ISAS-Tool Version 6: User's manual

E. Brion, F. Gaillard
Rapport LPO 12-01

Laboratoire de Physique de Océans, UMR 6523
<table>
<thead>
<tr>
<th>Auteur</th>
<th>Mise à jour</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>F. Gaillard</td>
<td>Création du document – V4 beta</td>
<td>03/02/2007</td>
</tr>
<tr>
<td>F. Gaillard</td>
<td>V4.01 – Version française</td>
<td>11/02/2008</td>
</tr>
<tr>
<td>F. Gaillard</td>
<td>V4.1b - English version</td>
<td>19/03/2008</td>
</tr>
<tr>
<td>F. Gaillard</td>
<td>Minor corrections</td>
<td>25/09/2008</td>
</tr>
<tr>
<td>F. Gaillard</td>
<td>V5.1</td>
<td>18/06/2009</td>
</tr>
<tr>
<td>E. Brion</td>
<td>V6 beta</td>
<td>10/06/2011</td>
</tr>
<tr>
<td>F. Gaillard</td>
<td>V6</td>
<td>05/01/2012</td>
</tr>
</tbody>
</table>
# Table of Contents

1. INTRODUCTION.............................................................................................................................................. 7  
2. General presentation.................................................................................................................................................... 8  
   2.1. The software directories...................................................................................................................................... 8  
   2.2. The data directories........................................................................................................................................... 10  
   2.3. The analysis directory........................................................................................................................................ 11  
3. Data Pre-processing.................................................................................................................................................. 13  
   3.1. Description.......................................................................................................................................................... 13  
   3.2. Running STD....................................................................................................................................................... 14  
   3.3. Outputs.............................................................................................................................................................. 15  
4. Analysis..................................................................................................................................................................... 18  
   4.1. Preprocessing (PREOA)....................................................................................................................................... 18  
   4.2. Analysis............................................................................................................................................................... 20  
   4.3. Post-Processing (POSTOA)............................................................................................................................... 23  
5. Matlab_isas............................................................................................................................................................... 25  
   5.1. The pre-processing of data files (STD program)................................................................................................. 25  
   5.2. Pre- and Post analysis (PREOA/POSTOA)...................................................................................................... 27  
   5.3. Tools................................................................................................................................................................. 29  
6. Optimal interpolation: (f90_isas)................................................................................................................................. 33  
   6.1. Program list........................................................................................................................................................ 33  
7. References.................................................................................................................................................................. 33  
8. ANNEXES................................................................................................................................................................... 34  
   8.1. Configuration file.................................................................................................................................................. 34  
   8.2. log files of the standardisation step.................................................................................................................. 39  
   8.3. output files of the pre-processing step............................................................................................................ 41  
   8.4. log and error file for the Analysis fortran step.............................................................................................. 43  
   8.5. log file of the post-processing steps.............................................................................................................. 44
1. INTRODUCTION

ISAS (In Situ Analysis System) is an analysis tool for producing gridded scalar fields. At the moment it is set to process temperature and salinity. Originally designed for the synthesis of ARGO dataset, it has been tested for the first time on the POMME area in the North-East Atlantic in 2000, it was later extended to the Atlantic and the Global ocean as the Argo array was setting up. It is developed and maintained at LPO (Laboratoire de Physique des Océans) within the Argo Observing Service (SO-ARGO) where it is used for research purposes on ocean variability. ISAS is made available to the Coriolis datacenter for exploitation in operational mode. The analysis is performed on the datasets prepared by Coriolis according to Argo recommendations for data quality control and NetCDF format. A set of background and statistical information required to complement the observations are provided with the software, as part of the configuration. For each analysis date, the results are provided in two NetCDF files, one holding the data and analysis residuals, the other holding the gridded fields and estimation error, expressed as percentage of a priori variance.

This document describes how to implement the ISAS software. The main steps of the process are detailed and examples of configuration files are provided. The statistical method used to produce the estimate and the specific choices performed to implement the method are described in the Method and Configuration document corresponding to the appropriate ISAS version.
2. General presentation

Producing a gridded field with a large dataset downloaded from a datacenter as Coriolis requires several preliminary operations before being able to perform the analysis. Two main tasks have been identified:

1. Definition of the configuration files that include: bathymetry, area limits and masks, as well as climatology (mean state and associated variance). They are provided in the `confstd` directory.

2. Preprocessing of the data files to perform elementary checks and interpolates the data on standard levels. This is performed by the `STD` function.

The analysis (optimal interpolation) itself is divided in 3 elementary operations:

- **PREOA**: preparation of files and assembly of the datasets needed by the analysis.
- **OA**: Optimal interpolation (or analysis)
- **POSTOA**: assembly of the final files

The various files and programs are organized as indicated in Figure 1. Note that STD, PREOA and POSTOA are written in Matlab while OA is in Fortran 90. The directories and configuration files

### 2.1. The software directories

These directories contain all programs and files required to perform the analysis. It is organized as follows.

![Figure 1: Program directories scheme](image)

Figure 1: Program directories scheme
2.1.1. Configuration directory (confstd)

The configuration proposed with ISAS-V6 is described in detail in (Gaillard et al, 2011). The horizontal grid is ½ degree Mercator limited to 80S-90N. The vertical levels are given in the Table 1.

<table>
<thead>
<tr>
<th>Standard levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0:3]</td>
</tr>
<tr>
<td>[5:5:100]</td>
</tr>
<tr>
<td>[110:10:800]</td>
</tr>
<tr>
<td>[820:20:2000]</td>
</tr>
</tbody>
</table>

Table 1: vertical standard levels

The directory « confstd » contains all files defining the standard configuration and statistics used by the analysis (Gaillard, 2011).

It is organised as follow in the directory:

- climref:
  - Bathymetry
  - Climatology (annual and monthly), for the parameter PSAL and TEMP
  - A priori variances (annual), for the parameter PSAL and TEMP
  - Covariance scales (annual, deduced from rossby radius calculated from the climatology)

- areadef:
  - Definition for analysis areas and masks.

2.1.2. Documentation directory (doc)

This directory contains

- the programs documentation (pdf files)
- models of configuration files for matlab scripts and perl script that allow to loop over different analysis dates and parameters.
- models of batch and config files for the fortran part.all files defining the standard configuration and statistics used by the analysis (Gaillard, 2011):

2.1.3. Matlab scripts (matlab_isas)

Contains the matlab scripts for version nn used for the data pre-processing and the analysis pre- and post-processing.

2.1.4. Fortran programs (f90_isas)

This directory contains the source codes. A makefile and executable for SGI-Altrix is provided. Although isas_f90 is part of ISAS chain, it has to be copied in the fortran computer tree structure (see 4.2.1).
2.2. The data directories

2.2.1. The datasets

We briefly describe here the characteristics of the data types taken into account at the moment. These dataset have different accuracy, resolution and sampling that depend mostly on the sensor and on the storage and transmission system used.

Temperature and salinity measurements are obtained from autonomous instruments, drifting or anchored or from instruments deployed with a ship. The data are transmitted in real time by satellite, or in delayed mode. The main characteristics of the most common instruments are given below.

- **Profiling floats**: The autonomous floats are part of the ARGO program, they collect vertical profiles of temperature and salinity as a function of pressure between their maximum pressure (usually 2000 dbars) and the surface. At the end of the profile that takes nearly 5 hours, the profiler transmits the data to a satellite and dives toward its parking depth (1000 dbars), waiting for the next cycle (10 days later). Nominal accuracy of the data is assumed to be 0.01°C and 0.01 PSU. At present time a vertical profile is described by approximately 100 pts.

- **XBT**: An eXpendable BathyThermograph is launched from a steaming ship. It measures temperature (and salinity in the case of XCTD). The measurement depth is deduced from the XBT fall rate. The accuracy is 0.1°C and most XBT reach 800 m.

- **CTD**: This high quality measurement is obtained from a research vessel in the context of a scientific cruise. Pressure and temperature sensors are carefully calibrated and water samples are taken to adjust the salinity measurement. Standard procedure were defined for the WOCE experiment, they lead to accuracies of 0.001°C and 0.001 PSU.

- **Time series**: Time series of pressure, temperature and salinity are recorded at high time resolution (hours) from sensors installed on fixed points (mooring) or drifting buoys. The sensors are similar to those used on the profiling floats. The thermosalinographs, installed aboard scientific or commercial ships, are another type of time series data. The sensors are different from profiling floats ones. Depth measurement is usually constant.

2.2.2. The directories

Two directories are defined, as shown in Fig. 2:

- **dir_raw/** contains all raw data downloaded from the Coriolis datacenter. There must be one subdirectory per year: the user have to create dir_raw and the subdirectory by year. The naming convention of the raw files must follow Coriolis recommendations: datasetname_yyyymmdd_PR_XX.nc.

  where yyyymmdd is the date of the measurements, and XX the type of data (PR for Profilers, CTD, ...)

- **dir_std/** this directory must be created before running the analysis. It will hold the data interpolated on the analysis levels. The subdirectory by year are created automatically by ISAS.

Provided by User

Provided by User
2.3. The analysis directory

It contains 3 subdirectories, as shown in the figure 3.
2.3.1. Confisas directory

This directory must be created by the user. The model files found in isas/doc/config_matlab should be copied here and adapted to the user needs and configuration.

- `isas_matlab.env`: defines the matlab path
- `config_isas_ana.txt`: holds the data paths and the parameters that define the configuration. This file is read by `isas_mat`.

Before starting the analysis, the matlab paths need to be defined. The environment is described in a file `isas_matlab.env` that we recommend be placed in the directory `confisas/`. See `isas_v6/doc/config_matlab` for an example.

To launch, type:

```
cd confisas
source isas_matlab.env
```

Hereafter, an example of the `isas_matlab.env` file:

```
#!/usr/bin/sh
setenv MATHOME /net/triagoz/export/home1/matlab/matlab_last
setenv TOOLBOXPATH /net/triagoz/export/home1/matlab/outils_matlab/m_map1.4
setenv MATLAB ${MATHOME}
set path=($MATLAB $path)

# Configuration ISAS #
setenv ISAS_ANA_HOME /home1/toto/lulu/prog_ISAS/svn/trunk/matlab_isas
setenv ISAS_ANA_PATH ${ISAS_ANA_HOME}/functions:${ISAS_ANA_HOME}/tools
echo " Toolbox ISAS-ANA svn OK "

setenv MATLABPATH ${ISAS_ANA_PATH}:${TOOLBOXPATH}
```

Furthermore, the various paths of the analysis, the file names and the different parameters must be defined. This is done through the configuration file: `config_isas_ana.txt`, that we recommend be placed in the directory: `confisas/`.

An example of configuration file is given in `isas_v6/doc/config_matlab`, and is reported in the annexe (8.1).
2.3.2. Results Directory (DIR_RESU)

This directory contains the analysis results in two subdirectories:

- **field**: contains the 3D field and error on the regular grid
- **data**: contains the data used by the analysis and the residuals.

2.3.3. Directory of calculation (DIR_RUN)

This directory contains the analysis run information in subdirectories:

- **preoa**: contains the pre-processed fields ready to be read in the analysis step.
- **logisas**: contains the log files
- **plotisas**: contains the plots created during the runs
- **alert**: contains the alert lists of suspicious profiles detected during the standardisation step STD.

3. Data Pre-processing

3.1. Description

The first step in the analysis is an interpolation of the raw data on the standard levels of the analysis grid. It is partly independent of the analysis, in the sense that the dataset produced can be used for different analysis. A new QC is introduced, it represents the quality of the interpolation (the closest to a measured value the lowest the QC flag value).

To avoid spoiling the analysis with erroneous data, a control is performed before the interpolation. Finally, oversampled points such as repeated fixed points CTD, drifting buoys, mooring, can be averaged (reduced) into super-profiles. The processing is detailed below.

3.1.1. Detection of erroneous data

Two different tests are successively applied:

**Distance to climatology:**

A data point will be accepted if the value X verifies \( |X_{obs} - X_{clim}| < \alpha_1 \cdot \text{STD} + \alpha_2 \cdot \frac{\delta X}{\delta z} \) where :

- The scalar \( \alpha_1 \) (crit_std in the configuration file) has been determined empirically, it defines the distance allowed to the climatology.
- The scalar \( \alpha_2 \) introduces an additional tolerance relative to the climatology. In the vicinity of very strong stratification, perfectly good data may differ strongly from the climatology. This is taken into account by introducing an additional tolerance proportional to the vertical gradient of the parameter.

**Spike detection:**

A data value is considered as a spike if the following conditions are filled:

- Change of sign of the first derivative for at least one point before or after the point.
• Second derivative criteria normalized by the median in the vicinity of the point:

\[
\left| \frac{\delta^2 P}{\delta^2 z} (z) \right| \left( \text{median} \left| \frac{\delta^2 P}{\delta^2 z} \right| \right)^{-1} \geq \text{crit}_\text{spike}
\]

3.1.2. Interpolation

The high resolution data are bin averaged on the standard levels, then the remaining levels are interpolated. A new Quality Code (QC) is set to the data, related to the interpolation quality. The more the initial data is close to the interpolated one, the better the QC is. To eliminate interpolated points which seem too far from the measurement levels, one can select the \text{str}_\text{qcmax}_\text{ana} level (see PREOA part in the configuration file).

3.1.3. Reduction (superobs)

Data from the same platform which are close in time and space are averaged. The control parameters, defined in the configuration file, are:

- \text{RED\_DXMAX}: Minimum distance in kilometers
- \text{RED\_DTMAX}: Minimum time difference in days
- \text{RED\_QCMAX}: maximum QC-flag (after standardization)

3.2. Running STD

After setting the parameters of the STD block in the configuration file, STD\_main can be launched in the matlab execution window.

STD files may remain on a daily sampling or data within a month can be grouped.

The climatological tests can be applied at two levels. The recommended procedure is:

1. perform a first pass with rather strict parameters:
   \[
   \text{crit\_std\_clim\_1}=6 \quad \text{(distance allowed to climatology)}
   \]
   \[
   \text{alpha\_clim\_1}=2 \quad \text{(tolerance to take into account strong stratification)}
   \]
   \[
   \text{crit\_spike\_1} = 100,100 \quad \text{(criteria for spike detection. Example for an anlaysis of temperature and Salinity)}.
   \]

2. check the alert plots and flag with \text{QC} 4 or 8 the data considered wrong

3. perform a second pass with high values of the criteria (example in the configuration file)

Data rejected by the tests are not used to produce the STD files, but the corresponding value might be interpolated from the neighboring valid points. When all data from a profile are rejected the profile no longer appears in STD.

Examples for running STD:

Define the configuration file: config\_fname = 'my\_DIR\_ANA/confisas/config\_isas\_ana.txt/';

Then, launch STD\_main (config\_fname, [dd mm yyyy], nb\_days, ipass, plot\_display) with the set of input argument as follow:
to process 10 days starting on july 14, 2011, first pass:

\[
\begin{align*}
dd &= 14; \\
mm &= 07; \\
yyyy &= 2011; \\
nb\_days &= 10; \\
ipass &= 1; \\
plot\_display &= 1;
\end{align*}
\]

■ to process a full month (ex july 2006):

\[
\begin{align*}
dd &= 0; \\
mm &= 07; \\
yyyy &= 2011; \\
nb\_days &= 0; \% or anything, this value is ignored
\end{align*}
\]

An example of perl script to run STD_main over several month and years is given in the doc directory.

3.3. **Outputs**

3.3.1. **Data files on standard levels**

Results are written as NetCDF files in the directory: DIR\_RESU/\textit{std/}. The naming convention is as follows:

\[
\text{ST\_CCCCCCCC\_YYYYMMDD\_PR\_TT.nc}
\]

- \textit{ST} identifies « STD » data
- \textit{CCCCCCCC} the dataset name
- \textit{YYYYMMDD} the date of observation, if day = 00, file contains the whole month
- \textit{PR} identifies « profile » data
- \textit{TT} the data types according to Coriolis convention

3.3.2. **Listing (log file)**

The log files can be found in DIR\_RUN /logisas/ . For an analysis standardization step, there is two log files which names start with:

- \textit{std\_mess}: the detailed log file
- \textit{std\_summary}: the summary log file

The suffix of each file name is composed of the date of the std data followed by the date of the run.

For example, the detailed log file for a standardisation of july 2011, running in August 2011, the 12\textsuperscript{th}, will be: \textit{std\_mess\_20110714\_20110812.asc}

Two examples of both log files are given in the Annexes (8.2.1, 8.2.2).

3.3.3. **Control plots**

Different types of plots can be found in DIR\_RUN/plotisas/\textit{std}. 

15
Standard plot level (PLOT_CONV=1):

The nomenclature is:

ST_myanalysis_yyyymmdd_PR_XX.png

with  yyyymmdd: the standardisation date, XX: the data type (PR for Profiler, CTD, ...)

A plot showing all profiles is produced, for both Temperature and Salinity (if it is the parameter analyzed), as shown in figure 4.

![Example of standard plot level](image1)

Figure 1: Example of standard plot level

High level plot (PLOT_CONV>1):

The plot shown in figure 5 shows, for each analyzed parameter (most of the time Salinity and Temperature), the data standardized (red) and the raw data (blue).

![Example of high level plot](image2)

Figure 5: example of high level plot. The red cross are the data standardized, the blue one
represents the raw data.

Alerts

When data points are excluded, a plot is created in the directory alert/std (Figure 6). The profile reference is added to the list in the directory alert/list.

Figure 6: example of alert plots. Temperature and salinity data points in blue, climatology in black, corrected standard deviation criteria as dashed line. In red, the points excluded by the climatology test and in green the points excluded by the spike test. The plot title gives the DC-reference of the profile.
4. Analysis

4.1. Preprocessing (PREOA)

PREOA select the data that will be used to perform the analysis over each area. All data within the area mask and the time interval defined by date +/-AMPL_OA are selected. At this stage, data might be excluded on the instrument type criteria (INST_EXCL_LIST).

4.1.1. Running PREOA

After setting the parameters in the preoa block of the configuration file, PREOA can be launched in the matlab window.

- config_fname = 'my_DIR_ANA/confisas/config_isas_ana.txt/';
- launch PREOA_main(config_fname, [dd mm yyyy], parameter, plot_display)

An example of perl script to run PREOA_main over several months and years is given in /isas_v6/perl/ : preoa.pl

4.1.2. The output

PREOA outputs are:

- the temporary files 'fld' and 'dat' for each area, placed in the directory DIR_RUN /preoa /.
  The 'fld' files contain the empty anomaly filled for the area on the grid. The 'dat' files contain the data to be used by the analysis.
  Naming convention are as follows:
  
  OA_YYYYMMDD _iarea _typ_PARAM.nc
  OA  identifier for «optimal analyse »
  YYYYMMDD  analysis date
  iarea  area number
  typ  identifier «dat » ou « fld »
  PARAM  TEMP ou PSAL

- A copy of the ‘dat’ and ‘fld’ files in the data/ subdirectory of DIR_OA_CALCUL if option copy_preoa=1 is set in the configuration file. If not, run the matlab script PREOA_copy_file separately. DIR_OA_CALCUL is the directory of analysis created in the fortran computer for the analysis step. It is defined in the configuration file.

- The files TEMP.in and PSAL.in that contain the list of the areas to be processed. These files are created by the program PREOA_creat_configin and copied in the subdirectory config of DIR_OA_CALCUL if option creat_in_preoa=1 is set in the configuration file. If not run PREOA_creat_configin separately. An example of TEMP.in and PSAL.in is given in the annexe (8.3.1).

- A log file (8.3.2), which naming convention is as follows:
  preoa_mess_PARAM_yyyymmdd_YYYYMMDDHHMMSS.asc
  where  PARAM is the parameter (PSAL, TEMP, ...) 
  yyyymmd is the analysis date
YYYYMMDD is the date when the pre-processing has been run
HHMMSS is the hour when the pre-processing has been run

Warning: Erase all files from previous runs in preoa directory before running PREOA!
The process that copies the files on the fortran computer takes all files found in the
directory, files from previous runs which have not been overwritten will be taken
into account and may produce inconsistencies.

Figure 2: Position of data selected by PREOA to be used for
analyzing one area
4.2. Analysis

The analysis must be run on a computer with fortran compiler, Netcdf library and lapack/linpack. The programs (source) must be copied from isas_f90 subdirectory into the ‘software’ directory and compiled for the machine. For files from the subdirectory confstd/climref must be copied into a directory ‘confstd’:

- arglv502_ann_STD_PSAL.nc: a priori variance for the Salinity
- arglv502_ann_STD_TEMP.nc: a priori variance for the Temperature
- ARV09FDS_ann_COVS.nc: Covariance scales calculated from the Rossby radius.
- bathy_GLOBAL05_v5c2.nc: bathymetry

A directory ‘my_DIR_ANA_F90’ must be created for each specific analysis. This directory must exist before launching preoa with option 1 for the file copying (see previous section).

4.2.1. Analysis subdirectory on the fortran computer

The analysis directory is shown in the Figure 7. « my_DIR_ANA_F90 » is automatically created by the pre-processing step (the path and name is defined in the configuration file as DIR_OA_CALCUL).

Two directories have to be created by the user in the fortran computer:

- ISAS_F90: contains all the fortran scripts
- confstd: contains the annual variance and covariance, and the bathymetry
4.2.1.1. config

Contains the list of area created and automatically copied here by the preprocessing step for each parameter. Nomenclature is PARAM.in, where PARAM is the parameter analysed. Theses files contains the list of NetCDF files to be processed. An example is given in the annexes (8.3.1).

It should also contain the configuration file for the analysis (TEMP.cnf or PSAL.cnf), created here by the user. An example is shown below.

```
TEMP
/home2/mycomputer/user/OA/run/CONFSTD/ISASW_51_STD_TEMP.nc
/home2/mycomputer/user/OA/run/CONFSTD/ISASW_5_ann_COVS.nc
/home2/mycomputer/user/OA/run/CONFSTD/bathyGLOBAL05_V5_0.nc
300 300 21 % covar_ls x, y t (in km, km, days)
21 % covar_ms_t (in days)
1 1 4 % var_weigh (LS, MS, UR)
1 1 0 1 % x, y, z, t covariance dependency (1 = yes, 0 = no)
1.2 % fact. Variance
5 12 % QC Max  Mx_std
1.1 % Cov_max (if > 1, no oversampling test)
2 11 % oversample: alpha, fct_test (If fct_test < 10 increases error in the whole area)
```
4.2.1.2. data
Contains the ‘fld’ and ‘dat’ files created (and optionnally copied) by PREOA. Those files will be completed by OA.

4.2.2. Running ISAS_f90
The program can be run in interactive mode:
```bash
cd my_DIR_ANA_f90
calculateur/isas_f90/OA_main < config/TEMP_2011.in
```

It can also be launched in batch mode, this allows to loop over dates and parameters. The way batches are run is machine dependent. Examples are given here for SGI – ICE 8200. To process, launch the batch with:
```bash
qsub my_DIR_ANA_F90/batch/my_batch
```
where `my_batch` is as follows (and be created in the directory « batch »):
```bash
#!/bin/csh
# cd to the directory you submitted your job
cd /home1/caparmor/toto/OARUN/myanalysis/
# get the path for library MKL
source /usr/share/modules/init/csh
module load intel-comp/11.1.073
module load netcdf-intel/3.6.3-11.1.073

date
foreach year (2011)
    foreach month (09 10 11)
        foreach param (TEMP PSAL)
            /export/home1/toto/OA/versions/isas_f90_V6.2_s8/OA_main <
            config/$param\_year\_month\15.in
        end
    end
end
```

At the end of the job, a log message like `my_batch.o######` appears in the directory `my_DIR_ANA_F90/batch`. It is a summary of the information written in the log files in the directory `err` and `log`, detailed in the following.

4.2.3. Outputs

4.2.3.1. err
Contains a short log file with the list of processes files and any error message issued by the
program. This file must be screened carefully to check that the processing has ended normally. An example is given in the Annexes (8.4.1).

### 4.2.3.2. log

The log file contain statistical information on the processing for each area and each level of analysis. An example is given in the Annexes (8.4.2).

### 4.2.3.3. data files

The analysis results are stored in the directory «data', with the ‘fld’ and ‘dat’ data files that now contain the gridded anomaly fields and corresponding error and the data residuals, respectively.

### 4.3. Post-Processing (POSTOA)

During this last part of the processing, the program concatenates all processed areas and datasets. It also convert anomalies to absolute values. The files are read in DIR_OA_CALCUL (here, my\_DIR\_ANA\_F90/data) and results are written in DIR\_ANA\_RESU (here, DIR\_RESU).

#### 4.3.1. Running POSTOA

POSTOA is launched with the same arguments as PREOA. In the matlab window:

- config\_fname = 'my\_DIR\_ANA\_confisas/config\_isas\_ana.txt/';
- launch POSTOA\_main (config\_fname, [dd mm yyyy], parameter, plot\_display)

An example of perl script to run POSTOA\_main over several month and years is given in the doc directory.

#### 4.3.2. Outputs

The results are saved in two files:

- In DIR\_RESU/data, the NetCDF ‘dat’ file that contains the data and residuals used by all the areas .
- In DIR\_RESU/field the NetCDF file ‘fld’ that contains the global 3D gridded fields and error.

File naming convention is as follows:

- myanalysis\_YYYYMMDD\_typ\_PARAM.nc
- myanalysis analysis identifier (ANA\_NAME in the config file)
- YYYYMMDD analysis day (if DD='00', analysis month)
- typ identifier «dat » ou « fld »
- PARAM the parameter TEMP or PSAL

The processing can be checked by looking at the log files and plots.

The log file has the following naming convention, and an example is given in the annexe (8.5).

postoa\_mess\_PARAM\_yyyyymmdd\_YYYYMMDDHHMMSS.asc
where PARAM is the parameter (PSAL, TEMP, ...)

yyyymmdd is the analysis date

YYYYMMDD is the date when the pre-processing has been run

HHMMSS is the hour when the pre-processing has been run

The plots are stored in DIR_RUN/plotisas/postoa. The nomenclature is:

postoa_myanalysis_yyyymmdd_PARAM.png

where myanalysis is the analysis identifier (ANA_NAME in the config file)

yyyymmdd is the analysis date

PARAM is the parameter (PSAL, TEMP, ...)

Each plot contains four axes, as shown in the figure 8:

- the data position
- the error
- the anomaly
- the final field

Figure 8: example of post-processing plot, from an analysis of june 2010, only in the Atlantic. Top, from left to right: data position, and anomaly at 5m depth. Bottom from left to right: field and error at 5m depth.
5. Matlab_isas

5.1. The pre-processing of data files (STD program)

This step reads the NetCdf files provided by Coriolis, or prepared by the user according to a specific format. Profiles with valid data are selected and depth is computed if not present in the file, then two automatic controls are performed: comparison to climatology and spike detection. Points that fail the control are not used. Finally the profiles are interpolated on the analysis standard levels and the final STD NetCdf files are saved.

The main program is STD_main.m.

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
<th>Calls</th>
</tr>
</thead>
<tbody>
<tr>
<td>STD_main</td>
<td>Main function to prepare the dataset for the analysis</td>
<td>STD_list_PR, STD_list_TS, STD_list_hyd, STD_conv_raw, STD_red_list, STD_red_mean, rep_definition, ISAS_MSG, flddat_init, var2analyse, count_data, NCW_data_hdr, NCW_data_var, NCW_data_mult</td>
</tr>
<tr>
<td>STD_list_PR</td>
<td>Builds list of raw files to read and the list of the corresponding standardised files. The files are downloaded from Coriolis database, in NetCDF Argo format for daily files.</td>
<td>none</td>
</tr>
<tr>
<td>STD_list_TS</td>
<td>Build list of Time series raw files to read, and the list of the corresponding standardised files. The raw files must be in NetCDF Gosud format.</td>
<td>none</td>
</tr>
<tr>
<td>STD_list_hyd</td>
<td>Not used yet</td>
<td></td>
</tr>
</tbody>
</table>
| STD_conv_raw     | STD_conv_raw Creates standard files from raw profiles (Coriolis or user defined NC-files)  
1 - Extracts valid profiles from 'RAW' file  
2 - Computes depths when only pressure is provided  
3 - Adds climatology mean and std profiles and checks profile against climatology and looks for spikes  
4 - Interpolates to standard levels: | STD_read_PR, STD_read_TS, STD_read_hyd, STD_count, STD_deph, STD_clim_check, STD_spike_check, STD_plt_err, STD_vert_bins, STD_vert_ext_TS, STD_vert_interp |
| STD_red_list | Provide a list of profiles to average  
1 - looks for platforms with multiple profiles  
2 - tests vicinity of multiple profiles  
3 - defines the groups of profiles | none |
|STD_red_mean | Compute average profiles:  
Averages profiles from the same platform according to list provided by STD_red_list | count_data UT_prof_avg |
|STD_read_PR | STD_read_PR:  
Reads and extracts valid profiles (QC, area and time period) from file FILE_raw_.i.  
Select raw values or adjusted values | ISAS_MSG var2analyse STD_read_PR_data count_data STD_red_list STD_red_mean |
|STD_read_PR_data | Select valid data according to QC list and fill_value.Raw data are « Profiles » data. | none |
|STD_read_TS | Reads and extracts valid profiles from time series files | ISAS_MSG var2analyse STD_read_TS_data count_data |
|STD_read_TS_data | Select valid data from « TimeSeries » Nc file according to area and QC list. | none |
|STD_read_hyd | Not used | none |
|STD_count | counts valid profiles and redefine flags: | none |
|STD_deph | Compute depth when only pressure is provided by interpolating the depth/pressure profile of the nearest grid point of the climatology | NCR_OA_field |
|STD_clim_check | Check each profile against climatology using the standard deviation.  
Eliminated data are flagged 4. | none |
|STD_spike_check | Check spikes in each profile.  
The spike test looks for:  
1) a significative change in the sign of the first derivative  
2) high values of the second derivative - (normalized)  
Eliminated data are flagged 5 | none |
|STD_plt_err | Plot and save error plots | none |
|STD_vert_bins | Averages the data into bins centered on STD | UT_vert_bin |
 Extrapolate the TSG data from the level of measurement up to the surface

Interpolate profiles when nb_data points > NB_INT_MIN
- Flags the data according to interval between data points
- Increase error by a factor sqrt(QC)

Control plot for profiles

Control plot for time series

### 5.1.1. STD_main.m

Main function to prepare the dataset for the analysis reads Raw files and creates STD files

**Syntax:**

```
[] = STD_main(config_fname, DATE_INPUT, DAY_NUMBER, i_pass, PLOT_DISP)
```

**Input arguments**

- **config_fname**: full name of the configuration file
- **DATE_INPUT**: date of the analyse (format : [dd mm yyyy]). dd could be equal to 0 if you consider the full month mm.
- **DAY_NUMBER**: number of day for the analysis (STD_main input argument). If you want to consider the full month mm, set DAY_NUMBER to 0.
- **i_pass**: i_pass = 1 for a strict window profile selection, in order to detect suspicious one. i_pass = 2 for an enlarged window, as the suspicious one are considered controlled and get off the raw set if needed. Generally, STD_main is made first with i_pass = 1. The alerted profiles are controlled. Then a second run of STD is made, with i_pass = 2.
- **PLOT_DISP**: Plotting option makes figure visible or not, the figure are saved anyway. PLOT_DISP=0 : no display, PLOT_DISP=1 : display

**Output arguments**

There is no output argument. The standardisation is made, and the standardised files are created with the name and in the directory defined in the config_fname.

### 5.2. Pre- and Post analysis (PREOA/POSTOA)

The main programs are PREOA_main and POSTOA_main.

<table>
<thead>
<tr>
<th>rep_definition</th>
<th>NCR_OA_field</th>
<th>PREOA_file_list</th>
<th>NCR_data</th>
<th>PREOA_append</th>
<th>PREOA_select</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PREOA_main</strong> prepares the dataset that will be used to perform the analysis for the parameter and the list of areas prescribed in the config file. All data within the area area mask and the time interval defined by date +/-dtime_ana are selected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Function</td>
<td>Description</td>
<td>Input arguments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREOA_file_list</td>
<td>Builds dates list corresponding to STD files.</td>
<td>PREOA_copy_file, globatt_init, NCW_data_hdr, NCW_data_var, NCW_OA_field</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREOA_append</td>
<td>Appends new dataset to structures OAHDR and OAVAR</td>
<td>none</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREOA_select</td>
<td></td>
<td>none</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREOA_copy_file</td>
<td>Copies dat and fld NetCdf files created by PREOA on the remote computer</td>
<td>PREOA_copy_list</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Creates the config file with list of NetCDF files to use and copy them to the remote computer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREOA_copy_list</td>
<td>Copy a list of files from dir_in to dir_out</td>
<td>none</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSTOA_main</td>
<td>1- Collects data from all area files and saves a unique data file.</td>
<td>rep_definition, flddat_init, NCR_data, PREOA_select, NCW_data_hdr, NCW_data_var, NCR_OA_field, POSTOA_filter, NCW_OA_field</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% 2 - Merge field areas to form the global anomaly. Filters the field and</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% error where error is above threshold. Add to climatology and save the</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% final global parameter field Filters field and error on area boundaries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>where error is above threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSTOA_filter</td>
<td>Filters field and error on area boundaries where error is above threshold</td>
<td>none</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2.1. PREOA_main.m

Main function to prepare the dataset that will be used to perform the analysis over each area, and for one parameter. All data within the area mask and the time interval defined by date +/-AMPL_OA are selected. Two classes of files are created: the temporary files 'fld' and 'dat'.

- The 'fld' files contain the empty anomaly fields for the area on the grid.
- The 'dat' files contain the data to be used by the analysis.

**Syntax:**

```matlab
[] = PREOA_main(config_fname, DATE_EST, PARAM, PLOT_DISP)
```

**Input arguments**

- **config_fname**: full name of the configuration file
- **DATE_EST**: date of the analyse (format: [dd mm yyyy])
- **PARAM**: parameter to be pre-processed (PSAL, TEMP, DOXY, ...)
- **PLOT_DISP**: Plotting option makes figure visible or not, the figure are saved anyway. PLOT_DISP=0 : no display, PLOT_DISP=1 : display
Output arguments

No output argument. The pre-processing is made. Pre-processed files are created with the name and in the directory defined in config_fname.

5.2.2. POSTOA_main.m

The area files created by PREOA and completed by the analysis OA are processed to form global *.fld* and *.dat* files.

**POSTOA_main**

- Collects data from all area files and saves a unique data file.
- Merges field areas to form the global anomaly. A smoothing of the areas boundaries can be made where error is above a specified threshold. Finally, the anomaly is added to climatology to form the final field of the analyzed parameter.

**Syntax**

[] = POSTOA_main(config_fname, DATE_EST, PARAM, PLOT_DISP)

**Input arguments**

- **config_fname**: full name of the configuration file
- **DATE_EST**: date of the analyse (format : [dd mm yyyy])
- **PARAM**: parameter to be pre-processed (PSAL, TEMP, DOXY, ...)
- **PLOT_DISP**: Plotting option makes figure visible or not, the figure are saved anyway. PLOT_DISP=0 : no display, PLOT_DISP=1 : display

**Output arguments**

No output argument.

Post-processed files are created with the name and in the directory defined in config_fname.

A control plot is saved in the directory ISAS_RUN/plotisas/postoa

5.3. Tools

These matlab programs are used in each of the three steps standardisation, pre- and post-processing. They are sorted in alphanumerical order.

<table>
<thead>
<tr>
<th>Program</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rep_definition</td>
<td></td>
</tr>
<tr>
<td>flddat_init</td>
<td></td>
</tr>
<tr>
<td>var2analyse</td>
<td></td>
</tr>
<tr>
<td>count_data</td>
<td></td>
</tr>
<tr>
<td>NCR_OA_field</td>
<td></td>
</tr>
<tr>
<td>NCW_data hdr</td>
<td></td>
</tr>
<tr>
<td>NCW_data_var</td>
<td></td>
</tr>
<tr>
<td>NCW_data_mult</td>
<td></td>
</tr>
</tbody>
</table>
5.3.1. NCR_data.m
Select usefull data read in STD files, or '.dat' files. Used in PREOA_main and POSTOA_main.m.

Function

[STDHDR, STDVAR, STDQC] = NCR_data(fnam_data, PARAM, iopt_nan, pltnum, jlim, tab_msk, lat_msk, lon_msk)

Input arguments

• fnam_data: full name of the file to be read
• PARAM: parameter name (PSAL, TEMP, ...)
• iopt_nan: If iopt_nan is set to 1, the program replaces fill_value with NaN

optional arguments:

• pltnum: platform number (character). If pltnum = '0': no platform selection
• jlim: time limits in julian days
• tab_msk: geographical mask of the considered area (only sea = 2, only land = 0, sea and land = 1)
• lat_msk: latitudes of the considered area
• lon_msk: longitudes of the considered area

Output arguments

• STDHDR: structure with metadata and global attributes of the file fnam_data
• STDVAR: structure of data of the file fnam_data
• STDQC: structure with QC and DCref information (reduction, ...) of the read file

5.3.2. NCR_OA_field.m
Reads parameter field and coordinates from nc file.

Function

[lon_ana, lat_ana, dep_ana, param_OA, pct_var] = NCR_OA_field(file_nc_OA,PARAM,list_level,ilim,jlim)

Input arguments

• file_nc_OA: full name of the file to be read
• PARAM: parameter name (PSAL, TEMP, ...)

optional arguments:

• list_level: list of depth levels to be read
  absent : read all the depth levels
0 : read only the coordinates  
-1 : read all the depth levels  
-n : read depth levels from 1 to n  
else: read depth levels listed in list_level

- \textbf{ilim}: min and max latitude to be read; if absent: all the latitude domain is read
- \textbf{jlim}: min and max longitude to be read; if absent: all the longitude domain is read

\textbf{Output arguments}
- \textbf{lon}_\text{ana}: longitude read in the field file
- \textbf{lat}_\text{ana}: latitude read in the field file
- \textbf{dep}_\text{ana}: depth read in the field file
- \textbf{param}_\text{OA}: data read, corresponding to the parameter PARAM
- \textbf{pct}_\text{var}: variance percentage, corresponding to the parameter PARAM

\subsection{5.3.3. NCW_data_hdr.m}
Write headers (metadata) and their atytributes in data file that is created here.
\texttt{[msg\_error] = NCW\_data\_hdr(ncfile\_name, GLOB\_ATT, STDHDR )}

\textbf{Input arguments}
- \textbf{ncfile\_name}: the full name of the file to be created
- \textbf{GLOB\_ATT}: global attributes structures.
- \textbf{STDHDR}: structure with metadata and global attributes of a standardised file or of a '.fld' analysed file.

\textbf{Output arguments}
No output argument, except a message. The file is completed.

\subsection{5.3.4. NCW_data_var.m}
Writes data of the variable PARAM in data file. Used after NCW_data_hdr that creates the data file.
\texttt{[msg\_error] = NCW\_data\_var(ncfile\_name, PARAM, STDVAR)}

\textbf{Input arguments}
- \textbf{ncfile\_name}: full name of the file to be completed. File created by NCW_data_hdr.
- \textbf{PARAM}: parameter to be written (PSAL, TEMP, ...)
- \textbf{STDVAR}: structure with the data of the file ncfile\_name, for the parameter PARAM

\textbf{Output arguments}
No output argument, except a message. The file is completed.
5.3.5. NCW_OA_field.m

This function creates and write the '.fld' NetCDF files. These files contains fields : 3D grid (time, lat, lon) and the time.

**Function**

[msg_error] = NCW_OA_field(OA_filename, PARAM, GLOB_ATT, longitude, latitude, depth, jul_rel, FIELD, ERROR)

**Input arguments**

- **OA_filename**: Full name of file to be created
- **PARAM**: parameter ('TEMP', 'PSAL', 'TPOT', 'SIG0', 'TEMP_STD', 'PSAL_STD')
- **GLOB_ATT**: structure setting the global attributes
- **longitude**: Vector of longitudes (its length defines the longitude dimension)
- **latitude**: Vector of latitudes (its length defines the latitude dimension)
- **depth**: Vector of depth (its length defines the depth dimension)
- **jul_rel**: Julian day referenced to the Reference Date Time read in the standardised file (most of the time, 01/01/1950).
- **FIELD**: 3D gridded field FIELD(longitude, latitude, depth) and time
- **ERROR**: 3D gridded field of error define as percent variance

**Output arguments**

No output argument, except a message. The file is written.
6. Optimal interpolation: (f90_isas)

6.1. Program list

<table>
<thead>
<tr>
<th>Program</th>
<th>Description</th>
<th>Dependencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>OA_main.f90</td>
<td>Main program for Objective Analysis Coordinates loop over areas</td>
<td>OA_anaarea</td>
</tr>
<tr>
<td>OA_anaarea</td>
<td>Analysis of a field on a regular grid for one area</td>
<td>OA_ncreaddim OA_ncreadstat OA_ncreaddata OA_covini OA_oversamp OA_calsol OA_ncwrite_fld OA_ncwrite_dat</td>
</tr>
<tr>
<td>OA_ncreaddim</td>
<td>Reads dimensions in nc file prepared by pre-OA</td>
<td>none</td>
</tr>
<tr>
<td>OA_ncreadstat</td>
<td>Reads field grid and covariances in nc file prepared by pre-OA</td>
<td>none</td>
</tr>
<tr>
<td>OA_ncreaddata</td>
<td>Reads data in nc file prepared by pre-OA</td>
<td>none</td>
</tr>
<tr>
<td>OA_covini</td>
<td>Prepares covariance matrices by computing the distance dependent part</td>
<td>none</td>
</tr>
<tr>
<td>OA_oversamp</td>
<td>Computes weights to increase error in case of oversampling</td>
<td>none</td>
</tr>
<tr>
<td>OA_calsol</td>
<td>Computes OA solution: (X_est - X_0) = Koa (Y - Y_0),</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>Koa = C_md (C_dd + R)**-1,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>diag(C_est) = R = diag(data_var),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>where:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y - Y_0 : dino</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X_est - X_0 : ana_fld</td>
<td></td>
</tr>
<tr>
<td></td>
<td>diag(C_est) : ana_var_ps</td>
<td></td>
</tr>
<tr>
<td>OA_ncwrite_fld</td>
<td>Writes analysis results in nc file prepared by pre-OA: gridded field</td>
<td>none</td>
</tr>
<tr>
<td>OA_ncwrite_dat</td>
<td>Writes analysis results in nc file prepared by pre-OA: data vector and residuals</td>
<td>none</td>
</tr>
</tbody>
</table>

7. References


8. ANNEXES

8.1. Configuration file

%===================================================================
% Configuration file for ISAS-V6
% F. Gaillard, LPO/Ifremer
% E. Brion, LPO/CNRS
%===================================================================
% Attention: no space within lines (unless necessary)
%===================================================================
% Language: EN or FR
LANG=En
%===================================================================
% Metadata information for NC files
DATA_SET=YYYYYYY
INSTITUTION=XX/IFREMER
PROJECT_NAME=MyProgName
DATA_MANAGER=You
SOURCE=ISAS-V6
%===================================================================
% Directory and file names for standard configuration
%===================================================================

% Configuration files: Climatology, STD, areas
%===================================================================
DIR_CONFSTD=/net/toto/export/home1/arivo/software/isas/isas_ana/confstd_v5c3/
nam_clim=arglv502
nam_std=arglv502_ann_STD
nam_bathy=bathy_GLOBAL05_v5c2.nc
%===================================================================

% RAW DATA
% -------
% Raw data directory and files prefix for profiles - Coriolis files
DIR_RAW_ROOT_PRC=/net/mymachine/export/home1/DATA_WORK/NRTOAGL01/
PRF_RAW_PRC=CO_NRTOAGL01_
% Raw data directory and files prefix for profiles - User files (hydro or profilers)
%DIR_RAW_ROOT_PRU=/net/mymachine/export/home1/DATA_WORK/NRTOAGL01/
%PRF_RAW_PRU=CO_NRTOAGL01_

% Raw data directory and files prefix for time series-TSG
DIR_RAW_ROOT_TSG=/net/toto/export/home1/SOERE-CTDO2-TSGRecherche/NRTGV2_nc/

% Raw data directory and files prefix for time series-buoys
% DIR_RAW_ROOT_TSB

% Raw data directory and files prefix for time series-moorings
% DIR_RAW_ROOT_TSM

% STD DATA
% --------
% STD data directory and files prefix for Profiles
DIR_STD_ROOT_PR=/net/mymachine/export/home1/std_data_NRT/isasv6xx_PR/
% PRF_STD : Prefix for std files
PRF_STD_PR=ST_isasv6xx_

% STD data directory and files prefix for TIMESERIES
DIR_STD_ROOT_TS=/net/mymachine/export/home1/std_data_NRT/isasv6xx_TS/
% PRF_STD_PR=ST_isasv6xx_

% Analysis
% --------
% Directory for Analysis results and logs
DIR_ANA_ALL=/net/mymachine/export/home1/analysis_NRT/isasv6xx/
DIR_ANA_RESU=/net/mymachine/export/home1/analysis_NRT/isasv6xx/ISAS_RESU/
% Directory for f90 computer
DIR_OA_CALCUL=/net/caparmor-nas/mnt/data2/home1/you/OARUN/arglv6XX/

%===================================================================
str_area_limits=[-81 +90 -180 +180]

%===================================================================

%===================================================================
% STD: Standardisation
%===================================================================

% TYP_LIST: List of file types to process
TYP_LIST=PR_CT,PR_PF,PR_TE,PR_XB,PR_MO,PR_BA
% TYP_LIST=PR_PF,PR_CT,PR_MO,PR_TE

GREY_LIST_STD=model_greylist.txt

% Definition of default measurement error associated with each file types
% Will be used if no error information is found in the file
PR_TE_ERR=0.03
PR_BA_ERR=0.05
PR_PF_ERR=0.01
PR_XB_ERR=0.03
PR_CT_ERR=0.01
PR_MO_ERR=0.01
PR_BH_ERR=0.002
TS_TS_ERR=0.02

% month_grp=1 : all data within a month are grouped, no group = 0
month_grp=1

% Variables to analyse (lower case)
VAR=psal,temp
% Variables to analyse, specific to timeSeries(lower case)
VARTimer=ssps,sstp

% use_adjust=1 : use adjusted value if exist (else: =0)
use_adjust_PR=1
use_adjust_TS=0

% Datacenter QC used
%---------------------------------------------------------------
% Datacenter QC are the following:
% 0 No quality control (QC) was performed.
% 1 QC was performed: good data.
% 2 QC was performed: probably good data.
% 3 QC was performed: probably bad data.
% 4 QC was performed: bad data.
% 5 The value was changed as a result of QC.
% 8 Interpolated value
% 9 The value is missing
% %
% % QC_TS: flags ok Temp and Psal
% QC_ZP: flags ok Pres and depth
% QC_XY: flags ok Position and date
% Temporarily two sets: char for Argo, num for TSG
% uses also error max to select data
QC_TS=125
QC_ZP=0 125
QC_XY=0 125
QC_TS_num=1,2,5
QC_ZP_num=0,1,2,5
QC_XY_num=0,1,2,5

ERR_MAX_PRES=20
ERR_MAX_TEMP=1
ERR_MAX_PSAL=1
%---------------------------------------------------------------

% Parameters for profiles controls
% ------------------------------------------
% Criteria for comparison with climatology and spike detection
% The processing is performed in two passes, so 2 values are proposed.
% The maximum deviation relative to the reference climatology is defined
% by
% crit_std_clim: that sets the number of standard deviations allowed for
% climatology comparison
% alpha_clim: that add a tolerance to take into account the strong
% stratifications
% crit_spike: that sets the number of standard deviations allowed for
% spike detection. Number of
% value depending on the number of variable analysed.
% Pass1 recommended values, for TEMP and PSAL parameter analysis
% crit_std_clim_1=6
% alpha_clim_1=2
% crit_spike_1=100,100
%
% Pass2
% crit_std_clim_2=80
% alpha_clim_2=10
% crit_spike_2=2000,2000
%
str_crit_std_clim_1=6
str_alpha_clim_1=2
str_crit_spike_1=100,100
%
str_crit_std_clim_2=80
str_alpha_clim_2=10
str_crit_spike_2=1000,1000
%
% INT_NB_MIN : Minimum number of data points for interpolating a profile.
INT_NB_MIN=5
%
% Parametres for the reduction of nearby profiles
% str_red_opt24 =1: applies 24h average of 1 level profiles
% RED_DXMAX: maximum allowed distance (km)
% RED_DTMAX: maximum allowed interval of time (days)
% str_red_qcmax: maximum quality flag of interpolation taking into account to make the average profiles.
str_red_opt24=1
RED_DXMAX=15
RED_DTMAX=7
str_red_qcmax=4
%
% NPROF_MAX: Maximum number of raw profiles found to form one std file
NPROF_MAX=90000
% NTIMS_MAX: Maximum number of raw TSG found to form one std file
NTIMS_MAX=100000
%
str_plot_std % Plotting option for STD
% 0 = no plots,
% 1 = summary plot + alert plot,
% 2 = 1 + ctrl plot
% 3 = 2 + plot each profile + pause
str_plot_std=1
%
% Specifique 'TSG'
% FILE_LIST : list of raw files to explore (file.txt) corresponding to
% the list of TSG files of the DIR_RAW_ROOT_TSG
FILE_LIST=TS_raw_list.txt
%
% LEV_SURF : level of standardisation
LEV_SURF = [0:20]
% Definition of the erro multiplicative factor for the surface processing
ERR_COEF=[1.5 2]

%===================================================================
%===================================================================
%===================================================================
%===================================================================

% PREOA: Preprocessing
%===================================================================

% ANA_AREA_LIST: list of geographical areas to analyze

% AMPL_OA: time interval around date of analysis in days
% data from jest-dtime_ana to jest+dtime_ana will be used
%str_dtime_ana=41
str_dtime_ana=41

% Levels selected for the analysis
% str_deph_ana=-1 : no selection, keep all the levels.
% str_deph_ana=[0, 5, 20, 100]: list of STD levels to analyze
% (the closest standard levels will be selected
str_deph_ana=-1

% str_qcmax_ana: max STD-QC taken into account (scale: QC_interp)
str_qcmax_ana=4

% Instrument type excluded
%INST_EXCL_LIST=[[1:800),900] : exclude XBTs of all types
%INST_EXCL_LIST=[]
str_list_excl_inst=[(1:800),900]

% After PREOA, copies NetCDF files on the remote computer
% and creates config files
% 1 = yes
str_copy_preoa=1

%===================================================================
%===================================================================
%===================================================================
%===================================================================

% POSTOA: Post-Processing
%===================================================================

% Analysis name
ANA_NAME=ARGLV6XX

% clim_ref_oa=M (month) or clim_ref_oa=Y (year)
clim_ref_oa=M

% Spatial filtering allowed on point with err>err_max
str_filter_errmax=80

%for plotting : level (in meter) to be plotted by post-oa, if PLOT_LEV < 0, no plot
str_plot_lev=5

%===================================================================
%===================================================================
%===================================================================
%===================================================================
8.2. log files of the standardisation step

8.2.1. message file

>>>>>> Running ISAS_V6/STD

Last update :20-Dec-2011 16:03:12
Configuration files :
climref: /home5/garo/arivo/software/isas/isas_ana/confstd_v5c3/climref/
bathy:   /home5/garo/arivo/software/isas/isas_ana/confstd_v5c3/climref/
area:    /home5/garo/arivo/software/isas/isas_ana/confstd_v5c3/areadef/

Type_list:
PR_PF
PR_CT
PR_MO
PR_TE

RAW data directory: /net/luarca/export/home1/DATA_WORK/NRTOAGL01/
STD data directory: /net/luarca/export/home1/std_data_NRT/arglv6NRT_PR/

STD: Type PR_PF, 31 files found

/net/luarca/export/home1/DATA_WORK/NRTOAGL01/2011/CO_NRTOAGL01_20110101_PR_PF.nc
/net/luarca/export/home1/DATA_WORK/NRTOAGL01/2011/CO_NRTOAGL01_20110102_PR_PF.nc
/net/luarca/export/home1/DATA_WORK/NRTOAGL01/2011/CO_NRTOAGL01_20110103_PR_PF.nc
/net/luarca/export/home1/DATA_WORK/NRTOAGL01/2011/CO_NRTOAGL01_20110104_PR_PF.nc
...
/net/luarca/export/home1/DATA_WORK/NRTOAGL01/2011/CO_NRTOAGL01_20110131_PR_PF.nc

******* File  1  *******

Starting STD_convprof

Starting STD_read_Prof - File processed: CO_NRTOAGL01_20110101_PR_PF
Number of profiles found in file : 271
Number of profiles with valid position/time : 271

STD_read_xxdata output: selection of raw/adjusted profiles
Variable:psal, Nb raw profiles QC ok : 127, Nb elements QC ok : 16063
  , Nb adj profiles QC ok : 144, Nb elements QC ok : 11293
Variable:temp, Nb raw profiles QC ok : 125, Nb elements QC ok : 16253
  , Nb adj profiles QC ok : 146, Nb elements QC ok : 11497
Variable:pres, Nb raw profiles QC ok : 125, Nb elements QC ok : 16255
  , Nb adj profiles QC ok : 146, Nb elements QC ok : 11518
the variable deph does not exist
STD_read_Prof end

Number of profiles with valid S/T data (good QC and no NaN) : 271
distribution per type: S: 0, T: 7, S/T: 264,

Number of profiles without deph: before 271, after 0
Number of profiles with valid S/T data (good QC and no NaN): 271
distribution per type: S: 0, T: 7, S/T: 264,

Number of profiles with std-alert S: 0, T: 0
Number of profiles with valid S/T data (good QC and no NaN): 271
distribution per type: S: 0, T: 7, S/T: 264,

Interpolation on std_levels
Number of profiles with valid S/T data (good QC and no NaN): 271
distribution per type: S: 0, T: 7, S/T: 264,
271 profiles, CPU time total (seconds): 87.84
CPU read: 0.36, deph: 26.62, STD check: 53.83, Stdlev: 34.01

********** File 2 **********

Starting STD_convprof

Starting STD_read_Prof - File processed: CO_NRTOAGL01_20110102_PR_PF
Number of profiles found in file :252
Number of profiles with valid position/time :252

.....

8.2.2. summary file

*** std_summary: 20110701_20111109102322

<<<
==================================
=====
<<< Final number of profiles per type:
<<<
==================================
=====
<<< before time/space reduction after reduction
<<< Nfound NokST NnoZ Nalert N_pstd NS NT NST Total
<-> 9494 9471 12 5 9459 0 236 8826 9062
ST_arglv6NRT_20110700_PR_PF.nc
<-> 1635 1426 131 1 1295 0 10 250 260
ST_arglv6NRT_20110700_PR_CT.nc
<-> 2796 2796 0 0 2796 0 55 395 450
ST_arglv6NRT_20110700_PR_MO.nc
<-> 113857 36266 0 0 36266 0 535 2048 2583
ST_arglv6NRT_20110700_PR_TE.nc
>>>
8.3. **output files of the pre-processing step**

8.3.1. **List of area processed : example with a PSAL.in**

```
data/2010/PSAL/
log/2010/PSAL/
config/PSAL.cnf
```

OA_20100615_101_dat_PSAL.nc
OA_20100615_103_dat_PSAL.nc
OA_20100615_105_dat_PSAL.nc
OA_20100615_106_dat_PSAL.nc
OA_20100615_107_dat_PSAL.nc
OA_20100615_109_dat_PSAL.nc
OA_20100615_110_dat_PSAL.nc
OA_20100615_111_dat_PSAL.nc
OA_20100615_112_dat_PSAL.nc
OA_20100615_113_dat_PSAL.nc
OA_20100615_115_dat_PSAL.nc
OA_20100615_116_dat_PSAL.nc
OA_20100615_117_dat_PSAL.nc
OA_20100615_118_dat_PSAL.nc
OA_20100615_120_dat_PSAL.nc
OA_20100615_122_dat_PSAL.nc
OA_20100615_123_dat_PSAL.nc
OA_20100615_124_dat_PSAL.nc
OA_20100615_125_dat_PSAL.nc
OA_20100615_126_dat_PSAL.nc
OA_20100615_127_dat_PSAL.nc
OA_20100615_129_dat_PSAL.nc
OA_20100615_130_dat_PSAL.nc
OA_20100615_131_dat_PSAL.nc
OA_20100615_133_dat_PSAL.nc
OA_20100615_134_dat_PSAL.nc
OA_20100615_135_dat_PSAL.nc
OA_20100615_136_dat_PSAL.nc
OA_20100615_137_dat_PSAL.nc
OA_20100615_138_dat_PSAL.nc
OA_20100615_140_dat_PSAL.nc
OA_20100615_141_dat_PSAL.nc

8.3.2. **log file**

```
>>> Running ISAS_V5.3b/PREOA

Last update :30-May-2011 14:18:32
STD_PRO:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLVPRT/EXP_MODIF_LEST/std/
STD_TIMSER:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLVPRT/EXP_MODIF_LEST/std/
PREOA:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLVPRT/EXP_MODIF_LEST/ISAS_RUN/preoa/

    nb area: 41, nb_std: 5, param: TEMP

File dates:
20100500
```
File types:
  PR_PF
  TSG

***** Starting area 101 *****
File:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLV5TEST/EXP_MODIF_LEST/std/2010/S
T_expmodlest_20100600_PR_PF.nc
  7 data
File:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLV5TEST/EXP_MODIF_LEST/std/2010/S
T_expmodlest_20100600_TSG.nc
  278 data

0 data removed on type criteria
Ok: NCW_data_hdr
Ok: NCW_data_var
Ok: NCW_data_mult
OK: NCW_OA_field

  area 101, nb_data:  285, nb_prof_tot:  3281, processing time:  33.55 sec
************************************************

***** Starting area 102 *****
File:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLV5TEST/EXP_MODIF_LEST/std/2010/S
T_expmodlest_20100600_PR_PF.nc
  0 data
File:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLV5TEST/EXP_MODIF_LEST/std/2010/S
T_expmodlest_20100600_TSG.nc
  0 data

  area 102, nb_data:  0, nb_prof_tot:  0, processing time:  0.04 sec
************************************************

***** Starting area 103 *****
File:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLV5TEST/EXP_MODIF_LEST/std/2010/S
T_expmodlest_20100600_PR_PF.nc
  0 data
File:
/net/spineg/export/home/ebrion/ISAS/ANALYSES/ARGLV5TEST/EXP_MODIF_LEST/std/2010/S
T_expmodlest_20100600_TSG.nc
  223 data

0 data removed on type criteria
Ok: NCW_data_hdr
Ok: NCW_data_var
Ok: NCW_data_mult
OK: NCW_OA_field

  area 103, nb_data:  223, nb_prof_tot:  1614, processing time:  26.75 sec
************************************************

...
8.4. log and error file for the Analysis fortran step

8.4.1. error file

Run: OA_20100615_PSAL
OA-Version 5.1

PARAM: PSAL, nb_area: 32

Configuration files used:
A-priori variance :
/home1/caparmor/fgaillar/OA/CONFSTD-V5C3/arglv502_ann_STD_PSAL.nc
Covariance scales :
/home1/caparmor/fgaillar/OA/CONFSTD-V5C3/ARV09FDS_ann_COVS.nc
Bathymetry :
/home1/caparmor/fgaillar/OA/CONFSTD-V5C3/bathy_GLOBAL05_v5c2.nc

Parameters for optimal estimation:
Large scale covariance (x,y,t): 300000.0 300000.0 30.0
Meso scale covariance (t) : 30.0
Variance weights (LS, MS, UR) : 1.0 2.0 6.0
Covariances used (x, y, z, t) : 1 1 0 1
Factor multiplying apr-var : 1.5
qc_max, mx_std: 5 12
Oversampling: cov_max, alpha, fct_test: 0.98000 2.00 11.0

Analysis name:OA_20100615_101_dat_PSAL
Analysis name:OA_20100615_103_dat_PSAL
Analysis name:OA_20100615_105_dat_PSAL
Analysis name:OA_20100615_106_dat_PSAL
Analysis name:OA_20100615_107_dat_PSAL
Analysis name:OA_20100615_109_dat_PSAL
...

Analysis name:OA_20100615_141_dat_PSAL
4.50828230381012

8.4.2. log file

Example for the file: OA_20100615_101_dat_PSAL.log
8.5. log file of the post-processing steps

ARGLV6NRT PSAL  20110115

ARGLV6NRT PSAL  20110115

filter_err_max = 80, AMPL_OA = 41.0

Area 504 missing on remote computer
Area 505 missing on remote computer
Area 506 missing on remote computer
    >>>> Copy terminated in 105.067581 sec, 176 area found

Number of OA data: 27136

Ok: NCW_data_hdr
Ok: NCW_data_var
    >>>> Processing time for data: 15.35

    ... Processing time for field anomaly: 55.92
    ... Processing time for smoothing: 347.15
    ... Processing time for absolute field: 76.18