Outil de copie des QC et données ajustées entre 2 fichiers NetCDF mono-profils

Cette note concerne l’outil *nc\_copy\_mono\_profile\_qc* (V1.01) utilisé pour recopier les QC et données ajustées d’un fichier NetCDF mono-profil dans un autre.

[Outil de copie des QC et données ajustées entre 2 fichiers NetCDF mono-profils 1](#_Toc387651374)

[1. Description de la livraison 1](#_Toc387651375)

[2. Pré-requis 1](#_Toc387651376)

[3. Installation/configuration de l’outil 1](#_Toc387651377)

[4. Mise en oeuvre 2](#_Toc387651378)

[5. Principe de fonctionnement 2](#_Toc387651379)

[5.1. Mise en correspondance des fichiers 2](#_Toc387651380)

[5.2. Comparaison du contenu des fichiers 2](#_Toc387651381)

[5.3. Copie des informations de QC et d’ajustement 3](#_Toc387651382)

[5.3.1. Utilisation du schéma NetCDF du fichier Input 3](#_Toc387651383)

[5.3.2. Copie des données du fichier Input 3](#_Toc387651384)

[5.3.3. Copie des données du fichier Qc 3](#_Toc387651385)

[6. Sorties 4](#_Toc387651386)

# Description de la livraison

L’outil est composé du fichier *nc\_copy\_mono\_profile\_qc.m*.

# Pré-requis

L’outil a été écrit et testé avec Matlab R2013b.

# Installation/configuration de l’outil

Aucune installation particulière n’est nécessaire.

L’outil se configure en renseignant les premières lignes du fichier *nc\_copy\_mono\_profile\_qc.m* :

* DIR\_INPUT\_NC\_FILES : répertoire des fichiers de référence pour les données brutes,
* DIR\_QC\_NC\_FILES : répertoire des fichiers de référence pour les QC et données ajustées,
* DIR\_OUTPUT\_NC\_FILES : répertoire des fichiers générés,
* DIR\_LOG\_FILE : répertoire de stockage du fichier *.log* produit à chaque lancement de l’outil,
* FLOAT\_LIST\_FILE\_NAME : liste des numéros WMO des flotteurs à traiter par défaut.

La structure des répertoires DIR\_INPUT\_NC\_FILES, DIR\_QC\_NC\_FILES et DIR\_ OUTPUT\_NC\_FILES doit être identique à celle du FTP Argo, à savoir :

* Un répertoire par flotteur (nommé du numéro WMO du flotteur),
* Pour chaque répertoire flotteur, un sous répertoire « *profiles* » contenant les fichiers NetCDF mono-profils.

# Mise en oeuvre

La mise en œuvre se fait en exécutant le programme *nc\_copy\_mono\_profile\_qc :*

* Avec arguments (Ex : nc\_copy\_mono\_profile\_qc(2901496, 6901027)) pour traiter les flotteurs dont les numéros WMO sont fournis en argument,
* Sans argument pour traiter tous les flotteurs de la liste FLOAT\_LIST\_FILE\_NAME.

# Principe de fonctionnement

L’outil crée un fichier **Output** dont :

* Le schéma NetCDF, la majorité des données et les mesures brutes proviennent du fichier **Input**,
* Les QC, mesures ajustées et variables associées proviennent du fichier **Qc** associé au fichier **Input**.

Le programme fonctionne en trois étapes :

* Mise en correspondance des fichiers **Input** et **Qc**,
* Comparaison du contenu des fichiers **Input** et **Qc** pour vérifier la faisabilité de la copie des informations de QC et d’ajustement,
* Copie des informations de QC et d’ajustement.

## Mise en correspondance des fichiers

Les jeux de fichiers **Input** et **Qc** pouvant être hétérogènes (différences de numéros de cycles ou de datation des profils), il est nécessaire d’associer à chaque fichier **Input** son fichier **Qc** le plus probable.

Cette association se fait sur les dates des profils (à défaut sur les dates des positions des profils).

Pour chaque profil **Input**, de date *ProfDateInput*, on recherche le profil **Qc** associé dont la date *ProfDateQc* est telle que :

*| ProfDateInput – ProfDateQc | ≤ 10 heures* (INTERVAL\_HOUR = 10).

**Un profil non daté est ignoré.**

Si aucun fichier **Qc** ne peut être associé à un fichier **Input**, le fichier **Output** est créé à partir d’une copie du fichier **Input** (et un message WARNING est généré dans le log).

## Comparaison du contenu des fichiers

La comparaison du contenu des fichiers **Input** et **Qc** porte sur les valeurs des paramètres mesurés.

Plus précisément, pour chaque niveau du profil, les données de tous les paramètres du profil sont transformées en chaîne de caractères (en utilisant l’attribut C\_format associé à chaque paramètre).

Lorsque les données sont différentes, la copie des informations de QC et d’ajustement ne peut être réalisée (et un message ERROR est généré dans le log).

Ainsi, l’étape de copie des informations de QC et d’ajustement n’est réalisée que si les fichiers **Input** et **Qc** ont :

* La même liste de paramètres,
* Le même nombre de niveaux,
* Les mêmes mesures de paramètres (à la précision fournie par l’attribut C\_format associé à chaque paramètre).

Le fichier **Input** et/ou **Qc** peut contenir deux profils (N\_PROF = 2). Dans ce cas, on s’assure que ces deux profils sont issus d’un même profil découpé en données de CTD pompées/non pompées.

Les critères utilisés pour reconnaitre un profil découpé sont :

* N\_PROF = 2,
* Les deux profils ont la même date,
* Les deux profils ont la même liste de paramètres.

Lorsqu’un profil a été découpé, il est reconstitué préalablement à l’étape de comparaison.

## Copie des informations de QC et d’ajustement

Le fichier **Output** est constitué au cours des étapes suivantes.

### Utilisation du schéma NetCDF du fichier Input

Le fichier **Output** est créé sur la base du schéma NetCDF du fichier **Input** dans lequel on a au préalable modifié la dimension N\_CALIB qui doit être identique à celle du fichier **Qc**.

### Copie des données du fichier Input

Le contenu du fichier **Input** est recopié dans le fichier **Output**, sauf :

* Les QC et données ajustées,
* Les QC globaux relatifs aux profils qui sont recalculés,
* Les informations de calibration,
* Les diverses données qui seront copiées du fichier **Qc** lors de l’étape suivante.

### Copie des données du fichier Qc

Au contenu du fichier **Output** issu de l’étape précédente sont ajoutées les données suivantes issues du fichier **Qc** :

* Les données de QC (<PARAM>\_QC) du fichier **Output** sont celles du fichier **Qc**,
* Les données d’ajustement (<PARAM>\_ADJUSTED, <PARAM>\_ADJUSTED\_QC et <PARAM>\_ADJUSTED\_ERROR) du fichier **Output** sont celles du fichier **Qc.**

**Attention, les données d’ajustement des paramètres intermédiaires ne sont pas recopiées. La règle d’exclusion de ces variables est fournie par le champ 'adjAllowed' associé à chaque paramètre dans la fonction get\_netcdf\_param\_attributes\_3\_1.**

* Les données de calibration (PARAMETER, SCIENTIFIC\_CALIB\_EQUATION, SCIENTIFIC\_CALIB\_COEFFICIENT, SCIENTIFIC\_CALIB\_COMMENT et SCIENTIFIC\_CALIB\_DATE ou CALIBRATION\_DATE selon la version) du fichier **Output** sont celles du fichier **Qc**,
* Les données historiques (HISTORY\_INSTITUTION, HISTORY\_STEP, HISTORY\_SOFTWARE, HISTORY\_SOFTWARE\_RELEASE, HISTORY\_REFERENCE, HISTORY\_DATE, HISTORY\_ACTION, HISTORY\_PARAMETER, HISTORY\_START\_PRES, HISTORY\_STOP\_PRES, HISTORY\_PREVIOUS\_VALUE et HISTORY\_QCTEST) du fichier **Output** sont celles du fichier **Input** auxquelles on ajoute celles du fichier **Qc**,
* Les données DC\_REFERENCE, DATA\_STATE\_INDICATOR et DATA\_MODE du fichier **Output** sont celles du fichier **Qc**,

Concernant les données historiques :

* Lorsque le fichier **Input** comporte deux profils et le fichier **Qc** un seul, les informations de ce profil sont dupliquées dans chaque profil du fichier **Output**,
* Lorsque le fichier **Input** comporte un profil et le fichier **Qc** deux profils, seules les informations du premier profil (N\_PROF = 1) sont copiées dans le fichier **Output**.

# Sorties

Le programme produit un fichier *log* contenant les messages d’information (type : INFO, WARNING ou ERROR) générés lors du traitement.