.schlegel@imev-mer.fr**](mailto:robert.schlegel@imev-mer.fr)) to have a new user account created.
       2. Provide them with your name and e-mail, as well as your desired username (ideally “Lastname\_firstname”, no special characters) and password.
    2. Prepare your data:
       1. Data must be in .txt or .csv file format.
       2. You can upload the raw files exported from the CTD, but do not use versions where you have manually changed the file or added additional text.
       3. You can upload a single profile, or multiple profiles at once. If multiple profiles are uploaded at once, they must come from the same CTD and have the same column structure. To streamline the process of adding metadata when uploading multiple files, consider these tips:
          1. If you have a lot of files from the common CTD sites in Kongsfjorden, it’s easiest to upload your data in pools of files, one station (profiles from the same lat/long) at a time.
          2. If you want to upload data from different stations in one go, make sure you have the lat/long data for each profile available (and note that you can copy-paste that data into the app just like Excel!).

Accessing the portal

* + Go to: http://193.50.85.71:3838/kongCTD/
  + Log in using your username and password assigned by the admin.
  + The web portal is structured by sequential pages, listed to the left. After logging in, you will see the first page of the app ( ‘1) Load file(s)’ ).
  + You can navigate between the different pages by clicking on the next tab once you are finished with the current one. You can move back and forth between the different tabs without losing or having to save your data in between.

**1) Load file(s)**

* + In the blue box (“File format”) you can identify the file(s) you want to upload (can be multiple files at once) and in which format they are.
    1. Click on browse and select the file(s) you want to upload.
    2. If the portal recognizes the file type as coming from one of the typical CTDs used in Kongsfjorden (e.g. for SAIV, Sea-Bird or RBR CTDs), it will suggest to you a pre-determined file format and parse the data accordingly. Otherwise the “CTD type” will stay as ‘Default’.
    3. Inspect in the “Uploaded data preview” box to the right. If the data are not parsed correctly, you can select a CTD type in the drop down menu and/or change the file format settings manually.
  + In the green box you can check how the data have been read. Double check e.g. if decimals are displayed correctly.
  + *NB:* If you want to upload several files for a CTD that is not of one of the commonly used types in Kongsfjorden (the Kings Bay SAIV or the AWIPEV RBR), contact us in advance to see whether it is possible to implement an automated upload scheme for your instrument type. Alternatively, you may use the manual controls to match your file format within the app. Or you may manually prepare your files to meet the default standard for a CSV file, as seen in the Figure 1.

**Figure 1:** The ‘Default’ CSV data structure understood by the Kongsfjorden CTD data upload app. The file is opened in excel on the left, and a text editor on the right. *NB:* The default column separator is ‘,’ and the default decimal place is ‘.’. The default file encoding is ‘UTF-8’, but note that on many computers in the EU files are encoded as ‘Latin’.

**2) Metadata**

* + Metadata (date, location, CTD type, etc.) is what makes the data useful, and it is thus critically required to provide this information.
  + In the light blue box (‘Fill values here that apply to all files’) you can add information to all currently uploaded files.
  + If your profiles have been collected at one of the pre-described, traditionally used stations in Kongsfjorden (e.g. KB3), you can select this station from the drop down menu. If your station is not in the dropdown menu or if your station has no fixed ID, the field can be edited manually in the blue box see below). Contact Robert Schlegel if you want your commonly-used stations added to the dropdown menu.
  + If you select a pre-described station, lat and long values will appear below. They can also be manually adjusted for all currently uploaded datasets if needed.
  + Add the name of the data owner, both the person and the institute (please use abbreviation, e.g. NPI instead of *Norwegian Polar Institute*)
  + The sensor owner (e.g. Kingsbay for the SAIV data format), brand and number are pre-described based on the file structure. If it is missing or incorrect, they can be changed/added here for all currently uploaded files
  + The dark blue box (‘Fill file-specific values here’) can be used to add or edit metadata for individual files (e.g. if you upload files from different stations together).
    1. It works like a normal spreadsheet (e.g. excel), where you can type, drag, and copy past from a different file.
    2. If you want to upload files from different stations together, it makes sense to prepare a table with site, lat, long info that can be directly copy-pasted into this box.
  + The map (orange box) shows you the lat/long position of your data. If the box around the map is green, all data are within the geographic range, if it is orange some data are located outside, and red means no data are located within the map. In this case you may want to double check your lat/long entries. For the future we plan to extend this to off the shelf locations, currently these profiles will appear as out of range, but can still be uploaded.
  + If your data have already been published, please add the DOI of the primary data publication (in this case, the data will be accessible via the database, with your associated DOI. Publishing of the entire dataset on Pangaea will also include the original DOI). You can also add a DOI to your dataset later via the edit page.

**3) Quality Control / Upload**

* + Before being uploaded, the data will undergo a basic QC and plausibility check procedure.
  + QC flag 1 stands for ‘implausible value’ and corresponds to data falling out of the following data ranges, that have been are agreed upon at during the flagship meeting:
    1. Temperature: -1.8°C to 15°C
    2. Salinity: 0 to 38
    3. Fluorescence: 0 to 50 µg L-1
  + QC flag 2 means that the measurement has been taken above the sea surface (negative depth values)
  + QC flag 3 stands for upcast data
  + QC flag 0 will be given to all remaining data (‘no problem’)
  + Your data will be published as part of the joint yearly dataset in the Pangaea repository via the automated pipeline provided by the app. If you have provided a DOI, the Pangaea dataset will refer to that primary data publication. The data will be published together with the other data from that year with all data owners as coauthors. The order of the author list will be dictated by the number of provided files.
  + If your data still needs be published in another database before being added to the Pangaea dataset, or if you wish to delay publication for other reasons, you can tick the ‘yes’ field under ‘embargo data for 2 years’. In this case, your data will not be included into the Pangaea publication for 2 years. The data will be added to the next Pangaea publication with you as the primary data owner. You can also provide a DOI via the edit page during the coming 2 years.
  + In the orange box ‘QC flags and missing meta-data in data to be uploaded’ you get an overview on the results from the automated checks. Under files you can check whether the embargo has been implemented.
  + Under QC flags (light orange), you will see the number of individual measurements that have been given the different QC flags. Ideally, you want a large number under flag 0 and no datapoints listed for flags 1 and 2.
  + Under ‘missing meta-data’ (dark orange) data points without full metadata are shown, so you want this to show all zeros.
  + In the green box (‘Meta-data for full database’) the meta data of all uploaded files is shown.

**4) Edit**

* + On this page, you are able to edit the metadata of all files that you have uploaded, e.g. if you realize that you used the wrong location for one of your files.
  + This is also the location where you can add a DOI to your datasets once they have been published in a repository.
  + To edit the metadata, just type into the black fields. And click on ‘save changes’ before leaving the window.

**Download**

* + On this page, all metadata that have been uploaded via this portal to the protected database online can be found and downloaded. This does not download the actual data from Pangaea. If you want to access data that is not yet published, the metadata will indicate the data owners to contact.
  + In the blue box (‘Download from database’), you can select the range of profiles you want to download in terms of 6 characteristics.
    1. All data and sensor owners are displayed, but can be omitted by clicking on the name and deleting it.
    2. Deleted data or sensor owners can be added back via the appearing drop down menu.
  + The selected data points are shown in the green data box (use drop-down menu to show different variables on x and y axis) and their locations are illustrated in the orange map box.
  + To illustrate data distribution and ranges, all selected data can be plotted in the red time series box by selecting the variables for the *x* and *y* axis of the plot.
  + Metadata can be downloaded as .csv or .rds files via the download data button.

**FAQ and Trouble Shooting**

* + Will be developed based on user feedback & occurring problems
  + If you have any technical problems with the app, contact Robert Schlegel ([robert.schlegel@imev-mer.fr](mailto:robert.schlegel@imev-mer.fr))
  + If you have questions regarding data usage, publication and collaboration, please contact the chairs of the Kongsfjorden Flagship, Clara Hoppe (clara.hoppe@awi.de) and Allison Bailey ([allison.bailey@npolar.no](mailto:allison.bailey@npolar.no))

A picture containing water, outdoor, sky, boat

Description automatically generated