

---

Ce travail doit être réalisé en binôme ; un contrôle de plagiat sera effectué sur les codes rendus.

## Contexte

Le réseau Centipede (<https://centipede.fr/>) est un réseau GNSS temps réel, *low-cost*, libre et participatif, mis en place en France en 2019 et en constant développement. Il permet de disposer gratuitement d'un positionnement RTK en réseau sur une grande partie du territoire Français, en Métropole et en Outre Mer et commence à s'étendre à l'international.

L'utilisation des données acquises par la station de référence constituant ce réseau pour des applications de géoscience est en cours d'évaluation ; l'intérêt est légitime : une station de référence *low-cost* coûte aux alentours de 500 € alors que les prix d'une station professionnelle « classique » peuvent facilement atteindre 15 000 €.

Dans ce contexte, on cherche à ici évaluer l'utilisation des stations de référence de ce réseau pour des applications de météorologie et de climatologie. En effet, initiés dès le début 1990, des travaux de recherche ont montré qu'il était possible d'évaluer la vapeur d'eau atmosphérique par GNSS via l'estimation de son influence sur la propagation des signaux mesurés pour le positionnement. Ainsi, depuis les années 2000, le GNSS est devenu un outil supplémentaire à disposition des communautés de la météorologie et climatologie, par la restitution de **contenu intégré en vapeur d'eau**, noté *IWV* (*Integrated Water Vapor* masse de vapeur d'eau au-dessus d'une surface d' $1 \text{ m}^2$ ) avec une précision de l'ordre de  $1 \text{ kg m}^{-2}$ .

**Le but de cette étude est d'évaluer les valeurs d'*IWV* estimées à partir d'un ensemble de stations de référence du réseau Centipede.**

On dispose pour cela de mesure d'*IWV* restitués par les stations du réseau le 30 décembre 2022 à 9h.

Les données sont contenues dans le fichier `FR_CTP_iwv.txt` et sont organisées de la manière suivante :

- Colonne 1 : coordonnée *E* de la station (RGF93 / Lambert93, en km)
- Colonne 2 : coordonnée *N* de la station (RGF93 / Lambert93, en km)
- Colonne 3 : *IWV* en  $\text{kg m}^{-2}$ .

Les données peuvent être chargées simplement comme suit :

```
Python
data = np.loadtxt('FR_CTP_iwv.txt')
E_obs = data[:,0].reshape((-1, 1))
N_obs = data[:,1].reshape((-1, 1))
IWV_obs = data[:,2].reshape((-1, 1))
```

On dispose également du fichier `contour_france.txt` contenant les coordonnées ( $E$ ,  $N$  en **km**) dans le système de coordonnées RGF93 / Lambert93 du contour de la France Métropolitaine (hors Corse).

On considère enfin les méthodes d'interpolation implémentées en **Python** lors des séances pratiques précédentes :

- `interp_lin` : interpolation linéaire à partir d'une triangulation.
- `interp_inv` : interpolation par inverse des distances (utilisation des points les plus proches possible).
- `interp_spl` : interpolation par splines (un coefficient de lissage de l'ordre de 50 est adapté).
- `interp_krg` : interpolation par krigeage avec un variogramme linéaire ; ce choix du variogramme devra être justifié ; il est inutile d'ajuster le variogramme analytique au-delà de 500 km.

## 1 Validation d'une méthode d'interpolation

On cherche à évaluer par validation croisée (il est conseillé de vous limiter à un nombre de point réduit, environ 50 – 100) chacune des méthodes d'interpolation proposées ; on se limitera au calcul des statistiques suivantes : erreur moyenne, écart-type de l'erreur, erreur moyenne quadratique.

Conclure quant à la méthode d'interpolation la plus adaptée à ces données en utilisant la méthode de validation croisée détaillée dans le cours.

Quelques fonctions qui pourront être utiles à cette étape de validation croisée :

- `np.delete` : Supprime un élément d'une matrice en fonction de son rang dans une liste.
- `np.random.randint` : Renvoie un entier aléatoire appartenant à une plage de valeurs possibles.
- `np.nanmean` : Réalise la moyenne d'un tableau de nombre sans tenir compte des valeurs NaN (*Not a Number*).
- `np.nanstd` : Réalise l'écart-type d'un tableau de nombre sans tenir compte des valeurs NaN (*Not a Number*).
- `np.isnan` : Renvoie 1 si le paramètre est un NaN (*Not a Number*).

L'exemple suivant permet de tirer aléatoirement un rang `rg` entre 0 et le nombre d'élément du tableau `vm_obs` défini précédemment, puis de supprimer cet élément :

```
Python
rg = np.random.randint(0,vm_obs.shape[0])
vm_del = np.delete(vm_obs,rg,axis=0)
```

## 2 Cartes de contenu intégré en vapeur d'eau

Visualiser la carte de contenu intégré en vapeur d'eau issue de la méthodes identifiée comme la plus pertinente à l'issue de la validation croisée.

Pour agrémenter vos cartes, quelques consignes à suivre :

- Compléter les cartes à l'aide du contour de la France ; la commande suivante permet d'afficher ce contour en noir avec un trait de taille 2 :

```
Python
import matplotlib.pyplot as plt
contour = np.loadtxt('FR_contour.txt')
plt.plot(contour[:,0],contour[:,1], 'k', lw=2)
```

- Ne pas interpoler de valeurs d'IWV au-delà du contour de la France. Pour cela, on peut utiliser la classe `Path` de `matplotlib` et sa méthode `contains_point`, appliquée au contour :

```
Python
from matplotlib.path import Path
path = Path(np.hstack( (xc,yc) ))
point = np.array( [[x, y]] )
in_path = path.contains_points(point)
```

- Utiliser une carte de couleur adaptée aux données météorologiques ; cette carte de couleur est contenue dans le fichier `nc1.rgb` et se charge de la manière suivante :

```
Python
import matplotlib as mpl
cmap_ncl = mpl.colors.ListedColormap(np.loadtxt('nc1.rgb')/255)
```

- Indiquer les positions des stations.

## 3 Comparaison aux données du RGP

On dispose des valeurs d'IWV restitués par des stations de référence du Réseau GNSS Permanent français (RGP, <http://rgp.ign.fr>) situées à moins de 20 km des stations Centipede ; ces données sont reconnues pour leur qualité (et sont assimilées de manière

routinière dans les modèles de prévision numérique du temps de Météo-France) et serviront donc de référence.

À l'aide de la méthodes identifiée comme la plus pertinente à l'issue de la validation croisée, interpoler les valeurs d'IWV issues du réseau Centipede au niveau de chacune des stations du RGP ; comparer les valeurs interpolées aux valeurs de référence :

- En calculant l'erreur moyenne, l'écart-type de l'erreur, l'erreur moyenne quadratique et le coefficient de détermination.
- En visualisant le diagramme de corrélation, l'histogramme et la carte des erreurs.

Conclure quant à la qualité des données issues du réseau Centipède.

## 4 Rendu

À rendre sur <https://cours-fad.ensg.eu> (un seul rendu par binôme) à la date convenue une archive compressée (nommée selon `interp_nom_eleve1_nom_eleve2`) contenant :

- Un rapport (format `pdf`), justifiant les choix (type de variogramme, lissage, etc.), synthétisant les calculs réalisés, présentant les résultats obtenus et concluant quant à l'objectif de l'exercice.
- Le code `Python` correspondant, commenté et fonctionnel.