

# Format des données SAFRAN et scénarios climatiques désagrégés au CERFACS

Christian Pagé

Climate Modelling and Global Change  
TR/CMGC/08/27

CERFACS  
Centre Européen de Recherche  
et de Formation Avancée en Calcul Scientifique  
42, avenue G. Coriolis, 31057 Toulouse Cedex, France

9 avril 2008

## Résumé

Les données SAFRAN sont des données horaires couvrant la France à une résolution de 8 km sur une projection Lambert-II étendue. Elles sont produites par Météo-France (Centre National de Recherches Météorologiques, CNRM). Une description du système SAFRAN appliqué à la France entière est décrite dans Le Moigne (2002).

Ces données, au CERFACS, sont utilisées dans le cadre de désagrégation de simulations climatiques afin de produire des données climatiques haute-résolution couvrant la France sur la grille SAFRAN. Historiquement, au CERFACS, les données SAFRAN sont utilisées sous un format binaire ou en format de sauvegarde IDL. De la même façon, les données des scénarios climatiques désagrégés étaient aussi dans ce format binaire. Pour des besoins pratiques, ces données ont été convertis dans le format NetCDF. Ce document décrit la convention NetCDF utilisée et la structure de stockage des données.

# 1 Structure des données et projection

Les données dans le format binaire SAFRAN ne sont pas enregistrées dans une matrice en deux dimensions malgré le fait que les données soient sur une grille régulière en projection Lambert. En effet, les données sont enregistrées comme une suite de points dans un vecteur une dimension. Il a donc été nécessaire de repenser la façon de structurer ces données afin de les convertir au format NetCDF, car ceci aurait limité significativement l'utilisation d'outils externes de visualisation et de traitement des fichiers NetCDF. Il a donc été décidé de replacer ces points sur une grille deux dimensions dans la projection Lambert native des données, malgré le fait que ceci augmente la taille des fichiers. Les points supplémentaires sont donc marqués comme ayant une valeur manquante.

La projection native des données est une projection conique Lambert-II étendue. Elle a donc les caractéristiques suivantes :

- Projection conique conforme
- Lambert-II étendue (Lambert-II Carto Zone Centre)
- Système NTF
- Parallèle 1 = 45°53'56.108" N
- Parallèle 2 = 47°41'45.652" N
- Latitude d'origine = 52 gr N (46°48' N)
- Longitude d'origine = 2°20'14.025" E
- X0 (False Easting) = 600 000
- Y0 (False Northing) = 2 200 000

La grille SAFRAN couvre donc la région suivante dans la projection Lambert-II étendue :

- $\Delta x = 8000.0$  m
- $\Delta y = 8000.0$  m
- Coordonnée X minimale = 60000 m
- Coordonnée X maximale = 1196000 m
- Coordonnée Y minimale = 1617000 m
- Coordonnée Y maximale = 2681000 m
- Nombre de points en X = 134
- Nombre de points en Y = 143
- Nombre de points total = 19162
- Nombre de points effectifs (valeurs manquantes omises) = 9892

À l'aide d'un fichier décrivant la position de chacun des points de données (parmi les 9892 points effectifs), chaque point a été placé au bon endroit dans la matrice deux dimensions (voir figure 1.1)

La figure

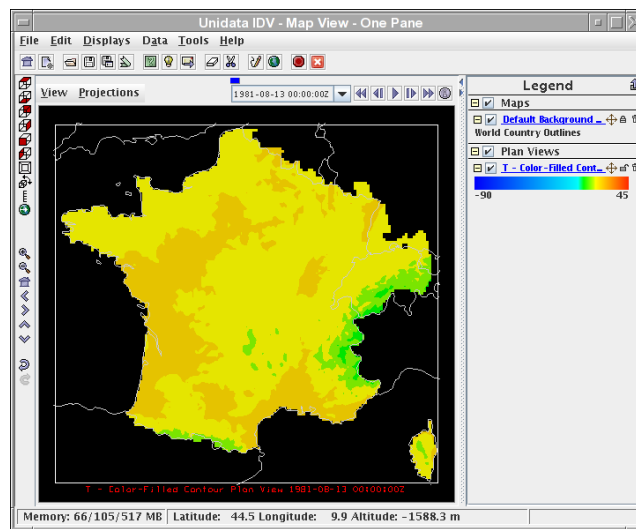


FIG. 1.1 – Couverture géographique des données SAFRAN et des scénarios climatiques désagrégés. Exemple pour la variable de température SAFRAN en moyenne journalière valide le 13/08/1981.

## 2 Convention NetCDF

La convention NetCDF qui a été utilisée est la convention CF-1.0 (Program for Climate Model Diagnosis and Intercomparison, 2008), incluant le "Data Discovery" (Ethan Davis, 2005). La convention CF-1.0 est une des conventions la plus récente et assure la compatibilité des fichiers NetCDF avec les outils de traitement et de visualisation tels que NetCDF Tools (Java) (Unidata, 2008), IDV (Unidata, 2007), etc. La convention "Data Discovery" permet de s'assurer que les fichiers peuvent être catalogués et documentés automatiquement s'ils sont mis à disposition par l'intermédiaire d'un serveur THREDDS.

## 3 Attributs globaux : méta-données

Les attributs globaux des fichiers SAFRAN et des scénarios climatiques désagrégés sont présentés ci-dessous. Les attributs sont auto-descriptifs, et des valeurs sont données en exemple.

- SAFRAN
 

```
// global attributes:
:Conventions = "CF-1.0" ;
:Metadata_Conventions = "Unidata Dataset Discovery v1.0" ;
:title = "Meteo-France SAFRAN data" ;
```

```

:title_french = "Donnees SAFRAN Meteo-France" ;
:summary_french = "Donnees SAFRAN Meteo-France" ;
:summary = "Meteo-France SAFRAN data" ;
:keywords = "climat,SAFRAN,Meteo-France,CERFACS" ;
:processor = "IDL scripts" ;
:description = "Meteo-France SAFRAN data" ;
:cdm_datatype = "Grid" ;
:institution = "CERFACS" ;
:creator_email = "globc@cerfacs.fr" ;
:creator_url = "http://www.cerfacs.fr/globc/" ;
:creator_name = "Global Change Team" ;
:time_coverage_start = "1980-08-01T00:00:00Z" ;
:time_coverage_end = "1981-07-31T23:59:59Z" ;
:timestep = "hourly" ;
:contact_email = "christian.page@cerfacs.fr" ;
:contact_name = "Christian PAGE" ;
:other_contact_email = "laurent.terray@cerfacs.fr" ;
:other_contact_name = "Laurent TERRAY" ;
:geospatial_lat_max = 51.1215157941104 ;
:geospatial_lat_min = 41.3181733938225 ;
:geospatial_lon_max = 10.7928123474121 ;
:geospatial_lon_min = -5.32915830612183 ;

```

- **Scénarios climatiques désagrégés**

```

// global attributes:
:Conventions = "CF-1.0" ;
:Metadata_Conventions = "Unidata Dataset Discovery v1.0" ;
:title = "Downscaling data from CERFACS" ;
:title_french = "Donnees de desagregation produites par le CERFACS" ;
:summary_french = "Donnees de desagregation produites par le CERFACS" ;
:summary = "Donnees de desagregation produites par le CERFACS" ;
:keywords = "climat,scenarios,desagregation,downscaling,CERFACS" ;
:processor = "IDL/NCL scripts" ;
:description = "Downscaling data from CERFACS" ;
:cdm_datatype = "Grid" ;
:institution = "CERFACS" ;
:creator_email = "globc@cerfacs.fr" ;
:creator_url = "http://www.cerfacs.fr/globc/" ;
:creator_name = "Global Change Team" ;
:time_coverage_start = "1991-08-01T00:00:00Z" ;
:time_coverage_end = "1992-07-31T23:59:59Z" ;
:version = "1.0" ;
:scenario = "SRESA1B" ;
:scenario_co2 = "A1B" ;
:model = "ARPEGE grille etiree" ;
:institution_model = "Meteo-France CNRM/GMGEC" ;
:country = "France" ;
:member = "1" ;
:downscaling_forcing = "SAFRAN 1970-2005" ;
:timestep = "daily" ;
:contact_email = "christian.page@cerfacs.fr" ;
:contact_name = "Christian PAGE" ;
:other_contact_email = "laurent.terray@cerfacs.fr" ;
:other_contact_name = "Laurent TERRAY" ;
:geospatial_lat_max = 51.1215157941104 ;
:geospatial_lat_min = 41.3181733938225 ;
:geospatial_lon_max = 10.7928123474121 ;
:geospatial_lon_min = -5.32915830612183 ;

```

## 4 Dimensions et coordonnées NetCDF

Les dimensions sont décrites en format illimité pour la dimension temps :

- time = UNLIMITED
- x = 143
- y = 134

Les variables de coordonnées sont décrites de la façon suivante, selon la convention CF-1.0. La variable Altitude est aussi renseignée dans chaque fichier (utilisée notamment par le logiciel IDV).

```
variables:
int Lambert_Conformal ;
    Lambert_Conformal:grid_mapping_name = "lambert_conformal_conic" ;
    Lambert_Conformal:standard_parallel = 45.89892f, 47.69601f ;
    Lambert_Conformal:longitude_of_central_meridian = 2.337229f ;
    Lambert_Conformal:latitude_of_projection_origin = 46.8f ;
    Lambert_Conformal:false_easting = 600000.f ;
    Lambert_Conformal:false_northing = 2200000.f ;
int x(x) ;
    x:units = "m" ;
    x:long_name = "x coordinate of projection" ;
    x:standard_name = "projection_x_coordinate" ;
int y(y) ;
    y:units = "m" ;
    y:long_name = "y coordinate of projection" ;
    y:standard_name = "projection_y_coordinate" ;
double lon(y, x) ;
    lon:units = "degrees_east" ;
    lon:long_name = "longitude coordinate" ;
    lon:standard_name = "longitude" ;
double lat(y, x) ;
    lat:units = "degrees_north" ;
    lat:long_name = "latitude coordinate" ;
    lat:standard_name = "latitude" ;
short Altitude(y, x) ;
    Altitude:units = "meters" ;
    Altitude:long_name = "Altitude" ;
    Altitude:standard_name = "Altitude" ;
    Altitude:grid_mapping = "Lambert_Conformal" ;
    Altitude:coordinates = "lon lat" ;
    Altitude:missing_value = -99s ;
int time(time) ;
    time:units = "hours since 1900-01-01T00:00:00Z" ;
    time:long_name = "time in hours since 1900-01-01T00:00:00Z" ;
```

Les variables lat et lon sont de dimensions x et y. Une variable définit la projection utilisée, et celle-ci est référencée comme attribut pour chacune des variables. La dimension time est définie dans un vecteur en unité d'heures selon la date de référence du 1er janvier 1900. Cette date de référence commune permet de traiter plus facilement les données selon le temps car la date de référence du temps est la même pour tous les scénarios. Cela permet, entre autres, d'utiliser ncrat (NCO) afin de concaténer des fichiers. Cependant, il serait tout-à-fait possible de choisir une date postérieure si les données décrites se situent dans le temps avant le 1er

janvier 1900.

## 5 Variables

Les variables utilisant la convention NetCDF décrite dans les sections précédentes sont définies avec les attributs suivants (valeurs en exemple) :

```
float tas(time, y, x) ;
    tas:long_name = "air_temperature" ;
    tas:units = "K" ;
    tas:height = "2m" ;
    tas:grid_mapping = "Lambert_Conformal" ;
    tas:coordinates = "lon lat" ;
    tas:missing_value = -9999.f ;
```

Dans les fichiers de données SAFRAN, le nom des variables utilisées sont les suivants, selon les noms qui sont utilisés dans la nomenclature des données SAFRAN (unités SI/mks) :

- GLO : Rayonnement visible incident à la surface ( $W/m^2$ )
- PRCP : Précipitations liquides ( $kg/m^2/sec$ )
- Q : Humidité spécifique à 2 m ( $kg/kg$ )
- RAT : Rayonnement infra-rouge incident ( $W/m^2$ )
- SNOW : Précipitations solides ( $kg/m^2/sec$ )
- T : Température à 2 m ( $K$ )
- Vu : Vitesse du vent horizontal à 10 m ( $m/sec$ )

Les noms des fichiers des scénarios désagrégés utilisent plutôt la convention IPCC PCMDI en unités SI/mks :

- evapn : Evapotranspiration potentielle (méthode Météo-France Penman modifiée) ( $mm/sec$ )
- rsds : Rayonnement visible incident à la surface ( $W/m^2$ )
- prr : Précipitations liquides ( $kg/m^2/sec$ )
- prtot : Précipitations totales ( $kg/m^2/sec$ )
- hus : Humidité spécifique à 2 m ( $kg/kg$ )
- hur : Humidité relative à 2 m (%)
- hurminmax : Humidité relative maximale et minimale journalière à 2 m (variables hurmin et hurmax) (%)
- rlds : Rayonnement infra-rouge incident ( $W/m^2$ )
- prsn : Précipitations solides ( $kg/m^2/sec$ )
- tas : Température à 2 m ( $K$ )
- tasminmax : Température maximale et minimale journalière à 2 m (variables tasmin et tasmax) ( $K$ )
- uvas : Vitesse du vent horizontal à 10 m ( $m/sec$ )

Les valeurs manquantes sont fixées à -9999.

## 6 Localisation des données localement au CERFACS

Les données de l'analyse SAFRAN ainsi que des scénarios climatiques désagrégés sont entreposés localement sur disques au CERFACS. Les données SAFRAN, en date d'avril 2008, sont entreposées dans cette arborescence de répertoires (sujet à changements) :

```
/home/vittel/SAFRAN/  
    binary/  
    netcdf/  
    daily/  
        hourly/  
        monthly/
```

Le format des noms de fichiers se décompose comme suit :

```
Forc{VARNAME}.DAT_france_{BYEAR}{EYEAR}_{freq}.nc
```

où VARNAME est le nom de la variable (GLO, PRCP, Q, RAT, SNOW, T, Vu), BYEAR/EYEAR sont les années où débutent/terminent les données (2 chiffres), et freq la fréquence des données (hourly, daily, monthly).

Les données des scénarios désagrégés, en date d'avril 2008, sont entreposés dans cette arborescence de répertoires (sujet à changements) :

```
/home/evian/Desag/  
    netcdf/  
        EXPERIENCE/  
            MODELE/  
                annual/  
                seasonal/  
                monthly/  
                climatology/
```

où EXPERIENCE est le nom de l'expérience (exemple : SRESA1B, etc.), MODELE sont les acronymes des différents modèles (exemple : arpege1, mpi\_echam5, etc.).

Dans ces répertoires modèle, les données sont à intervalle journalier, le répertoire annual contient les données annuelles, seasonal les données moyennées sur des



saisons, monthly les données mensuelles, et climatology les données moyennées sur de longues périodes (exemple : 1971-1990, 2046-2065, etc.) y compris des données en anomalie par rapport à la période 1971-1990.

# Bibliographie

Ethan Davis. Dataset Discovery NetCDF Attribute Convention, 2005. URL [www.unidata.ucar.edu/software/netcdf-java/formats/DataDiscoveryAttConvention.html](http://www.unidata.ucar.edu/software/netcdf-java/formats/DataDiscoveryAttConvention.html).

Patrick Le Moigne. Description de l'analyse des champs de surface sur la france par le système SAFRAN. Technical report, Centre national de recherches météorologiques, Météo-France, 2002.

Program for Climate Model Diagnosis and Intercomparison. CF Metadata, 2008. URL [cf-pcmdi.llnl.gov/](http://cf-pcmdi.llnl.gov/).

Unidata. Unidata Integrated Data Viewer (IDV), 2007. URL [www.unidata.ucar.edu/software/idv/](http://www.unidata.ucar.edu/software/idv/).

Unidata. NetCDF Java Library Home, 2008. URL [www.unidata.ucar.edu/software/netcdf-java/](http://www.unidata.ucar.edu/software/netcdf-java/).