

# ANALYSE DE DONNÉES NETCDF AVEC XARRAY

PREPRINT, COMPILED SEPTEMBER 24, 2025

Gwenaël CAËR<sup>1</sup>

<sup>1</sup>CNRS, France

## ABSTRACT

Cet article présente l'utilisation de la bibliothèque Python Xarray pour la lecture, l'exploration et l'analyse de fichiers NetCDF. Les fichiers NetCDF sont largement utilisés pour stocker des données multidimensionnelles en sciences environnementales et océanographiques.

## 1 INTRODUCTION

Les fichiers NetCDF (Network Common Data Form) sont un standard pour le stockage de données multidimensionnelles, très utilisés en océanographie, climatologie et sciences de l'environnement. Python offre plusieurs bibliothèques pour manipuler ce format, dont **Xarray**, qui fournit une interface intuitive pour explorer et analyser les données multidimensionnelles.

Xarray se distingue par sa capacité à manipuler des **objets DataArray et Dataset** avec des dimensions et des coordonnées nommées, facilitant la lecture et le traitement des données scientifiques complexes.

## 2 LECTURE D'UN FICHIER NETCDF AVEC XARRAY

Pour lire un fichier NetCDF, on utilise la fonction `xarray.open_dataset()` qui retourne un objet **Dataset** contenant toutes les variables et leurs dimensions.

```
import xarray as xr
```

```
# Charger un fichier NetCDF
ds = xr.open_dataset("exemple_data.nc")
```

```
# Afficher le résumé du dataset
print(ds)
```

Le résumé du dataset fournit les informations suivantes :

- Les dimensions (par exemple, time, latitude, longitude)
- Les variables (par exemple, temperature, salinity)
- Les attributs globaux et les unités associées aux variables

## 3 EXPLORATION DES DONNÉES

Xarray permet de sélectionner et extraire des sous-ensembles facilement :

```
# Sélectionner la température à un moment donné
temp_t0 = ds['temperature'].isel(time=0)
```

```
# Extraire une zone géographique spécifique
temp_region = ds['temperature'].sel(latitude=slice(-10,10))
```

Les objets Xarray peuvent être visualisés directement avec **Matplotlib** ou via la méthode `.plot()` :

```
import matplotlib.pyplot as plt
```

```
# Visualisation de la température à t0
temp_t0.plot()
plt.title("Température à t0")
plt.show()
```

## 4 TRAITEMENTS ET ANALYSES

Xarray permet également d'effectuer des calculs et des statistiques sur les dimensions :

```
# Moyenne temporelle
temp_mean = ds['temperature'].mean(dim='time')

# Maximum global
temp_max = ds['temperature'].max()
```

Il est possible de combiner Xarray avec **Pandas** et **NumPy** pour des analyses plus avancées et générer des visualisations interactives.

## 5 CONCLUSION

Xarray simplifie la manipulation et l'analyse des fichiers NetCDF, en fournissant des outils intuitifs pour explorer les données multidimensionnelles, sélectionner des sous-ensembles, calculer des statistiques et visualiser les résultats. Cette approche est particulièrement utile pour les sciences marines, l'océanographie et l'étude du climat, où les jeux de données sont souvent volumineux et complexes.

## 6 RÉFÉRENCES

1. Hoyer, S., Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5(1), 10.

2. Rew, R., Davis, G. (1990). NetCDF: an interface for scientific data access. *IEEE Computer Graphics and Applications*, 10(4), 76-82.
3. Xarray documentation: <https://xarray.pydata.org>