

**Inventaire suite de production de figures
Pour les fichiers de types postalt et evalalt et dbase**

**Chantal Côté
2024-04**

Table des matières

Exemple pour ajouter une famille	3
Production des figures.....	4
Production d'une page web.....	4
Exemple pour brancher le site de monitoring sur une suite	5
Les figures 6h (titres mauves de la page web)	7
1. Scatter_maps.....	7
2. map_boites.....	9
3. Timebin.....	11
4. Profils.....	13
Figures series (titres rouges de la page web)	15
1. series_nobs.....	15
2. series_temporelles	17
3. Zonaux	19
4. Series_anomalies.....	21
Séries mensuelles (titres jaunes de la page web)	24
1. Mensuelles_nprof_boites.....	24
2. Traces.....	26
3. Séries temporelles mensuelles O-P	27
4. Maps mensuelles O-P	29
5. Tables html du dernier rapport des stations.....	30
Statistiques combinées (titres verts de la page web).....	31
1. Stations suspectes	31
Bases de données d'agrégation.....	33
1. La table avg.....	33
2. La table count_profils_assim	34
3. La table resume	34
4. La table last_report_stn	34
5. Les tables soft_limits	35
Les bases d'agrégations mensuelles.....	35

1. La table moyenne_avg.....	35
2. La table moyenne_nprofs.....	35
3. La table resume	35

Pour produire les figures, on assume que la suite Burp2Rdb_conversion a été tournée préalablement sur les dates désirées. Les bases de données d'observations sont disponibles en format sqlite.

Sauf indication contraire, toutes les figures peuvent accepter les 3 types de données (dbase, evalalt et postalt) et toutes les familles ci-dessous :

ai ai_bcor ai_nobcor airs atms cris csr gp iasi metar pr ra ro sc sf ssmis sw sw_geos sw_polaire
sw_polaireDB swob to_amsua to_amsua_allsky to_amsua_rars to_amsua_allsky_rars to_amsub
to_amsub_rars ua

Les variables

MONITORING_POSTALT_FAM_LISTE

MONITORING_EVALALT_FAM_LISTE

MONITORING_DBASE_FAM_LISTE

devraient idéalement contenir toutes les familles qui ont été converties par la suite

Burp2Rdb_conversion. Il s'agit d'inclure la liste des familles qu'il est possible de traiter. Ceci est nécessaire entre autre pour la production du site web.

La variable MONITORING_RARS_LISTE contient le nom des familles provenant d'un réseau local ou dit « direct broadcast ». Ces familles sont produites lors de la conversion.

La variable MONITORING_SOUS_FAMILLES contient le nom des familles qui seront produites à partir d'une famille complète. Par exemple, les familles « ai_bcor » et « ai_nobcor » seront produites à partir de la famille souche « ai » dans le monitoring (pas la conversion).

Exemple pour ajouter une famille :

Certaines familles peuvent être créées pour un type particulier. Par exemple, on peut créer ua_tac ou ua_bufc pour les dbases seulement. Il faut alors s'assurer d'ajouter un fichier de configuration et le nom de ce fichier dans le container du type de figure que l'on veut avant de pouvoir utiliser ce nouveau nom. On peut alors ajouter le nouveau nom dans la liste des familles MONITORING_*_FAM_LISTE ou MONITORING_SOUS_FAMILLES et/ou dans la liste pour une figure en particulier (voir plus bas). Si la nouvelle famille doit être combinée à une autre sur la page web, il faut aussi ajouter son nom dans les configs combine.cfg.

Par exemple, on dans le cas d'un ajout de ua-tac, dans la section dbase, il faudrait ajouter

config/main/monitoring/dbase/figures_6h/ua_tac.cfg

config/main/monitoring/dbase/figures_series/ua_tac.cfg

config/main/monitoring/dbase/figures_mensuelles/ua_tac.cfg

ainsi que l'endroit où pointe ces configs dans

/modules/figures_6h/container.cfg

/modules/figures_series/container.cfg

/modules/figures_mensuelles/container.cfg

On pourrait ensuite ajouter le nom dans la liste de sous-famille dans la variable

MONITORING_SOUS_FAMILLES de config/main/monitoring.cfg. En effet, les données seront extraite de la famille souche i.e. ua.

Pour la page web, on aimerait peut-être avoir cette nouvelle famille avec les RAOBS. Il faut alors ajouter dans config/main/monitoring/dbase/fichiers_web/combine.cfg une entrée qui contiendrait :

FICHIERS_WEB_COMBINE[ua]="ua ua_tac "

Production des figures

Évidemment, certaines figures ne sont pas appropriées pour une famille en particulier. Par exemple, il ne sera pas possible de produire un profil vertical pour la famille sf. On aura le message qu'il n'y a pas suffisamment de niveaux dans le listing. Ou bien, une figure de série O-P sur un dbase retournera une figure vide, sauf pour le nombre d'observations qui sera zéro. Si on désire traiter une liste de familles différente de celle des 3 variables MONITORING_*_FAM_LISTE pour certaines figures, il faut l'indiquer à plus bas niveaux, dans les variables : FIGURES_6h_FAM_* et FIGURES_series_FAM_* FIGURES_mensuelles_FAM_* qui se trouvent dans les figures_6h.cfg, figures_series.cfg et figures_mensuelles.cfg des postalt, evalalt et dbase.

Si on ne désire pas de figures pour un type de données complet, dans le fichier config/main/monitoring.cfg on peut régler plusieurs variables à « no ». Les variables suivantes:

MONITORING_RUN_postalt

MONITORING_RUN_evalalt

MONITORING_RUN_dbase

contrôlent toutes les figures des postalt, evalalt ou dbase.

Les variables suivantes:

MONITORING_POSTALT_RUN_FIGURES_6h

MONITORING_POSTALT_RUN_FIGURES_series

MONITORING_POSTALT_RUN_FIGURES_mensuelles

MONITORING_EVALALT_RUN_FIGURES_6h

MONITORING_EVALALT_RUN_FIGURES_series

MONITORING_EVALALT_RUN_FIGURES_mensuelles

MONITORING_DBASE_RUN_FIGURES_6h

MONITORING_DBASE_RUN_FIGURES_series

MONITORING_DBASE_RUN_FIGURES_mensuelles

contrôlent les modules figures_6h, figures_series et figures_mensuelles.cfg de chaque type de données.

Les figures produites seront sauvegardées dans les répertoires définis dans le fichier

config/main/monitoring.cfg

MONITORING_POSTALT_GRAPH_DESTI

MONITORING_EVALALT_GRAPH_DESTI

MONITORING_DBASE_GRAPH_DESTI

Production d'une page web

Les figures produites peuvent être visualisées grâce à une page web. Si l'on veut produire un site web, on peut activer cette option grâce à la variable MONITORING_RUN_fichiers_web.

Les variables

MONITORING_POSTALT_WEB_DESTI

MONITORING_EVALALT_WEB_DESTI

MONITORING_DBASE_WEB_DESTI

représentent l'endroit où les fichiers de type json et html seront copiés afin d'être manipulés pour en faire une page par type de données. La page web complète qui se trouvera à l'endroit défini par MONITORING_WEB_PAGE.

On peut combiner des familles ensembles sur la page web (ex to_amsua et to_amsub dans une section atovs). Il suffit, pour les 3 types de données de modifier le tableau dans
config/main/monitoring/postalt/fichiers_web/combine.cfg
config/main/monitoring/evalalt/fichiers_web/combine.cfg
config/main/monitoring/dbase/fichiers_web/combine.cfg

Exemple pour brancher le site de monitoring sur une suite

À partir d'une suite en mode « Live » :

Modifier la source des données converties dans la variable

branche=/home/erv000/data_maestro/ppp5/maestro_archives/gdps-370/monitoring/banco/postalt
pour les fichiers suivants:

config/main/monitoring/arc_obs_evalalt_6h.darc
config/main/monitoring/arc_obs_postalt_6h.darc
config/main/monitoring/arc_obs_dbase_6h.darc

Modifier les variables

POSTALT_EXP=/fs/homeu2/eccc/cmd/cmda/erv000/.suites/gdps-370
EVALALT_EXP=/fs/homeu2/eccc/cmd/cmda/erv000/.suites/gdps-370
DBASE_EXP=/fs/homeu2/eccc/cmd/cmda/erv000/.suites/gdps-370
du fichier resources/resources.def

Selon le type de suite sur laquelle on est branché, les dépendances dans les fichiers suivant peuvent avoir à être modifiées

resources/main/monitoring/get_arcdata_postalt.xml
resources/main/monitoring/get_arcdata_evalalt.xml
resources/main/monitoring/get_arcdata_dbase.xml

À partir d'une suite d'expérience terminée ou de bases de données archivées À partir d'une suite qui n'est pas en mode « Live » :

Modifier la source des données converties dans la variable

branche=/home/erv000/data_maestro/ppp5/maestro_archives/gdps-370/monitoring/banco/postalt
pour les fichiers suivants:

config/main/monitoring/arc_obs_evalalt_6h.darc
config/main/monitoring/arc_obs_postalt_6h.darc
config/main/monitoring/arc_obs_dbase_6h.darc
ou

config/main/monitoring/arc_obs_evalalt_6h.rarc
config/main/monitoring/arc_obs_postalt_6h.rarc
config/main/monitoring/arc_obs_dbase_6h.rarc

modifier config/main/monitoring.cfg pour les variables suivantes (au besoin):

MONITORING_GETDATA_postalt_DIR
MONITORING_GETDATA_evalalt_DIR

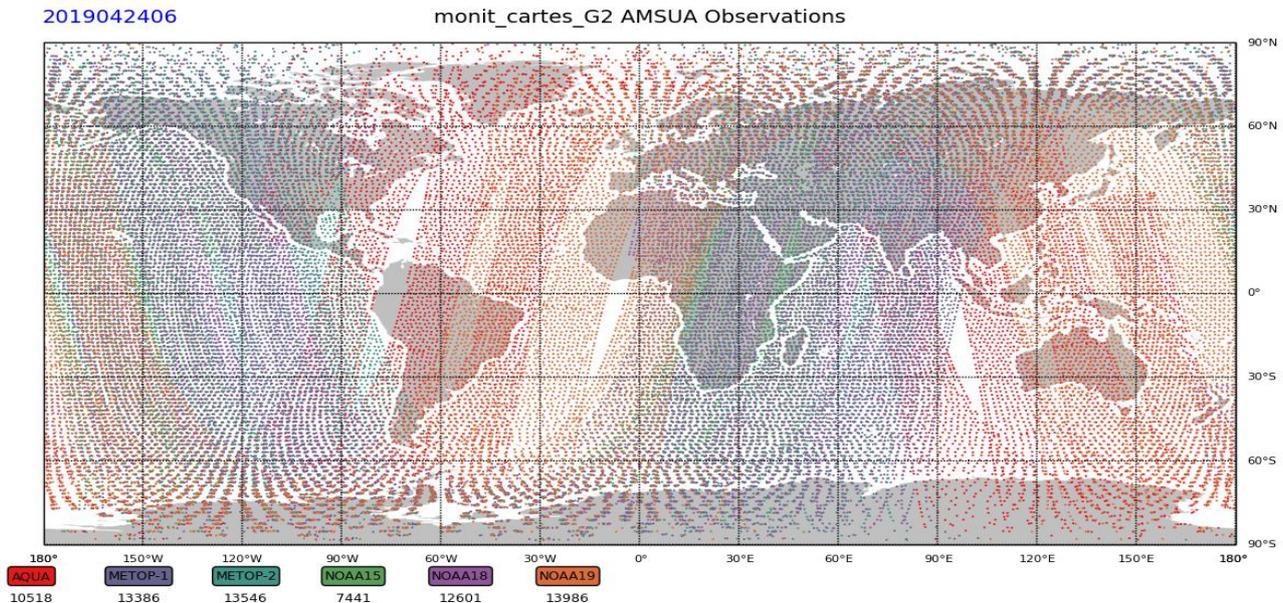
MONITORING_GETDATA_evalalt_DIR

Modifier le fichier config/main/cases

Les figures 6h (titres mauves de la page web) :

Ces figures utilisent les bases de données provenant de la conversion des fichiers burp à l'aide de burp2rdb.

1. **Scatter_maps** : Distribution géographique des observations. Figures produites sur une période de 6h et pour chaque bin temporelle. Correspond à « Distribution maps » des sections dbase evalalt et postalt ainsi qu'à « Distribution maps » sous l'onglet G2 En-Var Global Final Analysis du site web opérationnel. Les couches et/ou les bandes de latitude ne sont pas intégrés



présentement.

1.1. Comment régler les familles:

Variable FIGURES_6h_FAM_SQL2CSV de
config/main/monitoring/evalalt/figures_6h.cfg
config/main/monitoring/postalt/figures_6h.cfg
config/main/monitoring/dbase/figures_6h.cfg

1.2. Variables qui peuvent être changées par figures_6h/{fam}.cfg et utilisées pour produire le fichier csv (sql2csv.tsk).

PROJLIST cyl reg npstere spstere Canada
Tstep=15. (pour définir BIN). Temps en minutes
OFFSET=13 offset de temps pour les BINS (voir plus bas)
UA_SFC =condition sur les éléments de surface (ua seulement)
FLAG= Flags désirés
NO_ELEMS=" and varno not in(11001,11002,11012...)" Elements non-désirés.
BIN=cast(round((julianday(((julianday(isodatetime(date,time)))))) -
julianday(isodatetime({SEQ_SHORT_DATE}))) *24.*60./{Tstep},0) as int) +
{OFFSET}
LAT="lat"
LON="lon"
type=codtyp, id_stn ou rele
CODELIST=liste de cotype à vérifier

A_EXCLUDE= pour sql, condition supplémentaire.

1.3. Variable qui peuvent être changées par figures_6h/{fam}.cfg et utilisées pour contrôler la

figure (scatter_map.tsk).
PROJLIST cyl reg npstere spstere Canada
NBINS=25
FAMILLE_LABELS=\${fam}
Scatter_ptsize=Grosneur des points
Scatter_plotbins= «yes», «no» Pour faire tracer les figures pour chaque bins temporelles

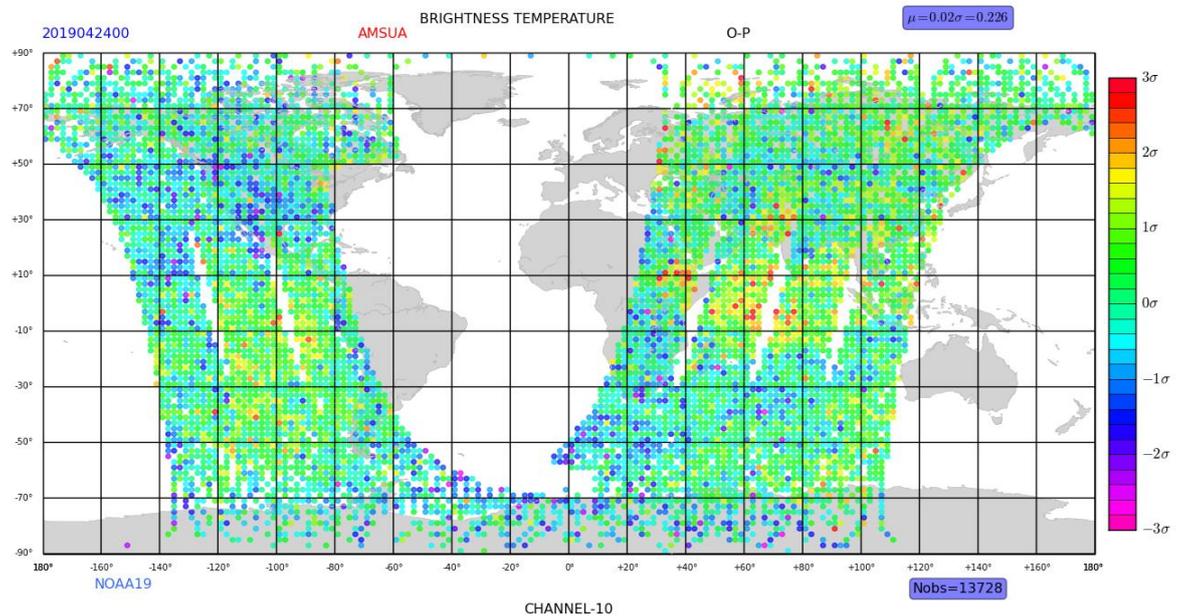
1.4. Shell script:

sql2csv.tsk
scatter_map.tsk
extension_projection
regions

Programmes python :

scatter_map_proj.py
constantes.py
LesProjections.py

2. `map_boites` : Distribution des O-P moyennée sur des boites. L'échelle correspond à des multiples de sigma, l'écart-type moyen. Figure produite par instrument ou satellite et coordonnées verticale ou canal. Correspond à " Maps of O-P values before/after thinning" des



sections evalalt et postalt.

2.1. Comment régler les familles

Variable `FIGURES_6h_FAM_SQL2CSVBOITES` de
`config/main/monitoring/evalalt/figures_6h.cfg`
`config/main/monitoring/postalt/figures_6h.cfg`

2.2. Variables qui peuvent être changées par `figures_6h/{fam}.cfg` et utilisées pour produire le fichier csv (`sql2csv_boites.tsk`).

`VCOORD_CRIT=" 1=1 "` Critère de coordonnées verticales

`CHANNELS=`liste de canaux ou variable vide

`FLAG="flag & 4096=4096 "` Données assimilées

`choffset=0` Nombre à soustraire du numero de canal (ex 27 pour amsua)

`LISTV=12163` élément burp à vérifier

`VCOORD=" vcoord - ${choffset} as vcoord "` Coordonnée verticale

`STN_IDS="id_stn like '%'"`

`FONCTION_boites="omp"` S'assurer qu'il s'agit d'une colonne dans les bases de données d'observations.

`REJ=" flag & 512=512 "`

`COUNTFONC=" distinct id_obs "` Methode pour compter le nombre d'obs

`Bias=0.0`

`type=id_stn, codtyp` ou rele

`SCAL=1`

`boxsizex=002` Longueur horizontale de la boite en degrés

`boxsizey=002` Longueur verticale de la boite en degrés

`mode=O-P, NOBS, OBSERVATIONS, O-P Std-dev, undefined`

`NMIN=1`

`A_EXCLURE=""`

`PROJLIST=cyl, npstere, spstere, Canada, reg`

`LATLONCRIT=""`

2.3. Variable qui peuvent être changées par figures_6h/{fam}.cfg et utilisées pour contrôler la figure (map_boites.tsk).

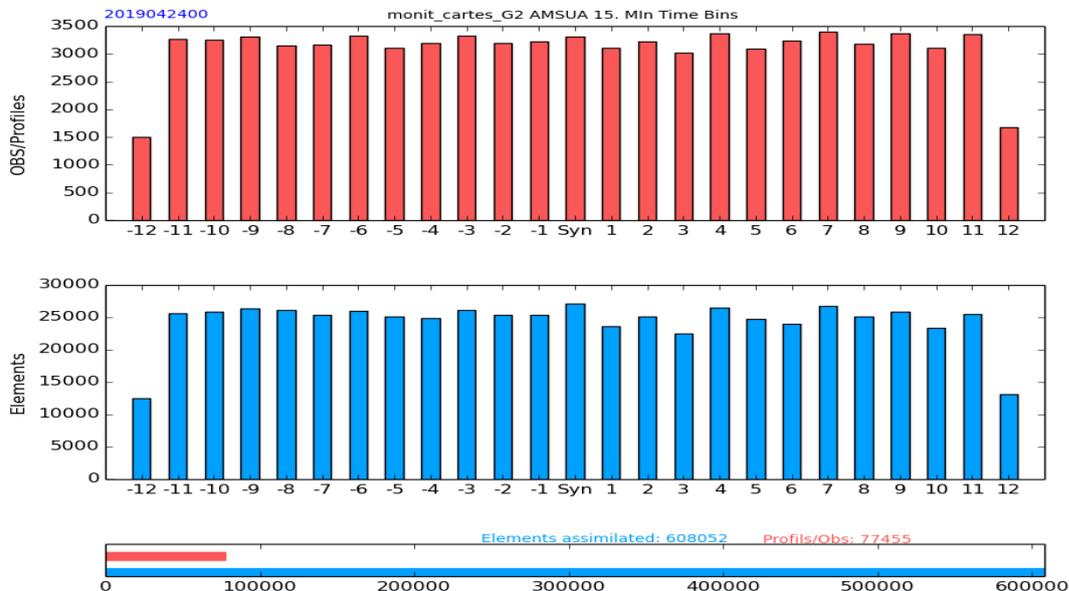
```
PROJLIST: cyl npstere, spstere, Canada, reg
VCOTYP: SURFACE, HEIGHT, CHANNEL, PRESSURE
VCOORD_LABEL=" "
CHANNELS
LISTV=12163
choffset=0
FAMILLE_LABELS=" "
Map_boites_ptsize=1.0
Mode=O-P, NOBS, OBSERVATIONS, O-P Std-dev, undefined
```

2.4. Shell scripts

```
sql2csv_boites.tsk
map_boites.tsk
extension_projection
regions
```

```
programmes python
maps.py
constantes.py
LesProjections.py
```

3. **Timebin** : Nombre d'observation par bin de temps.. Correspond a « No. Of observations per time bin » sous l'onglet « G2 En-Var Final Analysis». Mais elles peuvent être faites pour les evalalt et dbase aussi. Comme la tâche timebin.tsk partage la même tâche de préparation des données en fichier csv que scatter_plot, (sql2csv.tsk), les valeurs données aux variables dans les fichiers de configuration doivent correspondre. Par exemple, si on a fait des bins temporelles de 15 minutes pour scatter_plot, on aura un histogramme aux 15 minutes pour timebin.tsk.



3.1. Comment régler les familles

Variable FIGURES_6h_FAM_SQL2CSV de
 config/main/monitoring/evalalt/figures_6h.cfg
 config/main/monitoring/postalt/figures_6h.cfg
 config/main/monitoring/dbase/figures_6h.cfg

3.2. Variable qui peuvent être changées par figures_6h/{fam}.cfg et utilisées pour produire le fichier csv (sql2csv.tsk)

Même que pour scatter_plot

3.3. Variable qui peuvent être changées par timebin (mais doivent être réglées comme pour scatter_plot)

PROJLIST=cyl npstere, spstere, Canada, reg

NBINS=25

Tstep=15.

FAMILLE_LABELS=\${fam}

3.4. Shell scripts

sql2csv.tsk

timebin.tsk

extension_projection

regions

programmes python

timebins.py
constantes.py

FONCTION2=oma S'assurer qu'il s'agit de colonnes dans les bases de données
d'observations
VCOTYP: SURFACE, HEIGHT, PRESSURE, CHANNEL, RANGE
choffset=0
LISTV=12163

4.3. Variables qui peuvent être changées par par figures_6h/{fam}.cfg et utilisées pour
contrôler la figure (plot_prof.tsk)

FONCTION=omp
FONCTION2=oma
VCOTYP=SURFACE, HEIGHT, PRESSURE, CHANNEL, RANGE
FAMILLE_LABEL =
CHANNELS="ALL"
LISTV=12163
choffset=0
type=id_stn, codtyp, rele

4.4. Shell Scripts

sql2csv_prof.tsk
plot_prof.tsk

Programmes python
Profils_vert.py
Constantes.py

Figures series (titres rouges de la page web)

Ces figures utilisent les bases de données d'agrégation. Ces bases contiennent un sous-ensemble de données pour accélérer les requêtes sql. Les flags, les fonctions ainsi que les extra (si nécessaire) utilisés doivent être réglés lors de l'agrégation des fichiers burp.

Ces figures sont une accumulation des bases d'agrégation. Pour configurer le nombre de cas minimum et maximum que contiendront les « accum_agg », les variables suivantes doivent être définies dans config/main/monitoring.cfg :

MONITORING_NBCAS_MIN_SERIES="1" (ici, on a 2 dates. On compte le nombre de cas autre que \$SEQ_SHORT_DATE)

MONITORING_NBCAS_MAX_SERIES="100"

Les accum_agg créés par la tâche accum_agg.tsk sont sauvegardés dans

MONITORING_POSTALT_AGG_DESTI

MONITORING_EVALALT_AGG_DESTI

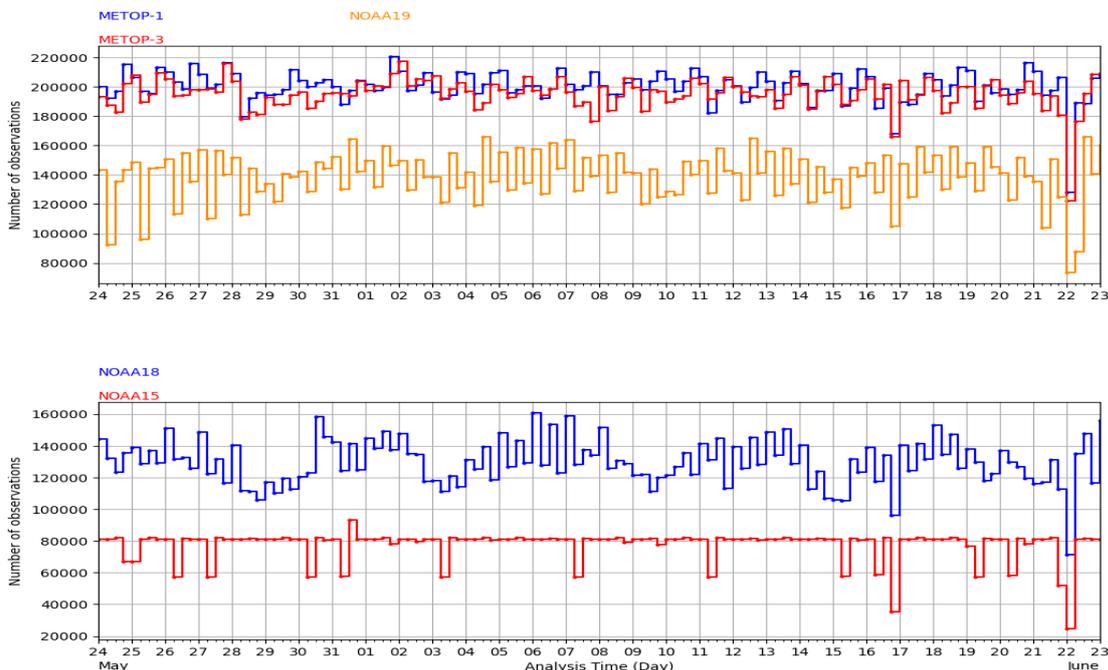
MONITORING_DBASE_AGG_DESTI

et seront utilisés par toutes les tâches de production de figures en série et certaines figures mensuelles.

1. **series_nobs** : Série du nombre d'observations. Correspond à « No. Of observations » pour evalalt et postalt et dbase.

2022052400 - 2022062300

TO_AMSUA_ALLSKY
90S 90N



1.1. Comment régler les familles

Variable FIGURES_series_FAM_ACCUM_AGG de
config/main/monitoring/dbase/figures_series.cfg
config/main/monitoring/evalalt/figures_series.cfg
config/main/monitoring/postalt/figures_series.cfg

- 1.2. Variables qui peuvent être changées par figures_series/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_nobs.tsk)

A_EXCLURE=""
AGG_CRIT=id_stn ou codtyp
NCOURBE_NOBS=1 : nombre de courbe sur la figure.

1.3. Variables qui peuvent être changées par par SQLITE_serie/\${fam}.cfg et utilisées pour contrôler la figure (series_nobs.tsk)
aucune

1.4. Shell scripts

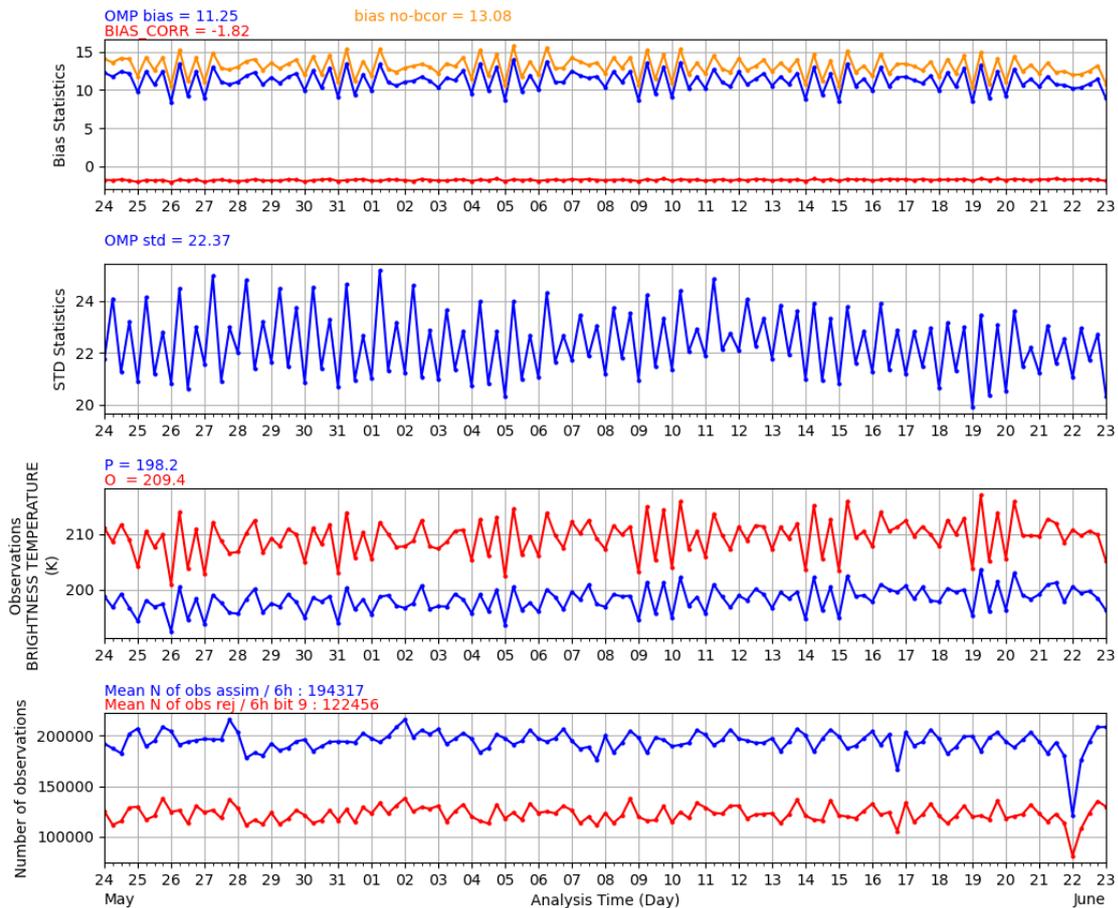
accum_agg.tsk
submit_nobs.tsk
accum_sql2csv_nobs.tsk
series_nobs.tsk

programmes python
constantes.py
parse.py*
plot.py*
plotTimeSeries.py

2. series temporelles : figures des biais, std O-P O-A (si applicable), valeurs des observations et nombre d'observation assimilées et/ou rejetées. Correspond à « O-P O-A statistics » dans la section postalt et « O-P of data before thinning » pour evalalt.

2022052400 - 2022062300

METOP-3 BRIGHTNESS TEMPERATURE (K) 28
90S 90N AMSUA



2.1. Comment régler les familles

Variable FIGURES_series_FAM_ACCUM_SQL2CSV_TEMPORELLE de
 config/main/monitoring/dbase/figures_series.cfg
 config/main/monitoring/evalalt/figures_series.cfg
 config/main/monitoring/postalt/figures_series.cfg

2.2. Variables qui peuvent être changées par figures_series/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_temporelle.tsk)

```
A_EXCLUDE=""
CHANNELS=""
VCOORD_CRIT=" 1=1 "
LISTV=12163
STN_IDS="id_stn like '%"
CODELIST="164"
BANDE_LAT="lat between -90 and 90"
LAT_LABEL="90S 90N"
```

2.3. Variables qui peuvent être changées par figures_series/\${fam}.cfg et utilisées pour contrôler la figure (series_temporelles.tsk)

VCOORD_CRIT=" 1=1 "

LISTV=12163

BANDE_LAT="lat between -90 and 90"

CODELIST="164"

CHANNELS=""

2.4. Shell scripts :

accum_agg.tsk

submit_temporelle.tsk

accum_sql2csv_temporelle.tsk

Series_temporelle.tsk

Programmes python

constantes.py

parse.py

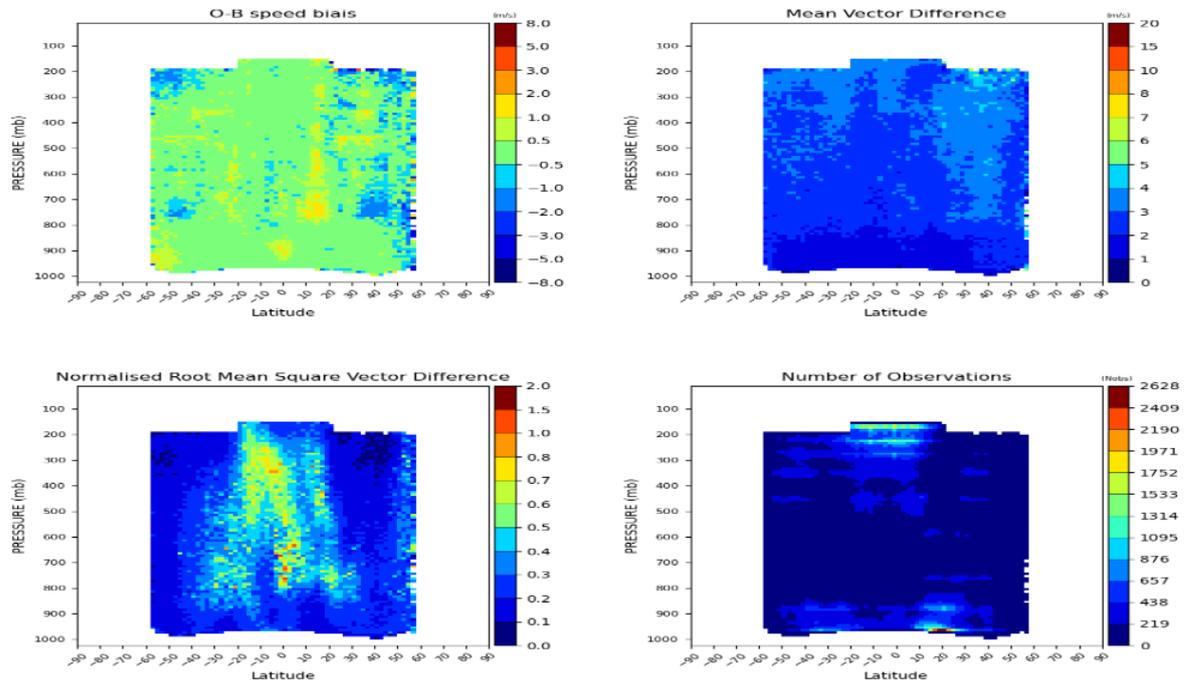
plot.py

plotTimeSeries.py

3. Zonaux Non-disponible pour dbase puisqu'il n'y a pas de OMP ou OMA.

2023020500 - 2023030700

HMWARI-9 WINDS (m/s) IR
AMVS



3.1. Comment régler les familles

Variable FIGURES_series_FAM_ACCUM_SQL2CSV_ZONAUX de
config/main/monitoring/dbase/figures_series.cfg
config/main/monitoring/evalalt/figures_series.cfg
config/main/monitoring/postalt/figures_series.cfg

3.2. Variables qui peuvent être changées par figures_series/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_zonaux.tsk)

```
A_EXCLUDE=""  
CHANNELS=""  
CRIT_ZONAUX=" 1=1 "  
STN_IDS="id_stn like %"  
CODELIST="164"  
VCOTYP=PRESSURE (mb), CHANNEL  
boxsizex=2.0 ( 2 deg de latitude)  
boxsizey=10 ( 10 mb).  
CRIT_ZONAUX_LABEL=""  
LISTV_ZONAUX=winds ou élément BURF. (Si « winds », s'assurer que FONCTION[3]  
existe dans les bases d'agrégation.)  
chffset=0
```

3.3. Variables qui peuvent être changées par par SQLITE_serie_{fam}_cfg et utilisées pour contrôler la figure (series_zonaux.tsk)

```
CRIT_ZONAUX_LABEL=""  
LISTV_ZONAUX=winds  
CODELIST="188"
```

3.4. Shell scripts

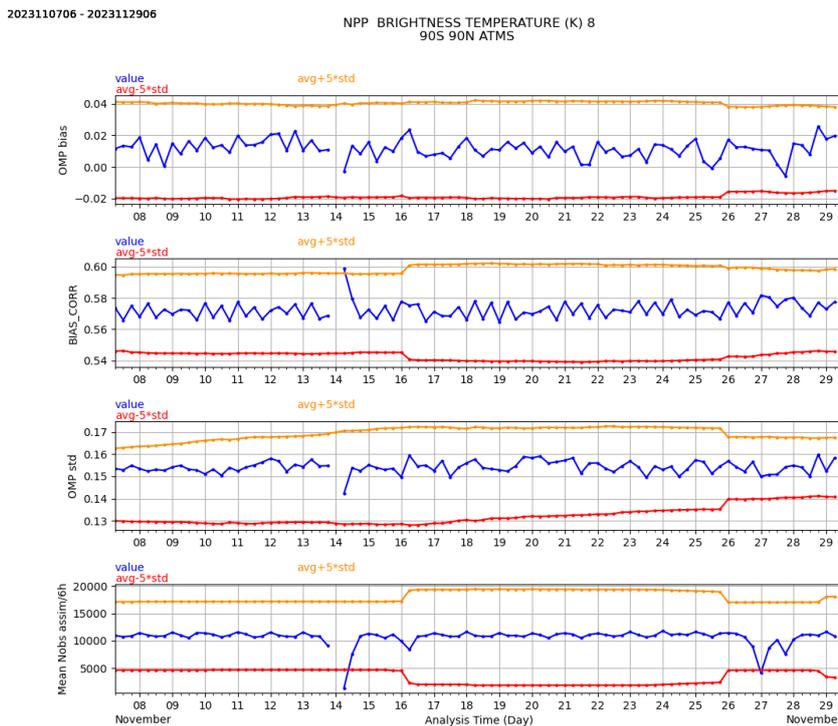
accum_agg.tsk
submit_zonaux.tsk
accum_sql2csv_zonaux.tsk
Series_zonaux.tsk

Programmes python :

plot_zonaux.py
parse_contour.py
parse.py
plotContourGraphs.py
constantes.py

4. **Series_anomalies** Figures pour montrer les anomalies dans la qualité ou la quantité des observations. Les limites sont calculées par la moyenne plus ou moins 5 fois l'écart-type sur une période allant de t-22 jours jusqu'à t-2 jours. Cette méthode provient de <https://www.ecmwf.int/sites/default/files/elibrary/2014/17338-automatic-checking-observations-ecmwf.pdf>

Pour produire la moyenne et l'écart-type, on demande un minimum de 10 données sur les 20 jours utilisés (80 cas) pour le calcul. Pour certaines familles, on ne vérifie pas les données manquantes (ai et ua). Pour les autres familles, lorsqu'une donnée est manquante ou dépasse les limites, un graphique et une page html sont créés. Le graphique est créé pour une période de 22 jours soit du début de l'influence des données sur le dernier point. Quant à la page web « Table of events », elle contient les valeurs des 35 derniers jours.



Canadian Meteorological Centre

List of ai incidents.
in file.
2023120500

DATE	ID_STN	CODTYP	VARNO	VCOORD	TYPE_STATS	MESSAGE
2023120406	AI	AIREP	TEMPERATURE/DRY BULB	701-1050hPa	OMP bias	Missing data low limit=-4.93 high limit=4.86
2023120406	AI	AIREP	TEMPERATURE/DRY BULB	701-1050hPa	BIAS_CORR	Missing data low limit=-0.28 high limit=-0.24
2023120406	AI	ADS	TEMPERATURE/DRY BULB	301-700hPa	BIAS_CORR	Slightly: outside +/-5 stdev from the mean val=-0.65 low limit=-1.01 high limit=-0.67
2023120400	AI	AIREP	TEMPERATURE/DRY BULB	701-1050hPa	OMP bias	Missing data low limit=-4.85 high limit=4.8
2023120400	AI	AIREP	TEMPERATURE/DRY BULB	701-1050hPa	BIAS_CORR	Missing data low limit=-0.28 high limit=-0.24
2023120400	AI	AIREP	TEMPERATURE/DRY BULB	701-1050hPa	Mean Nobs assim/6h	Missing data low limit=-2.77 high limit=4.38
2023120400	AI	AIREP	VECTORIAL DIFFERENCE OF WIND, VERIF	701-1050hPa	OMP MVD	Missing data low limit=-10.57 high limit=18.74
2023120400	AI	AIREP	VECTORIAL DIFFERENCE OF WIND, VERIF	701-1050hPa	OMP SPEED	Missing data low limit=-15.85 high limit=20.39
2023120400	AI	AIREP	VECTORIAL DIFFERENCE OF WIND, VERIF	701-1050hPa	Mean Nobs assim/6h	Missing data low limit=-3.05 high limit=6.39
2023120400	AI	AIREP	U-COMPONENT	701-1050hPa	OMP bias	Missing data low limit=-17.72 high limit=18.21
2023120400	AI	AIREP	U-COMPONENT	701-1050hPa	OMP std	Missing data low limit=-4.96 high limit=6.57

4.1. Comment régler les familles

Variable FIGURES_series_FAM_ANOMALIES de
config/main/monitoring/dbase/figures_series.cfg
config/main/monitoring/evalalt/figures_series.cfg
config/main/monitoring/postalt/figures_series.cfg

- 4.2. Variables qui peuvent être changées par figures_series/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_graph_anomalies.tsk). Attention, ce sont les mêmes variables que pour les figures_series_temporelles (sauf CHECK_MISSING) puisque les données proviennent de accum_sql2csv_nobs.tsk ainsi que accum_sql2csv_temporelle.tsk.

```
A_EXCLUDE=""  
CHANNELS=""  
VCOORD_CRIT=" 1=1 "  
LISTV=12163  
STN_IDS="id_stn like '%"  
CODELIST="164"  
BANDE_LAT="lat between -90 and 90"  
LAT_LABEL="90S 90N"  
CHECK_MISSING[STNID,Codtyp]="and message NOT LIKE '%Missing%' "
```

- 4.3. Variables qui peuvent être changées par par SQLITE_serie_{fam}_cfg et utilisées pour contrôler la figure (series_graph_anomalies.tsk)

```
VCOORD_CRIT=" 1=1 "  
LISTV=12163  
BANDE_LAT="lat between -90 and 90"  
CODELIST="164"  
CHANNELS=""
```

- 4.4. Shell scripts

```
accum_agg.tsk  
submit_series_anomalies.tsk  
accum_sql2csv_graph_anomalies.tsk  
Series_graph_anomalies.tsk  
Accum_sql2csv_html_anomalies.tsk  
Series_html_anomalies.tsk
```

Programmes python :

```
constantes.py  
parse.py  
plot.py  
plotTimeSeries.py  
Csv2HtmlTable.py
```

5. **nobspercent** Tableau représentant la variation du nombre de profils sur 24h et sur une semaine. Il s'agit d'une moyenne sur 6h des profils. Pour « 1 day », on aura 4 cas et pour « 1 week », on aura 28 cas. La valeur « ALL » de id_stn correspond à une famille complète. Quant à la valeur « ALL » de la colonne family, elle représente la somme totale de toutes les familles.

Canadian Meteorological Centre

```
Avg of nobs profiles per 6h
in postalt file.
today= date > 2024062418 and date <= 2024062518
yesterday=date > 2024062218 and date <= 2024062318
this week=date > 2024061818 and date <= 2024062518
last week=date > 2024061018 and date <= 2024061718

Difference 1 day = (Today - Yesterday) / Yesterday * 100
Difference 1 week = (This_week - Last_week) / Last_week * 100
```

FAMILY	ID_STN	CODTYP	NB PROFILS TODAY	NB PROFILS YESTERDAY	DIFFERENCE 1 DAY (%)	NB PROFILS THIS WEEK	NB PROFILS LAST WEEK	DIFFERENCE 1 WEEK (%)
AI	AI	AMDAR	2082.75	2159.75	-3.57	2142.89	2145.18	-0.11
AI	AI	AIREP	711.00	704.75	0.89	660.89	629.29	5.02
AI	AI	BUFR	30436.25	80567.00	-62.22	76237.25	85904.25	-11.25
AI	AI	ADS	4855.00	3818.50	27.14	3633.46	3460.89	4.99
AI	ALL	ALL	38085.00	87250.00	-56.35	82674.50	92139.61	-10.27
ATMS_ALLSKY	NOAA20	ALL	14486.75	15212.25	-4.77	15091.82	15181.44	-0.59
ATMS_ALLSKY	NPP	ALL	12623.50	14854.25	-15.02	14413.89	14156.00	1.82
ATMS_ALLSKY	ALL	ALL	27110.25	30066.50	-9.83	29505.71	6396.46	361.28
CRIS	NOAA20	ALL	11232.25	11501.75	-2.34	11305.21	11425.79	-1.06
CRIS	ALL	ALL	11232.25	11501.75	-2.34	11305.21	11425.79	-1.06
CSR	GOES16	ALL	7124.25	9093.00	-21.65	8397.04	9367.68	-10.36
CSR	GOES18	ALL	6772.25	7259.75	-6.72	7637.96	8508.46	-10.23
CSR	HMWARI-9	ALL	7995.75	8269.25	-3.31	8600.46	8817.96	-2.47
CSR	ALL	ALL	21892.25	24622.00	-11.09	24635.46	26694.11	-7.71
CS	CS	GROUND BASED	3262.25	3211.75	1.57	3216.66	3228.50	-0.37

5.1. Comment régler les familles

Variable FIGURES_series_FAM_ANOMALIES de
 config/main/monitoring/dbase/figures_series.cfg
 config/main/monitoring/evalalt/figures_series.cfg
 config/main/monitoring/postalt/figures_series.cfg

5.2. Variables qui peuvent être changées par figures_series/\${fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_html_nobspercent.tsk).

Les données proviennent de la table soft_limits. On ne fait que sommer le nombre de profils.

5.3. Variables qui peuvent être changées par par SQLITE_serie_\${fam}_cfg et utilisées pour contrôler la figure (nobspercent.tsk)

5.4. Shell scripts

Accum_csv2sql_nobs.tsk
 Accum_csv2sql_temporelles.tsk

Programmes python :
 Csv2HtmlTable.py

Séries mensuelles (titres jaunes de la page web)

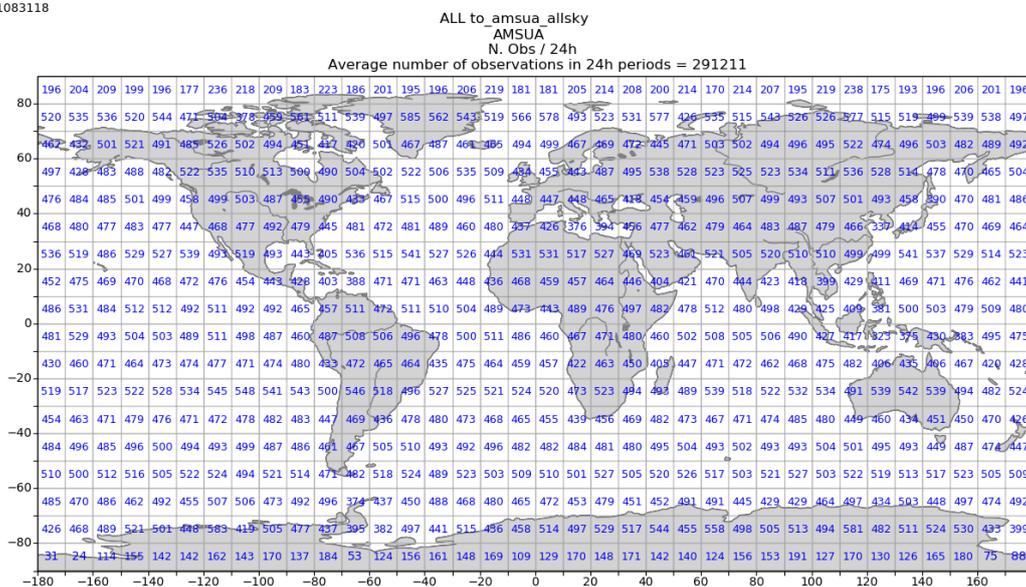
Ces figures utilisent soit les bases de données d'agrégation ou bien les bases d'agrégation mensuelles. Les bases d'agrégation mensuelles contiennent les moyennes sur 6h ou 24h pour le mois complet afin d'accélérer les requêtes sql.

1. Mensuelles_nprof_boites : Figures du nombre de profils par boites. Moyennes sur 24h.

Utilise la base de donnée d'agrégation.

Note concernant la config à utiliser : Étant donné qu'il a été démontré qu'il n'est pas possible, dans certaines conditions, de recalculer précisément le nombre de profils à partir de la table d'agrégation « avg », il est souhaitable d'utiliser la table « count_profils_assim » qui elle, est basée sur le nombre de id_obs distinct. Il est donc mieux, si cela est possible, de ne pas demander d'élément (LISTV) ou de niveaux (VCOORD_CRIT). Ceci est particulièrement vrai lorsque les données ont été regroupées par bandes de latitude ou de longitude dans la base d'agrégation (lors de la conversion) et contiennent beaucoup d'éléments et/ou de niveaux.

2021080100 - 2021083118



1.1. Comment régler les familles

Variable FIGURES_mensuelles_FAM_NPROF_BOITES de
config/main/monitoring/dbase/figures_mensuelles.cfg
config/main/monitoring/evalalt/figures_mensuelles.cfg
config/main/monitoring/postalt/figures_mensuelles.cfg

1.2. Variables qui peuvent être changées par figures_mensuelle/\${fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_nprof_boites.tsk)

AVG_AGG_CRIT=id_stn ou codtyp
CODELIST="164"
LISTV[0]=12163
A_EXCLUDE=""
VCOORD_CRIT=" 1=1 "
boxsizex=10.
boxsizey=10.
SQL_EXTRA_GROUPBY_NPROF_BOITES=""

1.3. Variables qui peuvent être changées par par figures_mensuelles/\${fam}.cfg et utilisées pour contrôler la figure (mensuelles_nprof_boites.tsk)

LISTV

VCOORD_CRIT=" 1=1 "

CODELIST="164"

1.4. Shell scripts

accum_agg.tsk

submit_nprof_boites.tsk

accum_sql2csv_nprof_boites.tsk

mensuelles_nprof_boites.tsk

Programmes python :

plot_nprof_boites.py

parse_contour.py

parse.py

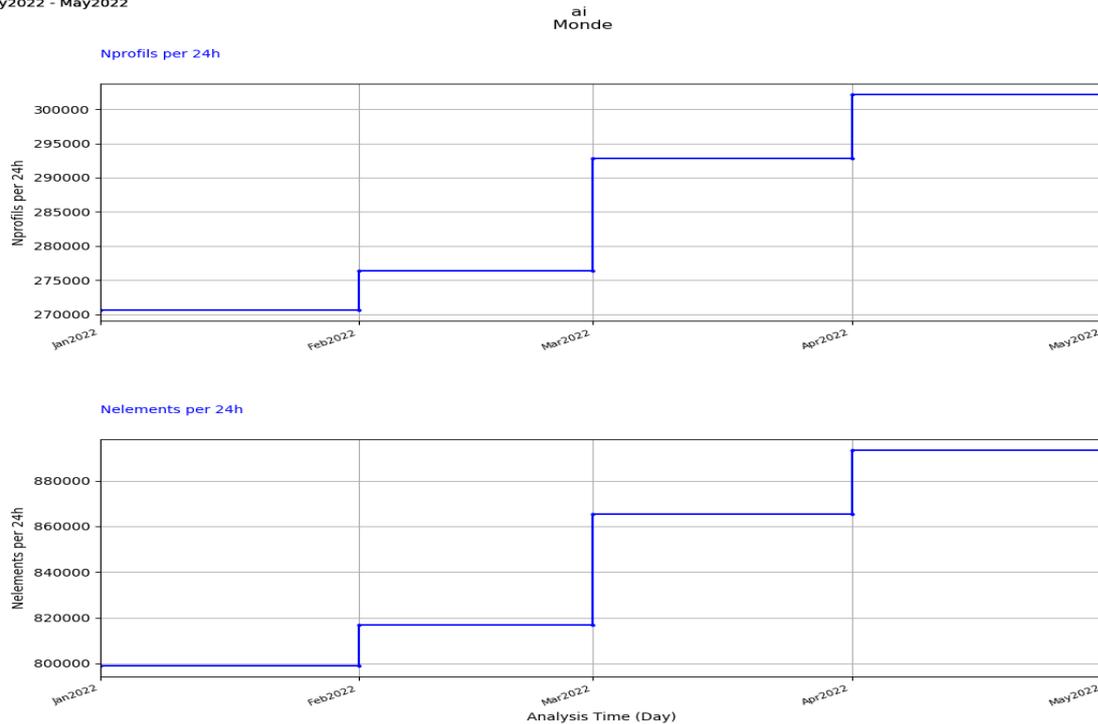
plotContourGraphs.py

constantes.py

LesProjections.py

2. **Traces** : Figures du nombre d'observations dans le mois. On y retrouve le nombre de profils ainsi que le nombre d'observations. Utilise la base d'agrégation mensuelle.

January2022 - May2022



2.1. Comment régler les familles

Variable FIGURES_mensuelles_FAM_TRACES_NOBS de
 config/main/monitoring/dbase/figures_mensuelles.cfg
 config/main/monitoring/evalalt/figures_mensuelles.cfg
 config/main/monitoring/postalt/figures_mensuelles.cfg

- 2.2. Variables qui peuvent être changées par figures_mensuelle/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_traces_nobs.tsk)

NA

- 2.3. Variables qui peuvent être changées par figures_mensuelles/{fam}.cfg et utilisées pour contrôler la figure (mensuelles_traces_nobs.tsk)

Idem à figures_series/series_nobs.tsk

2.4. Shell scripts

accum_mois.tsk
 submit_traces_nobs.tsk
 accum_sql2csv_traces_nobs.tsk
 mensuelles_traces_nobs.tsk (= figures_series/series_nobs.tsk)

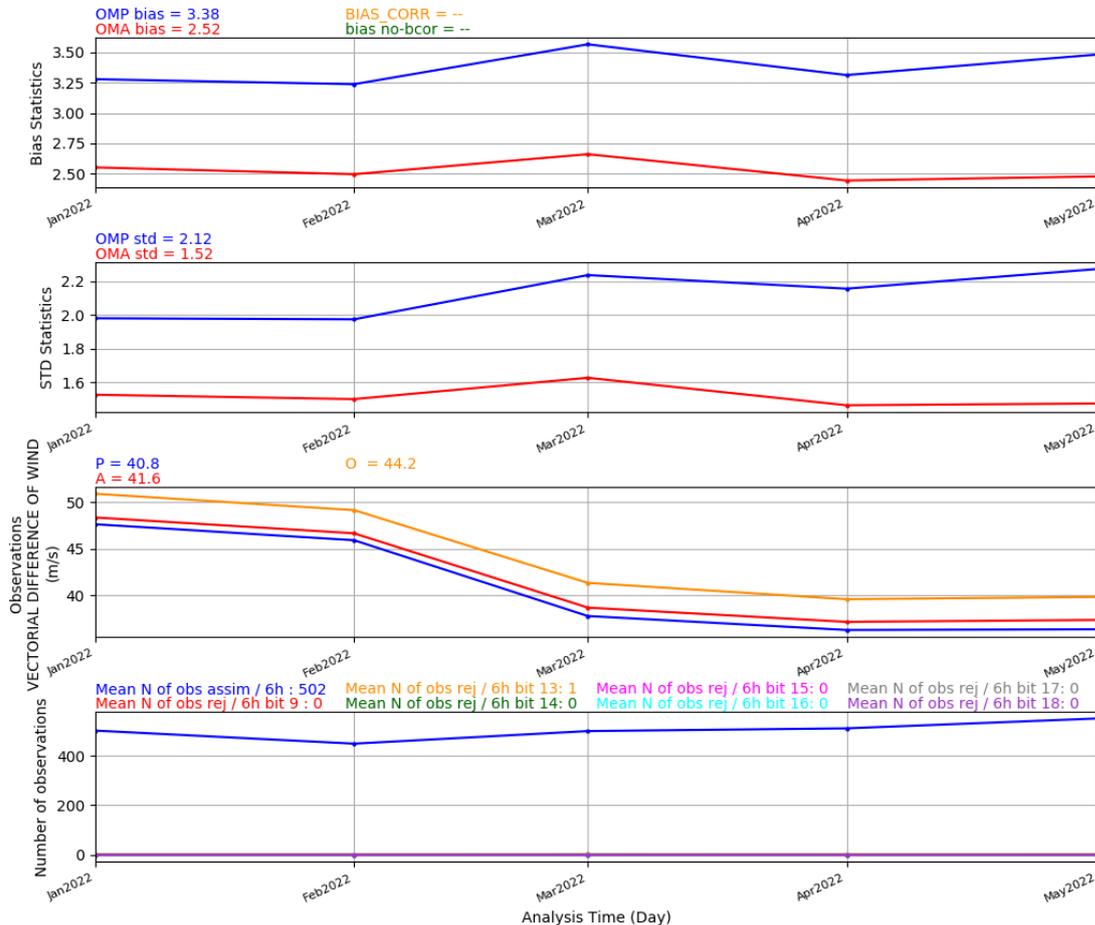
Programmes python :

Constantes.py
 parse.py*
 plot.py*
 plotTimeSeries.py

3. **Séries temporelles mensuelles O-P** : figures des biais, std O-P O-A (si applicable), valeurs des observations et nombre d'observation assimilées et/ou rejetées. Utilise la base d'agrégation mensuelle.

January2022 - May2022

AI VECTORIAL DIFFERENCE OF WIND (m/s) 100-300hPa
90S 90N AMDAR



3.1. Comment régler les familles

Variable FIGURES_mensuelles_FAM_SERIES_TEMPORELLES de
config/main/monitoring/dbase/figures_mensuelles.cfg
config/main/monitoring/evalalt/figures_mensuelles.cfg
config/main/monitoring/postalt/figures_mensuelles.cfg

3.2. Variables qui peuvent être changées par figures_mensuelle/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_series_temporelles.tsk)

A_EXCLUDE=""
CHANNELS=""
VCOORD_CRIT=" 1=1 "
LISTV=12163
STN_IDS="id_stn like "%"
CODELIST="164"
BANDE_LAT="lat between -90 and 90"
LAT_LABEL="90S 90N"

3.3. Variables qui peuvent être changées par par figures_mensuelles/{fam}.cfg et utilisées pour

contrôler la figure (mensuelles_series_temporelles.tsk)
Idem à figures_series/series_temporelles.tsk

3.4. Shell scripts

accum_mois.tsk
submit_series_temporelles.tsk
accum_sql2csv_series_temporelles.tsk
mensuelles_series_temporelles.tsk (= figures_series/series_temporelles.tsk)

Programmes python :

constantes.py
parse.py
plot.py
plotTimeSeries.py

5. Tables html du dernier rapport des stations (extrait) Utilise la base de donnée d'agrégation.

Canadian Meteorological Centre

List of RADIOSONDES stations.
Gives the date of the last observation in postalt file.
May 2022

ID_STN	CODTYP	LAT	LON	DATE
01001	35	70.94	-8.67	2022-05-31 23:11
01004	35	78.92	11.92	2022-05-31 10:58
01010	35	69.31	16.13	2022-05-31 23:02
01028	35	74.50	19.00	2022-05-31 23:10
01241	35	63.70	9.61	2022-05-31 23:02
01400	35	56.54	3.22	2022-05-31 23:02
01415	35	58.87	5.66	2022-05-31 23:03
01492	35	59.94	10.71	2022-05-31 10:42
02365	35	62.52	17.45	2022-05-27 12:00
02527	35	57.67	12.50	2022-05-31 23:30
02836	35	67.37	26.63	2022-06-01 00:00
02963	35	60.82	23.50	2022-06-01 00:00
03005	35	60.13	-1.19	2022-05-31 11:15
03238	35	55.01	-1.88	2022-05-31 23:15
03354	35	53.00	-1.25	2022-05-31 23:00
03502	35	52.13	-4.58	2022-05-26 11:22

5.1. Comment régler les familles

Variable FIGURES_mensuelles_FAM_LAST_STN de
config/main/monitoring/dbase/figures_mensuelles.cfg
config/main/monitoring/evalalt/figures_mensuelles.cfg
config/main/monitoring/postalt/figures_mensuelles.cfg

5.2. Variables qui peuvent être changées par figures_mensuelle/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_last_stn.tsk)

FAMILLE_LABEL=\${fam}

5.3. Variables qui peuvent être changées par figures_mensuelles/{fam}.cfg et utilisées pour contrôler la figure (mensuelles_last_stn.tsk)

5.4. Shell scripts

accum_agg.tsk
submit_last_stn.tsk
accum_sql2csv_last_stn.tsk
mensuelles_last_stn.tsk

Programmes python :
Csv2HtmlTable.py

Statistiques combinées (titres verts de la page web)

Ces figures utilisent les données des 3 types de données dbase, evalalt et postalt pour les calculs. Ce type de statistique n'existe que dans le postalt. Utilisez la base de données d'agrégation.

1. Stations suspectes

Liste des stations suspectes selon les règles de la WMO définies dans le document [Manual on the Global Data processing and Forecasting System](#)

Canadian Meteorological Centre

List of AIRCRAFT stations with suspect observations.
February 2023
WIND codtyp=128
100-300hPa

To be considered as suspect, the number of observations must meet minimal counts (20) and the data statistics versus the guess must exceed at least one criterion:
abs(bias) >= 2.5 or rms >= 10. or 100.* (NG*1.0)/NA >= 2.

NA = Total number of available observations
NT = Number after data thinning for assimilation
NG = Observations differing from the guess by amounts larger than gross check limits (40.)
% Gross = NG/NA * 100
NR = The number of rejected observations excluding thinning
% NR = NR/NA * 100
NC = The number of exactly calm winds (obs < 50.) m/s
Wind speed (elem BUFR 11002) used in calculation

STN	ELEM	LEVEL	NA	NT	NG	% GROSS	NR	% NR	NC	RMS	BIAS
ANZ6071	WIND	100-300hPa	37	17	0	0	15	40.5	37	4.77	3.00
TWY10	WIND	100-300hPa	30	11	0	0	0	0.0	30	4.41	2.95
UAL1721	WIND	100-300hPa	32	24	0	0	2	6.3	27	6.00	2.64

1.1. Comment régler les familles

Variable STATS_combinees_FAM_STATIONS_SUSPECTES de config/main/monitoring/postalt/stats_combinees.cfg

1.2. Variables qui peuvent être changées par stats_combinees/{fam}.cfg et utilisées pour produire le fichier csv (accum_sql2csv_stations_suspectes.tsk)

type=codtyp

CODELIST="42 157 177"

FAMILLE_LABELS="AIRCRAFT"

AJOUTE_LATLON="non"

CHECK_HEURE="oui"

COMBINE_VCOORD="oui"

agg_in_postalt=\${MONITORING_POSTALT_AGG_DESTI}

agg_in_evalalt=\${MONITORING_EVALALT_AGG_DESTI}

agg_in_dbase=\${MONITORING_DBASE_AGG_DESTI}

LISTV[1]=12001

LISTV_LABEL[1]="TEMP"

```
VCOORD[0]=" '100300' as vcoord "  
VCOORD_CRIT[0]=" vcoord > 10000. and vcoord <= 30000. "  
VCOORD_LABEL[0]="100-300hPa"  
#NMIN[vcoord] (nombre min d'obs necessaire pour faire le calcul)  
#GROSS[vcoord,listv]  
#BIAS_LIMIT[vcoord,listv]  
#RMS_LIMIT[vcoord,listv]  
#PGRERR[listv] (gross error check)  
#CALM_LIMIT[codtyp] (limite de vents calmes)  
#WEIGHT [vcoord,listv] (poids à donner à chaque niveaux)  
#listeniv[listv] (liste de niveaux à considerer pour chaque élément)  
#extra_sql[numero,listv]
```

1.3. Variables qui peuvent être changées par par stats_combinees/{fam}.cfg et utilisées pour contrôler la figure (stations_suspectes.tsk)

1.4. Shell scripts

```
accum_agg.tsk des 3 types de données (dbase, evalalt et postalt)  
submit_stations_suspectes.tsk  
accum_sql2csv_stations_suspectes.tsk  
stations_suspectes.tsk
```

Programmes python :
Csv2HtmlTable.py

Bases de données d'agrégation

Les bases de données d'agrégation sont un pré-calcul qui sera utilisé dans les programmes de visualisation. On y garde des sommes et sommes qui serviront, par exemple, au calcul des moyennes et écarts-types. Pour chacune des familles, un fichier de configuration est nécessaire afin de spécifier les paramètres du calcul. On peut y définir, entre autre, la région (ex. Monde, Amérique du Nord, etc.), les flags pris en compte pour les rejets ou l'acceptation des données, les éléments utilisés dans les calculs, ainsi que les équations mathématiques utilisées pour produire les statistiques. Certaines variables des fichiers de configuration sont représentées sous forme de tableau. L'indice du tableau correspondra au numéro des colonnes dans les tables. Les tables seront automatiquement créées selon le nombre d'éléments des tableaux avec un minimum qui est définie dans la tâche d'agrégation. Les bases de données d'agrégation contiendront 4 tables. Voici un résumé? des différentes tables produites lors de l'agrégation.

1. La table avg

Avg contient les éléments suivants : date, id_stn, codtyp, varno, lat, vcoord, sumobs, sum_fonc0, sum_fonc0carre, Nrej0, Extra0 Nassim.

Dans la table avg les sum_fonc0 et sum_fonc0carre sont défini par la variable FONCTION qui est un tableau dans les fichiers de configuration. On peut mettre autant de FONCTION que l'on désire. Toutefois, le minimum est réglé à 3 éléments. Si FONCTION contient moins de 3 éléments, des valeurs NULL seront inscrites dans les colonnes. Les FONCTION sont groupées par id_stn, codtyp, varno, lat et vcoord et Extra0. Le ou les varno sélectionnés sont définis par la variable LISTV dans le fichier de configuration.

Note sur les FONCTION. En principe, on peut changer les FONCTIONS dans les fichiers de configuration. Toutefois, il faut porter une attention particulière à ce choix et à l'interaction qu'on veut créer entre les fonctions. Par exemple, si on veut faire un calcul à partir de OMP et de BIAS_CORR (trouver la valeur non corrigée), il faut s'assurer que les sommes de OMP et BIAS_CORR sont effectuées sur les mêmes éléments. Pour pallier à ce problème, lors de la création de la base de données d'agrégation, on vérifie ASSIM lors de l'insertion des valeurs dans toutes les FONCTIONS. Autre exemple, pour un fichier de type evalalt, il n'y a pas de OMA calculés. Les valeurs sont donc toutes manquantes dans le fichier burp. Si on met OMA dans l'espace FONCTION2, la production d'images de séries temporelles sera affectée puisque toutes les valeurs seront NULLES.

sumobs est la somme des obsvalues.

lat et lon sont respectivement la latitude et la longitude. Elles sont réglées dans les fichiers de configuration et peuvent être arrondies (par exemple à la dizaine près) par

```
LAT="(round(lat/10.)*10)"  
LON="(round(lon/10.)*10)"
```

Nassim représente le nombre de profils où les OMP, OMA ne sont pas 'NULL' et où le flag défini par ASSIM est allumé .dans le fichier de configuration

Nrej0 représente les éléments rejetés. Il y a 7 espaces pour la sélection des bits de rejets au minimum. Ces bits sont définis dans le fichier de configuration. Par défaut, ils sont réglés à 'NULL'. Il n'est donc pas besoin de tous les utiliser.

Extra0 représente une sélection autre qui peut être utilisées dans la production de figure. Pour l'instant, seuls les sw utilisent un Extra qui contient le WIND_COMP_METHOD nécessaire au calcul des figures zonales. Extra est sous forme de tableau, on peut donc en ajouter d'autres. Par contre, il faudra se rappeler qu'Extra entre dans la liste pour grouper les fonctions.

2. La table count_profils_assim

Pour la famille ua, la table count_profils_assim contient : date, codtyp, N_data_ua, N_data_sf.

Pour toutes les autres familles, la table count_profils_assim contient : date, id_stn, codtyp, lat, Nobs.

La table count_profils_assim des ua contient le nombre de données assimilées tel que défini par la variable ASSIM du fichier de configuration. N_data_ua est le décompte pour un btyp=9312 alors que N_data_sf a un btyp=98.

La table count_profiles_assim des autres familles contient le nombre de données assimilées tel que défini par la variable ASSIM du fichier de configuration. C'est Nobs qui est le décompte des données assimilées.

3. La table resume

Resume contient les colonnes suivantes : région, famille, fonction0, exper, date, Drej0, Extra0. Elle ne contient qu'une seule entrée, pour toutes les familles.

Par exemple, pour la famille sw, nous aurons les valeurs suivantes :

Region	Monde
Famille	Sw
Exper	Burp2Sql_conversion
Date	2020041512
Fonction0	OMP
Fonction1	OMA
Fonction2	OMP/OMA
Fonction3	(OBSVALUE-OMP)
Drej0	bit 9
Drej1	bit 13
Drej2	bit 14
Drej3	bit 15
Drej4	bit 16
Drej5	bit 17
Drej6	bit 18
Dextra0	WIND_COMP_METHOD

Par défaut, les descriptions de rejet sont mises à 'NULL'. Ils sont définis par le tableau descript_rejet du fichier de configuration de chaque famille. Ils seront utilisés lors de la production d'images.

4. La table last_report_stn

La table last_report_stn contient la date du dernier rapport d'une station. Elle contient : id_stn, codtyp, lat, lon, date. À noter, la date est inscrite en format *ISO standard date and time*.

5. Les tables `soft_limits`

Afin de pouvoir calculer les anomalies dans la quantité et la qualité des données, on garde une table contenant le calcul des moyennes et écart type ainsi que le message associé. Ces valeurs proviennent des fichiers csv produits pour faire les graphiques des séries temporelles ou séries nobles.

Date, id_stn, codtyp, varno, vcoord, type_stats, value, avg_limits, std_limits, message

La moyenne et l'écart-type sont calculés en utilisant les valeurs du « type_stats » allant de t-22 jours jusqu'à t-2 jours. On demande un minimum de 10 données sur cette période de 20 jours (80 cas) pour faire le calcul. Type_stats défini d'où viennent les valeurs soit des séries temporelles ou séries nobles.

Les bases d'agrégations mensuelles

Ces bases de données sont produites le 1^{er} de chaque mois à 00Z et portent sur le mois précédent.

1. La table `moyenne_avg`

Annee, mois, id_stn, codtyp, varno, vcoord, sumobs_6h, sum_fonc0_6h, sum_fonc0carre_6h, Nrej0_6h, Nassim_6h, lat_label

Les colonnes sont définies comme pour les accum_agg produites à chaque 6h. La différence est que l'on conserve la moyenne par 6h pour le mois.

Lat_label correspond à la bande de latitude considérée en format texte.

2. La table `moyenne_nprofs`

annee, mois, Nprofils_24h, Nelem_24h

Nous conservons ici que le nombre de profils ainsi que le nombre d'élément par 24h durant le mois.

3. La table `resume`

Resume contient les colonnes suivantes : annee, mois, région, famille, fonction0, exper, date, Drej0, Extra0. Elle ne contient qu'une seule entrée, pour toutes les familles.

Par exemple, pour la famille sw, nous aurons les valeurs suivantes :

Annee	2022
Mois	6
Region	Monde
Famille	Sw
Exper	MonitoringLive
Fonction0	OMP
Fonction1	OMA
Fonction2	OMP/OMA
Fonction3	(OBSVALUE-OMP)
Drej0	bit 9
Drej1	bit 13
Drej2	bit 14
Drej3	bit 15
Drej4	bit 16
Drej5	bit 17

Drej6	bit 18
Dextra0	WIND COMP METHOD

Par défaut, les descriptions de rejet sont mises à 'NULL'. Ils sont définis par le tableau `descript_rejet` du fichier de configuration de chaque famille. Ils seront utilisés lors de la production d'images.