
Marine user guide

Release v1.0

IPG

Jun 26, 2020

CONTENTS:

1	Indices and tables	1
2	Introduction	3
3	Tool set-up	5
3.1	Code set up	5
3.2	Data directory setup	5
3.3	Paths setup	6
4	Initialize a new user guide	7
5	Data summaries	9
5.1	Monthly grids	9
5.2	Monthly time series of selected quality indicators	9
6	Figures	11
6.1	Number of reports time series plot	11
6.2	Duplicate status time series plot	11
6.3	Report quality time series plot	11
6.4	Seasonal and monthly Hovmöller plots	12
6.5	ECV nreports and converge time series	12
6.6	Maps with number of observations and number of months observed	12
6.7	Maps with ECVs mean observations value	12
7	Appendix. Configuration files	13
7.1	level2 file format	13
7.2	Marine user guide data configuration file	14
7.3	monthly_grids configuration	15
7.4	qi_counts configuration	16
7.5	nreports_ts configuration	16
7.6	nreports_dup_ts configuration	16
7.7	nreports_qc_ts configuration	16
7.8	nreports_hovmoller configuration	17
7.9	ecv_coverage configuration	17
7.10	map_mean configuration	17
7.11	map_nobs configuration	19

INDICES AND TABLES

- `genindex`
- `modindex`
- `search`

INTRODUCTION

This project contains the necessary code to produce the data summaries that are included in the Marine User Guide. These help to document the status of the marine in situ data in the CDS after every new data release.

The marine data available in the CDS is the result of a series of data releases that are stored in the marine data file system in different directories. This project uses the data in the marine file system, rather than accessing to the CDS data.

Every new data release can potentially be created with a different version of the marine processing software. The current version of this project is compatible with the glamod-marine-processing code up to version v1.2.

TOOL SET-UP

3.1 Code set up

Clone the remote repository:

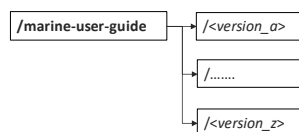
```
git clone git@git.noc.ac.uk:iregon/marine-user-guide.git
```

Build the python environment using the requirements.txt file in marine-user-guide/env. This step system dependent. The following code block described the steps to follow in CEDA JASMIN, using the Jaspy toolkit.

```
cd marine-user-guide/env
module load jaspy/3.7
virtualenv --system-site-packages mug_env
source mug_env/bin/activate
pip install -r requirements.txt
```

3.2 Data directory setup

The data that the tools in this project use and the products created are all stored in the marine-user-guide data directory, which is organized in separate directories to host the different versions of the Marine User Guide.



This general directory needs to be created before starting using the tool.

```
cd <parent_data_directory>
mkdir marine-user-guide
```

3.3 Paths setup

Edit file `marine-user-guide/setpaths.sh` and modify as needed the following fields:

- `code_directory`: parent path of the repository installation.
- `data_directory`: parent path of the Marine User Guide data directory.

INITIALIZE A NEW USER GUIDE

Every new version of the Marine User Guide (MUG) needs to be initialized in the tools' data directory (*Marine User Guide data directory and its relation to the data releases directories.*). This means:

- Creating the appropriate data configuration files to reflect the combination of individual data releases that generate the version.
- Creating a subdirectory for the version in the data directory.
- Create a view of the merged data releases: rather than copying all the files to a common location, this is done by linking the level2 data from the different releases to the marine-user-guide data directory.

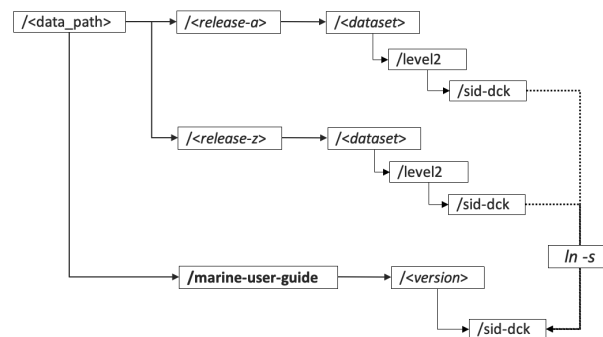


Fig. 1: Marine User Guide data directory and its relation to the data releases directories.

These steps initialize a new user guide:

1. Create the data configuration file (*mug_file*, ref: *mug_config*) by merging the level2 (*level2 file format*) configuration files of the different data releases included in the new version.

```
python marine-user-guide/init_version/init_config.py
```

2. Use the sid-dck keys of the *mug_config* file to create a simple ascii file with the full list source-deck IDs of the release merge (*mug_list*).

3. Create a view of the merged data releases. The launcher script also initialises the subdirectory for the new version.

```
./marine-user-guide/init_version/merge_release_data_launcher.sh version mug_config_  
↪ mug_list
```

Check that the copies really reflect the merge of the releases. Edit the following script to add the corresponding paths and run. If any does not match, it will prompt an error.

```
./marine-user-guide/init_version/merge_release_data_check.sh
```


DATA SUMMARIES

The data summaries are monthly aggregations of all the source-deck ID partitions in the data.

5.1 Monthly grids

These are monthly aggregations in a lat-lon grid which are stored in nc files:

- header table: number of reports per grid cell per month
- observations tables: number of observations and observed_value mean per grid cell per month.

All the aggregations are configured in a common configuration file, `monthly_grids.json` (*monthly_grids configuration*). The current configuration excludes reports not passing all the quality checks, but this can be configured in the configuration file.

A launcher bash script configures the LSF job for each table and logs to `/log/monthly_grids/table.log` in the data directory.

```
./marine-user-guide/data_summaries/monthly_grids_launcher.sh version monthly_grids.  
↪ json
```

5.2 Monthly time series of selected quality indicators

Monthly summaries of categorical counts of quality indicators aggregated over all the source-deck IDs. These are additionally, split in counts by main platform types (ships and buoys) and include the total number of reports. They are stored in ascii pipe separated files.

The configuration file, `qi_counts_ts.json` (*qi_counts configuration*), includes very limited parameterization, with the platform type segregation pending to be parameterized.

A launcher bash script configures the LSF job for each quality indicator (currently only `report_quality` and `duplicate_status`) and logs to `/logqi_counts_ts/quality_indicator.log` in the data directory.

```
./marine-user-guide/data_summaries/qi_counts_ts_launcher.sh version qi_counts_ts.json
```


FIGURES

The data summaries generated are used to generate a series of maps and time series plots. The following sections give the necessary directives to create them, with references to the configuration files used.

6.1 Number of reports time series plot

- Data summary used: report_quality quality indicators time series.
- Configuration file: nreports_ts.json (*nreports_ts configuration*)
- Command:

```
source /marine-user-guide/setpaths.sh
source /marine-user-guide/setenv.sh
python /marine-user-guide/figures/nreports_ts.py nreports_ts.json
```

6.2 Duplicate status time series plot

- Data summary used: duplicate_status quality indicators time series.
- Configuration file: nreports_dup_ts.json (*nreports_dup_ts configuration*)
- Command:

```
source /marine-user-guide/setpaths.sh
source /marine-user-guide/setenv.sh
python /marine-user-guide/figures/nreports_dup_ts.py nreports_dup_ts.json
```

6.3 Report quality time series plot

- Data summary used: report_quality quality indicators time series.
- Configuration file: nreports_qc_ts.json (*nreports_qc_ts configuration*)
- Command:

```
source /marine-user-guide/setpaths.sh
source /marine-user-guide/setenv.sh
python /marine-user-guide/figures/nreports_qc_ts.py nreports_qc_ts.json
```

6.4 Seasonal and monthly Hovmöller plots

- Data summary used: monthly grids (counts, header and observation tables)
- Configuration file: `nreports_hovmoller.json` (*nreports_hovmoller configuration*)
- Command:

```
source /marine-user-guide/setpaths.sh
source /marine-user-guide/setenv.sh
python /marine-user-guide/figures/nreports_hovmoller.py nreports_hovmoller.json
```

6.5 ECV nreports and converge time series

- Data summary used: monthly grids (counts, header and observation tables)
- Configuration file: `ecv_coverage_ts_grid.json` (*ecv_coverage configuration*)
- Command:

```
source /marine-user-guide/setpaths.sh
source /marine-user-guide/setenv.sh
python /marine-user-guide/figures/ecv_coverage_ts_grid.py ecv_coverage_ts_grid.
↪ json
```

6.6 Maps with number of observations and number of months observed

- Data summary used: monthly grids (counts, header and observation tables)
- Configuration file: `map_nobs_from_monthly_nc.json` (*map_nobs configuration*)
- Command:

```
source /marine-user-guide/setpaths.sh
source /marine-user-guide/setenv.sh
python /marine-user-guide/figures/map_nobs_from_monthly_nc.py map_nobs_from_
↪ monthly_nc.json
```

6.7 Maps with ECVs mean observations value

- Data summary used: monthly grids (mean, observation tables only)
- Configuration file: `map_mean_from_monthly_nc.json` (*map_mean configuration*)
- Command:

```
source /marine-user-guide/setpaths.sh
source /marine-user-guide/setenv.sh
python /marine-user-guide/figures/map_mean_from_monthly_nc.py map_mean_from_
↪ monthly_nc.json
```


APPENDIX. CONFIGURATION FILES

7.1 level2 file format

Listing 1: Release_2.0 level2 file extract

```
{
  "001-110": {
    "year_init": 1945,
    "year_end": 1949,
    "exclude": false,
    "params_exclude": []
  },
  "001-116": {
    "year_init": 1945,
    "year_end": 1949,
    "exclude": false,
    "params_exclude": []
  },
  "171-711": {
    "year_init": 1889,
    "year_end": 1899,
    "exclude": false,
    "params_exclude": []
  },
  "params_exclude": ["observations-wbt"],
  "year_init": 1851,
  "year_end": 1949
}
```

Listing 2: r092019 level2 file extract

```
{
  "year_init": 1950,
  "year_end": 2010,
  "params_exclude": [
    "observations-wbt"
  ],
  "001-110": {
    "exclude": false,
    "params_exclude": [],
    "year_end": 1951,
    "year_init": 1950
  },
}
```

(continues on next page)

(continued from previous page)

```
"001-116": {
  "exclude": false,
  "params_exclude": [],
  "year_end": 1963,
  "year_init": 1950
},
"001-999": {
  "exclude": false,
  "params_exclude": [],
  "year_end": 1969,
  "year_init": 1967
}
}
```

7.2 Marine user guide data configuration file

Listing 3: Marine User Guide v4 configuration file extract

```
{
  "release_names": [
    "r092019",
    "release_2.0"
  ],
  "dataset_names": {
    "r092019": "ICOADS_R3.0.0T",
    "release_2.0": "ICOADS_R3.0.0T"
  },
  "year_init": 1851,
  "year_end": 2010,
  "params_exclude": [
    "observations-wbt"
  ],
  "001-110": {
    "year_init": {
      "r092019": 1950,
      "release_2.0": 1945
    },
    "year_end": {
      "r092019": 1951,
      "release_2.0": 1949
    },
    "exclude": false,
    "params_exclude": []
  },
  "063-714": {
    "year_init": {
      "r092019": 1978
    },
    "year_end": {
      "r092019": 2010
    },
    "exclude": false,
    "params_exclude": [
      "observations-ws",

```

(continues on next page)

(continued from previous page)

```

        "observations-dpt",
        "observations-at",
        "observations-wd"
    ]
},
"152-721": {
    "year_init": {
        "release_2.0": 1851
    },
    "year_end": {
        "release_2.0": 1868
    },
    "exclude": false,
    "params_exclude": []
}
}

```

7.3 monthly_grids configuration

```

{ "dir_data": "/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4",
  "dir_out": "/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "start": 1851,
  "stop": 2010,
  "id_out": "optimal",
  "region": "Global",
  "resolution": "lo_res",
  "header":
  {
    "filter_by_values": {"header.report_quality": [0]}
  },
  "observations-at":
  {
    "filter_by_values": {"header.report_quality": [0], "observations-at.quality_flag": [0]}
  },
  "observations-dpt":
  {
    "filter_by_values": {"header.report_quality": [0], "observations-dpt.quality_flag": [0]}
  },
  "observations-slp":
  {
    "filter_by_values": {"header.report_quality": [0], "observations-slp.quality_flag": [0]}
  },
  "observations-sst":
  {
    "filter_by_values": {"header.report_quality": [0], "observations-sst.quality_flag": [0]}
  },
  "observations-wd":
  {
    "filter_by_values": {"header.report_quality": [0], "observations-wd.quality_flag": [0]}
  },
  "observations-ws":
  {
    "filter_by_values": {"header.report_quality": [0], "observations-ws.quality_flag": [0]}
  }
}

```

(continues on next page)

(continued from previous page)

```
}  
}
```

7.4 qi_counts configuration

```
{ "dir_data":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4",  
  "dir_out":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/  
↪quicklooks",  
  "start":1851,  
  "stop":2010  
}
```

7.5 nreports_ts configuration

```
{ "file_data":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/  
↪quicklooks/report_quality-ts.psv",  
  "file_out":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/  
↪quicklooks/nreports-ts.pdf"  
}
```

7.6 nreports_dup_ts configuration

```
{ "file_data":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/  
↪quicklooks/duplicate_status-ts.psv",  
  "file_out":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/  
↪quicklooks/duplicate_status-ts.pdf"  
}
```

7.7 nreports_qc_ts configuration

```
{ "file_data":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/  
↪quicklooks/report_quality-ts.psv",  
  "file_out":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/  
↪quicklooks/report_quality-ts.pdf"  
}
```

7.8 nreports_hovmoller configuration

```
{
  "dir_data":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "dir_out":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "file_in_id" : "-no_reports_grid_ts-optimal",
  "file_out_id_sea":"nreports_hovmoller_seasonal.png",
  "file_out_id_mon":"nreports_hovmoller_monthly.png",
  "start":1851,
  "stop":2010,
  "tables":["header","observations-at","observations-sst","observations-slp",
           "observations-dpt","observations-wd","observations-ws"],
  "colorbar":"magma"
}
```

7.9 ecv_coverage configuration

```
{
  "dir_data":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "dir_out":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "file_in_id" : "-no_reports_grid_ts-optimal",
  "file_out":"ecv_coverage_ts_grid.pdf",
  "start":1851,
  "stop":2010
}
```

7.10 map_mean configuration

```
{
  "dir_data":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "dir_out":"/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "projection" : "Mollweide",
  "tables":
  {
    "observations-at":
    {
      "nc_file":"observations-at-mean_grid_ts-optimal.nc",
      "out_file":"observations-at-mean-map-optimal.png",
      "scale":1,
      "offset":-273.15,
      "colorpalette":"jet",
      "colorbar_title":"AT ($^\\circ$C)",
      "colorbar_orien":"h",
      "cmax_value":35,
      "cmin_value":-10
    }
  }
}
```

(continues on next page)

(continued from previous page)

```

    },
    "observations-sst":
    {
        "nc_file": "observations-sst-mean_grid_ts-optimal.nc",
        "out_file": "observations-sst-mean-map-optimal.png",
        "scale": 1,
        "offset": -273.15,
        "colorpalette": "jet",
        "colorbar_title": "SST ( $^{\circ}\text{C}$ )",
        "colorbar_orien": "h",
        "cmax_value": 35,
        "cmin_value": -5
    },
    "observations-dpt":
    {
        "nc_file": "observations-dpt-mean_grid_ts-optimal.nc",
        "out_file": "observations-dpt-mean-map-optimal.png",
        "scale": 1,
        "offset": -273.15,
        "colorpalette": "jet",
        "colorbar_title": "DPT ( $^{\circ}\text{C}$ )",
        "colorbar_orien": "h",
        "cmax_value": 35,
        "cmin_value": -10
    },
    "observations-slp":
    {
        "nc_file": "observations-slp-mean_grid_ts-optimal.nc",
        "out_file": "observations-slp-mean-map-optimal.png",
        "scale": 0.01,
        "offset": 0,
        "colorpalette": "jet",
        "colorbar_title": "SLP (hPa)",
        "colorbar_orien": "h",
        "cmax_value": 1030,
        "cmin_value": 980
    },
    "observations-ws":
    {
        "nc_file": "observations-ws-mean_grid_ts-optimal.nc",
        "out_file": "observations-ws-mean-map-optimal.png",
        "scale": 1,
        "offset": 0,
        "colorpalette": "jet",
        "colorbar_title": "WS ( $\text{ms}^{-1}$ )",
        "colorbar_orien": "h",
        "cmax_value": 15,
        "cmin_value": 0
    },
    "observations-wd":
    {
        "nc_file": "observations-wd-mean_grid_ts-optimal.nc",
        "out_file": "observations-wd-mean-map-optimal.png",
        "scale": 1,
        "offset": 0,
        "colorpalette": "twilight_shifted",
        "colorbar_title": "WD (degrees)",

```

(continues on next page)

(continued from previous page)

```

        "colorbar_orien": "h",
        "cmax_value": 360,
        "cmin_value": 0
    }
}
}

```

7.11 map_nobs configuration

```

{
  "dir_data": "/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "dir_out": "/group_workspaces/jasmin2/glamod_marine/data/marine-user-guide/v4/
↪quicklooks",
  "projection" : "Mollweide",
  "colorpalette_nobs": "magma",
  "colorpalette_nmonths": "magma",
  "colorbar_title_nobs": "# reports",
  "colorbar_title_nmonths": "# months",
  "normalization_nobs": "log",
  "normalization_nmonths": "log",
  "colorbar_orien": "h",
  "tables":
  {
    "header":
    {
      "nc_file": "header-no_reports_grid_ts-optimal.nc",
      "out_file_nobs": "header-nobs-map-optimal.png",
      "out_file_nmonths": "header-nmonths-map-optimal.png"
    },
    "observations-at":
    {
      "nc_file": "observations-at-no_reports_grid_ts-optimal.nc",
      "out_file_nobs": "observations-at-nobs-map-optimal.png",
      "out_file_nmonths": "observations-at-nmonths-map-optimal.png"
    },
    "observations-sst":
    {
      "nc_file": "observations-sst-no_reports_grid_ts-optimal.nc",
      "out_file_nobs": "observations-sst-nobs-map-optimal.png",
      "out_file_nmonths": "observations-sst-nmonths-map-optimal.png"
    },
    "observations-dpt":
    {
      "nc_file": "observations-dpt-no_reports_grid_ts-optimal.nc",
      "out_file_nobs": "observations-dpt-nobs-map-optimal.png",
      "out_file_nmonths": "observations-dpt-nmonths-map-optimal.png"
    },
    "observations-slp":
    {
      "nc_file": "observations-slp-no_reports_grid_ts-optimal.nc",
      "out_file_nobs": "observations-slp-nobs-map-optimal.png",
      "out_file_nmonths": "observations-slp-nmonths-map-optimal.png"
    }
  },
}

```

(continues on next page)

(continued from previous page)

```
"observations-ws":
{
  "nc_file":"observations-ws-no_reports_grid_ts-optimal.nc",
  "out_file_nobs":"observations-ws-nobs-map-optimal.png",
  "out_file_nmonths":"observations-ws-nmonths-map-optimal.png"
},
"observations-wd":
{
  "nc_file":"observations-wd-no_reports_grid_ts-optimal.nc",
  "out_file_nobs":"observations-wd-nobs-map-optimal.png",
  "out_file_nmonths":"observations-wd-nmonths-map-optimal.png"
}
}
```