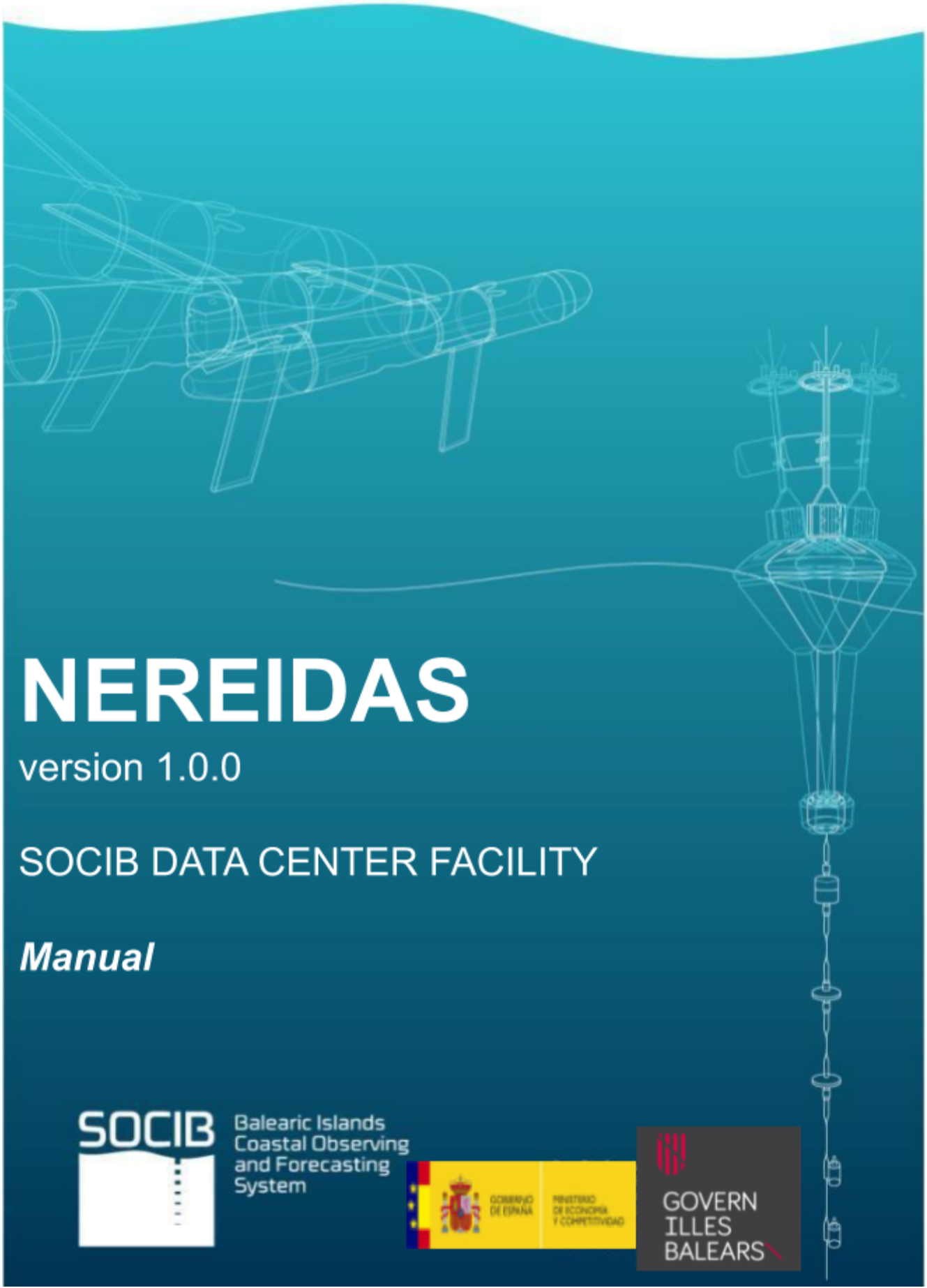

NEREIDAS

version 1.0.0

SOCIB DATA CENTER FACILITY

Manual

SOCIB Balearic Islands Coastal Observing and Forecasting System

GOBIERNO DE ESPAÑA · MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

GOVERN ILLES BALEARS

DOCUMENT VERIFICATION LIST

Date:	Checked by (name)	SOCIB division	Ref.
2019-08-19	C. Muñoz	SOS	

DOCUMENT DISTRIBUTION LIST

Date:	Distribution to:
2019-08-19	public release

CHANGE RECORD

ver	Date	Description	Author	Checked by
1.0.0	2019-07-27	First version document	M. Charcos	C. Muñoz

Overview **4**

SCBVessel Package Description **4**

Configuration 5

Services 7

Bash Scripts 8

Common Scripts 8

SADO Support Scripts 8

NEREIDAS Scripts 9

MISC Scripts 9

Requirements and Installation **10**

Requirements 10

Installation of NEREIDAS 10

The R/V SOCIB NEREIDAS system is a Near-Real Time System developed to show different variables (meteorological, navigational information and oceanographic parameters) acquired in the vessel during a cruise. Such data is displayed in SOCIB Seaboard application. NERIDAS replaces the previous system used from the UTM-CSIC - namely SADO - to record data from the onboard sensors. Such acquisition system presented several limitations, as it is a closed system designed by the UTM-CSIC. It was able to cover their needs, but is not able to provide response and flexibility to the SOCIB needs.

1. Overview

The Coastal Ocean Observing and Forecasting System (SOCIB for Sistema de Observación Costero y de Predicción de las Islas Baleares) is a multi-platform distributed and integrated system that provides streams of oceanographic data and modelling services to support operational oceanography in a European and international framework (see www.socib.eu). Among the observing facilities, R/V SOCIB is critical to support glider regular operations and other specific scientific missions. These missions may have specific requirements of the data formats. Thus, NEREIDAS responds to the needs of the regular processing of the data center and to the specific needs of the PI on board of the vessel. It also allows recording of the raw data as output from the sensor for future re-processing or processing in different formats.

NEREIDAS is part of the SCBVESSEL_UTILS package. This package contains two utilities. The SADO scripts were used to fill the gaps in the framework of the SOCIB usage of the SADO. These scripts allow manipulation of the output files from the UTM-CSIC system. The NEREIDAS scripts allow collecting the data from the sensors that are sent on the network. Both sets of scripts are described in the package description section.

2. SCBVessel Package Description

The SCBVessel package contains three main folders: configuration, services and scripts. The installation of the package is described in the requirements and installation section. In this

section, we explain the content of each of these three folders. The NEREIDAS component diagram illustrates the interaction between the various components of the package.

a. Configuration

The configuration folder contains the files that are used by NEREIDAS to define the network frames, the output data, the sensors and their calibrations. The content is at scbvessel_utils/config as follow:

- File **env.def**: Contains the environment variables that are set at each call as explained in the bash script section. Most of the definitions are intrinsic to the implementation and must not be changed. A few variables may be changed with caution if the requirements of the software change such as data path or variable name conventions. A few examples of these variables are:
    - SCBVESSEL_UTILS_SCBDCDATA_PATH="/datos/BaseDatosContinua"
    - SCBVESSEL_UTILS_SADO_DATA="$SCBVESSEL_UTILS_SCBDCDATA_PATH/SCB"
    - SCBVESSEL_UTILS_RAWBASENAME="$SCBVESSEL_UTILS_SADO_DATA/{MM-YYYY}/{$SCBVESSEL_UTILS_LABEL_PORTNAME}.raw/{DDMMYYYY}.{$SCBVESSEL_UTILS_LABEL_PORTNAME}.raw"
    - SCBVESSEL_UTILS_BINBASENAME="$SCBVESSEL_UTILS_SADO_DATA/{MM-YYYY}/{$SCBVESSEL_UTILS_LABEL_PORTNAME}.bin/{DDMMYYYY}.{$SCBVESSEL_UTILS_LABEL_PORTNAME}.bin"
    - SCBVESSEL_UTILS_PROCBASENAME="$SCBVESSEL_UTILS_SADO_DATA/{MM-YYYY}/{$SCBVESSEL_UTILS_LABEL_PORTNAME}.proc/{DDMMYYYY}.{$SCBVESSEL_UTILS_LABEL_PORTNAME}.proc"
    - SCBVESSEL_UTILS_TMPSAMPLE="$SCBVESSEL_UTILS_SCBDCDATA_PATH/tmp/{MM-YYYY}/{YYYYMMDD}.nereidas_sample"
- File **nmea_standard.txt**: Contains the description of the NMEA variables, the binary frames and the outputs of the interpretation of the sensor frames. This file is referred in the env.def configuration as SCBVESSEL_UTILS_NMEA_STANDARD_FILE. If you would like to use another NMEA description file you can either edit the env.def or call the scripts with the appropriate options. You can see an example of NMEA description below. New descriptions can be added to the file if other sensors are used. Each line of the nmea file contains the name of the variable follows by each of the fields that show in the sensor frame. If the name of the variable starts with the $ sign, the frame is interpreted as an ascii frame. If it starts with the ! sign, the frame is a binary frame. The bash scripts use these descriptions to interpret the frames that are gathered from the sensors or to output the data in a standard format.
- File **port_config.txt**: Contains the description of the sensors that are sending data into the network. Each line explains the information of the data that should be collected and the routine that collects this data. The scbvessel_nereidas_raw service will use this information to start background processes that will write the data in the specified network ports. The file name is defined in env.conf as SCBVESSEL_UTILS_PORT_CONFIG_FILE. The format of the lines of the file are as shown below. X indicates that the values are not used.

    - Field 1: Name given by the user
    - Field 2: Type of frame (NMEA or BINARY)

- Field 3: Sniff IP where the sensor data is listen
- Field 4: Sniff port where the sensor data is listen
- Field 5: IP where data can be forwarded
- Field 6: port where data can be forwarded
- Field 7: Name of variables to be collected (comma separated)
- Field 8: Script name of the function collecting the data

For example, the line

```
gps:NMEA:X:5602:X:X:GPGGA,GPRMB,GPZDA,GPVTG:scbvessel_getudp_nmeasensor
```

Indicates that the variable named gps (user defined) is of type NMEA, should be listen in port 5602, and the scbvessel_getudp_nmeasensor script will read only the NMEA variables GPGGA, GPRMB, GPZDA and GPVTG. Another example is:

```
meteo:BINARY:172.16.137.44:30000:X:X:NEREIDAS_METEO:scbvessel_gettcpip_geonica2000
```

The variable named meteo is a binary frame that is read by scbvessel_gettcpip_geonica2000 in the ip: port 172.16.137.44:30000.

Optionally, a forward IP, port can be set by adding values to field 5 and 6.

- Folder **proc**: Contains the definition of the outputs to be created by NEREIDAS. These files are used by the crontab scbvessel_nereidas_proc to create proc files in a specific format. They are json files with the following fields:
  - Name: name of the product
  - Outproc: Name of the file containing the product with the data of the whole day
  - Partproc: Name of the file containing the product with the data partial data since last execution
  - Header: Array with the names of the header keywords to consider in the same order as in the product files
  - Required_fields: List of columns that must exist in order to consider the row as valid. A valid row is written to the product file and a non valid row is ignored.
  - Empty_fields: List of columns that must have at least one no NaN values in order to consider the row as valid. A valid row is written to the product file and a non valid row is ignored.
  - Sort_field: Reference column that is used to sort the data
  - Fields: Name and definition of the columns and the way their values are calculated:
    - Name: Name of the column
    - Cmd: Command to calculate the value
    - Params: Parameters of the command CMD. Order is important since they will be used as arguments to the funcion.

                - Constant: Value of a constant to be used
                - A column definition. The Name of the NMEA column is made as SENSOR:NMEA:COL ignoring the variables that are not defined. COL is mandatory.
                    - Sensor: Name of the sensor
                    - Nmea: Name of NMEA variable
                    - Col: Name of the column of the NMEA definition
            - Default: Default value to use if it cannot be calculated with the CMD
            - Constant: Constant value to be used. It is an alternative to the CMD calculation when a specific value should be used.

- Folder **calibration**: JSON file that contains the calibration values of specific columns. The dataframe is checked against the definition in this file and the specified columns calibrated accordingly. It contains two fields (pre_calibration and calibration) defining the pre-calibration and the calibration processes, each are an array of structure with the following fields:
    - Input: Field definition of the input column in the dataframe
        - Sensor: Name of the sensor
        - Nmea: Name of NMEA variable
        - Col: Name of the column of the NMEA definition
    - Output: Field definition of the column to be created in the dataframe
    - cmd : command name to be used to calibrate the input
    - Params: Structure to use as input of the calibration command. The calibration command should take the data frame followed by the param structure as a dictionary. Thus, it should understand the fields of the param structure.

b. Services

The folder scbvessel_utils/services contains the definitions of the files that should be run to support SADO or NEREIDAS systems.

- scbvessel_watchnasfile.service, scbvessel_watchnasmeteo.service, scbvessel_watchnastermosal.service and scbvessel_watchsensor.service support SADO. They sniff the creation of new files in the file tree. Scbvessel_watchnasfile.service sniff all the sensor at once while the other are sniffers for specific sensors.
- **Scbvessel_nereidas_raw.service** is a main component of the NEREIDAS package. It sniffs the sensor data in the network that were defined in the port_config.txt file (see Configuration and Bash Scripts sections).

In order to use a service, the service file should be copied at /etc/systemd/system. The service will start automatically at reboot but can be manage as a regular service by calling service scbvessel_nereidas_raw start, stop or restart.

c. Bash Scripts

The scripts are contained under the scbvessel_utils/bin folder. The folder contains three scripts that are used by the whole package, a misc folder containing the supporting scripts, a sado folder containing the scripts of the SADO system utilities and a nereidas folder containing the scripts of the NEREIDAS system.

i. Common Scripts

- **Scbvessel**: Sets the environment of the scripts. All routines must be called within this environemnt. For example, to call the routine scbvessel_nereidas_raw you should call **scbvessel nereidas_raw**. The scbvessel script sets the environement and calls scbvessel_nereidas_raw within this environment. Other routines outside the package may also be called with this environment the same way. If a script is not recognized by scbvessel, it will just set the environment and call it. If it is recognized, the scbvessel_ prefix is not needed and the scbvessel will interpret that the routine starts with that prefix.
- **Scbvessel_checkenv**: Checks the environment. It is used at the beginning of each nereidas script.
- **Scbvessel_logevent**: Log events. It outputs in the log files with a specific format containing the routine name, date and message. All log messages within the package must use this function to log message for consistency.

ii. SADO Support Scripts

The sado supporting utilities have two groups as indicated in the SADO support tools diagram. The first one watch the data that is created by SADO from the sensors and updates the NAS filesystem accordingly. The second one watches the files in NAS and adds coordinates to the files that are updated. The description of the scripts is as follow:

- GROUP1:
    - **Scbvessel_watchsensor**: Watch creation or updates of files in the data filesystem. These are the files that are created by the SADO system. It is used when the transfer is performed as a service. Unfortunately, the watcher fails sometimes when the access to disk is slow.
    - **Scbvessel_transfer2nas**: Transfer files of the input to the NAS structure. If a directory is input, files are search recursively inside the tree. It is used when calling the transfer as a crontab.
    - **Scbvessel_transferfile**: Transfer a file to the NAS structure.
- GROUP2:

    - **Scbvessel_watchnasfile**: Watch creation or updates of files in the NAS filesystem. These are the ones that are transferred from the /datos/BaseDatosContinua/SCB folder. It is used when the transformation is performed as a service. Unfortunately, the watcher fails sometimes when the access to disk is slow.
    - **Scbvessel_transformnas**: Transform the files of the input and create new files with the coordinates. If a directory is input, files are search recursively inside the tree. It is used when calling the transformation as a crontab.
    - **Scbvessel_addcoord**: Add coordinate columns from sensor and coordinate files.
    - **Scbvessel_getfcoord**: Finds the coordinate file correspondign to the input file.
- **Scbvessel_newpath**: Creates the equivalent path in the NAS data storage of the Vessel of the input file in the SADO directory structure. It allows to maintain the directory tree in a specific format. It is a supporting script for both groups.

iii. NEREIDAS Scripts

The NEREIDAS Scripts can be divided in three groups as illustrated in the NEREIDAS component diagram. Configuration files are described in configuration section and their role is represented in the diagram. The first group of functions collect data in the network from the sensors. The second group reads and process the data.

- GROUP1: This group runs a service
    - **Scbvessel_nereidas_raw**: Start background processes to collect data from sensors
    - **Scbvessel_gettcpip_geonica2000**: Collects the raw data from the meteo
    - **Scbvessel_getudp_nmeasensor**: Collects the raw NMEA data from the gps and the gyro
    - **Scbvessel_getudp_sbe21**: Collects the raw data from the termosal.
    - **Scbvessel_getudp_sonda**: Collects the raw data from the sonda

- GROUP2: This group runs as crontabs
    - **Scbvessel_nereidas_proc**: Used to create proc files based on the proc definition list. It calls a python script that uses pandas to manage and process the data

3. MISC Scripts

These scripts are in the bin/misc folder. The intend is to support all the scripts in the package although in practice it only supports the NEREIDAS scripts.

- SUPPORTING FUNCTIONS

- **Fill_coefficients**: Given an equation and a calibration file, it returns the equation where the coefficients are replaced by their values. It is mainly used for the calibration of termosal data.
  - **Get_calibration_label**: Gets the value in a calibration file of a specific coefficient. It is used by fill_coefficients to get the values of the coefficients in an equation.
- **Filldate**: Given a path with a {DATE} type string, it fills them with their value
- **Hex2float**: Converts a hex to float value.

4. Requirements and Installation

a. Requirements

- Linux and bash
- Libraries:
  - Inotifywait for SADO supporting routines
  - XXD, netcat, socat and tcpdump
- Hardware setup as shown in the RV Nereidas component diagram. (Configure accordingly)
- Python and virtualenv (preferably virtualenvwrapper)

b. Installation of NEREIDAS

- Clone scbvessel_utils repository.
- Create virtual environment. Requirement file is at bin/nereidas/requirement.txt
- Set configuration files:
  - Configure port_config.txt to match hardware.
  - Configure proc files to match required proc outputs.
  - Configure json files.
- Set raw watcher services. Copy scbvessel_nereidas_raw.service to /etc/systemd/system/
  - Start service: service scbvessel_nereidas_raw start
  - Service must be restart when configurations are changed
- Set crontabs:

```
lpath=/home/socib/scbvessel_utils/logs
spath=/home/socib/scbvessel_utils/bin
*/1 * * * * $spath/scbvessel nereidas_sample -N 60 -i 1 -r 10 -v >> $lpath/crontab_nereidas_sample.log
*/5 * * * * $spath/scbvessel nereidas_proc -N 300 -v >> $lpath/crontab_nereidas_proc.log
```

```
*/1 * * * * $spath/scbvessel nereidas_clean_samples -vr 1200 >> $lpath/crontab_nereidas_clean_samples.log
```