



TD– analyse de température

Etude de l'évolution du climat mondial Exploitation des données CRUTEM4

Remarques préliminaires / consignes :

- vous rédigerez vos réponses sur feuille de copie ;
- le programme python résultant de ce travail sera envoyé par mail à Mr Pernot : nicolaspernot@hotmail.com

Contexte du travail :

L'étude de l'évolution du climat passé, au cours des deux derniers siècles, est un enjeu majeur pour anticiper les évolutions futures, à court et moyen termes, du climat mondial. En effet, les projections les plus pessimistes font état d'un réchauffement global, en l'an 2100, de l'ordre de 4°C à 6°C ce qui induirait des changements radicaux dans l'organisation de nos sociétés.

L'Unité de Recherche sur le Climat (CRU) de l'Université East Anglia propose un ensemble de données relatives à l'historique du climat sur le globe, qui sont mises à disposition des chercheurs pour être analysées¹. Ces données sont : la température de surface, la pression, la pluviométrie. Les données sont agrégées à partir des mesures provenant de plus de 5500 stations météo du monde entier. Elles sont traitées, mises en forme et mises à disposition du public sous la forme de fichier. On propose ici de travailler sur l'exploitation des données de température de surface (TEM), recueillies entre janvier 1850 et septembre 2013, et disponibles dans un fichier appelé texte très volumineux : CRUTEM4².

Objectif du travail :

L'objectif du travail demandé est de répondre aux 2 questions suivantes :

- le climat s'est-il réchauffé au cours des **163 dernières années** ?
- Si oui, dans quelle mesure ?

Contenu du fichier CRUTEM4 :

Le fichier CRUTEM4 est un fichier "d'anomalies mensuelles de températures"³. En effet, ce ne sont pas les températures absolues qui sont reportées, mais la différence vis-à-vis d'une température de référence, calculée comme la moyenne des températures sur une période suffisamment longue, entre 1961 et 1990. Les données présentes dans le fichier sont donc des températures relatives, en °C. Par exemple :

- une donnée de +0.1°C en septembre 1924 indique que, ce mois-là, la température moyenne a été 0.1°C plus élevée que la température moyenne des mois de septembre entre 1961 et 1990.
- une donnée de 0°C en avril 1994 indique que la température moyenne ce mois-là est la même que celle de tous les mois d'avril entre 1961 et 1990.

¹ <http://www.metoffice.gov.uk/hadobs/crutem4/>

² <http://www.metoffice.gov.uk/hadobs/crutem4/data/download.html>

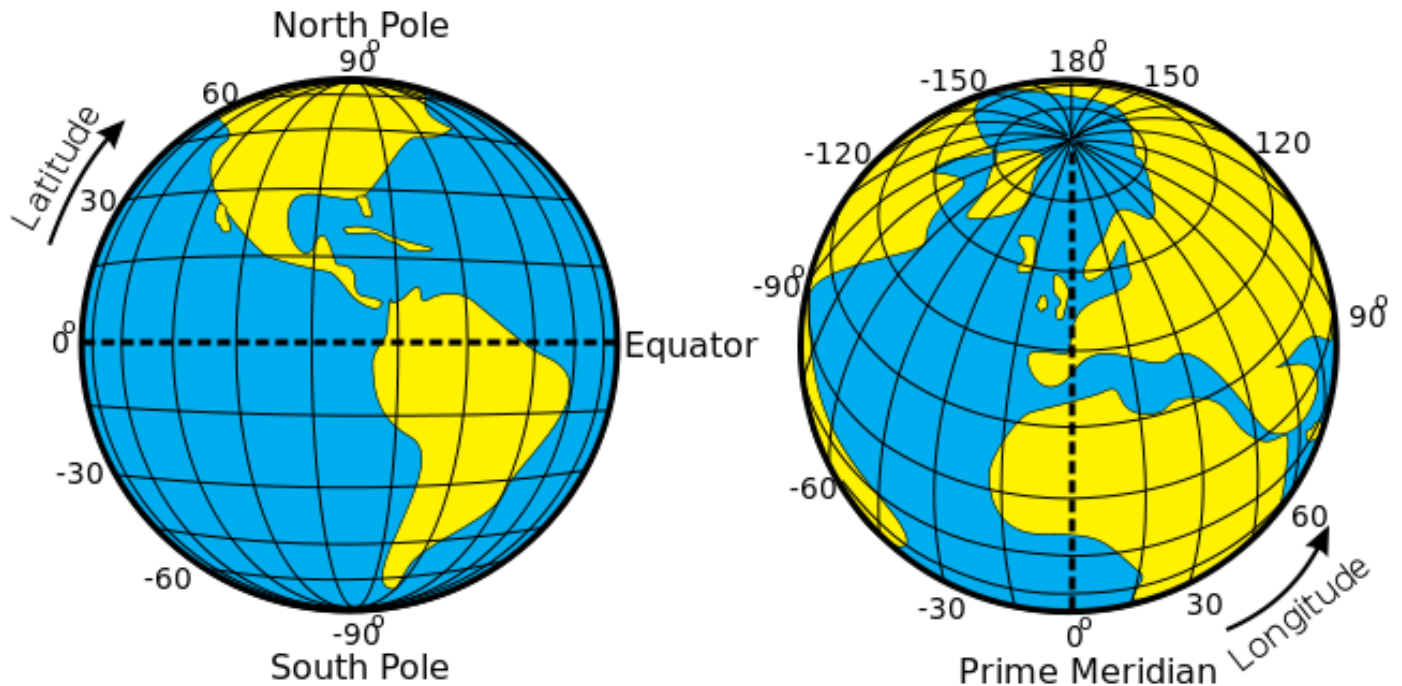
³ http://www.odlt.org/dcd/ballast/temperature_anomaly.html



TD– analyse de température

Maillage du globe :

La localisation des températures est donnée par une grille de 5° de résolution, qui définit le maillage d'un planisphere. Cette résolution de 5° vaut pour la longitude et la latitude.



Pour rappel :

- la latitude évolue entre 0° (équateur) et 90° (le pôle nord) pour l'hémisphère nord ;
- la latitude évolue entre 0° (équateur) et -90° (le pôle sud) pour l'hémisphère sud ;
- la longitude évolue de -180° (ou 180° Ouest) à $+180^\circ$ C (ou 180° Est) en passant par 0° (le méridien de greenwich).



TD– analyse de température

Travail demandé

Problématique 1 : étude de l'adéquation de la taille du fichier aux ressources de la machine de traitement. L'objectif est de vérifier que le fichier de mesure peut être ouvert, chargé en mémoire, et traité par un ordinateur personnel de type PC disponible au laboratoire de SII.

Question 1.

Soit r_{long} , resp. r_{lat} , la résolution en longitude, resp. en latitude, du maillage du globe. Exprimer le nombre de carreaux permettant de mailler le planisphère en fonction de r_{long} et r_{lat} . Vous ferez une application numérique pour $r_{long}=5^\circ$ et $r_{lat}=5^\circ$, valeurs utilisées dans ce problème. Calculer l'ordre de grandeur, en km, de la longueur d'un carreau au niveau de l'équateur (avec rayon équatorial = 6378km).

Question 2.

Si n est le nombre d'années pour lesquelles des mesures ont été relevées, et sachant que les relevés de température sont mensuels, exprimer N_t le nombre total de mesures en fonction de n , de r_{long} et de r_{lat} .

Question 3.

On a reproduit ci-dessous 3 valeurs telles qu'elles apparaissent dans le fichier. Un relevé de température est écrit, dans le fichier de mesure, sous la forme de chaînes de caractères suivantes (extrait):

"9.125e-01"

"1.155e+00"

"-2.000e-01"

Quelles sont les valeurs réelles correspondantes ?

Question 4.

L'encodage utilisé pour le fichier de données est ASCII. Quel est le nombre d'octets requis pour encoder une anomalie de température positive ? négative ? Dans la suite, vous considérerez qu'en moyenne, il y a autant d'anomalies positives que négatives, et vous pourrez prendre la moyenne du nombre d'octets = 9.5.

Question 5.

En considérant le résultat de Q2 étant : $777600n / r_{long} * r_{lat}$, déterminer le nombre d'octets requis pour encoder l'ensemble des données de 1850 à 2013. Vous exprimerez le résultat en ko et Mo.

Question 6.

Vérifiez la taille du fichier CRUTEM4 = 56311005 octets. Comparer le résultat précédent par rapport à la taille du fichier. Quelle est l'erreur, en %, de votre résultat par rapport à cette valeur ? Expliquer les écarts observés.

Question 7.

En annexe, on a figuré les caractéristiques techniques d'un ordinateur de bureau de type PC, en environnement Windows, sur lequel on souhaite analyser les données du fichier de mesure. Quelle est la taille de la mémoire vive disponible? Quelle est la capacité de stockage du disque dur? Ces valeurs permettent-elles de garantir que l'on pourra sauvegarder le fichier et le charger en mémoire ?

Question 8.

Dans l'hypothèse où la moitié de la mémoire vive disponible peut être allouée au chargement du fichier de mesure, quelle serait la taille maximale du fichier correspondant ? Dans ce cas, en considérant la période de 1850 à 2013 inchangée, quelle serait la résolution en longitude et latitude ? Quelle serait, dans ce cas, la longueur, en km, du carreau au niveau de l'équateur ?

Question 9.

Proposer une conclusion synthétique à cette partie.



TD– analyse de température

Problématique 2 : tracé des courbes mensuelles, annuelles, décennales de température. On souhaite, procéder à l'analyse des mesures portées dans le fichier CRUTEM4. L'objectif est de tracer 3 courbes avec MATPLOTLIB :

- la courbe donnant l'évolution, en fonction du temps, des températures mensuelles ;
- la courbe donnant l'évolution, en fonction du temps, des températures annuelles ;
- la courbe donnant l'évolution, en fonction du temps, de la température moyenne des 10 années précédentes.

Structure du fichier CRUTEM4:

Dans le fichier CRUTEM4, les résultats mensuels sont présentés sous la forme de blocs (il y a, dans le fichier, autant de blocs que de mois). Chaque bloc est constitué de 37 lignes (1 ligne commentaires + 36 données) :

- la première ligne de chaque bloc a la forme suivante :
`1902 6 1 36 rows 72 columns. Missing=-1.000e+30 Version=CRUTEM.4.2.0.0 Field_status=f`

Le premier nombre indique l'année, le deuxième le mois (de 1 à 12). *Le 1 après ne sert à rien mais est présent dans le fichier.* Les nombres de lignes et de colonnes correspondent au maillage du globe vu dans la première partie. Il est, dans tout le fichier, de 36 lignes et 72 colonnes. L'information "Missing=-1.000e+30" est la valeur qui est utilisée, dans le fichier, pour signifier que l'information concernant l'anomalie de température est manquante. Il est important donc de ne pas prendre en compte cette valeur dans les questions qui suivent et il y en a pas mal (au risque de fausser complètement les résultats que vous trouverez). Les autres informations n'ont pas d'influence sur le travail demandé.

- les 36 lignes suivantes correspondent chacune aux 72 anomalies de température (*incluant des valeurs manquantes* : -1.000 e +30)

Question 10.

Créez un script « temperature_votre nom ».

Commencer votre script en ouvrant le fichier CRUTEM.4.2.0.0.anomalies.txt et chargez le contenu du fichier dans la variable `file_content` avec l'instruction `readlines()` puis fermer le fichier

Question 11.

Sachant que le nombre de mois écoulés entre le 1er janvier 1850 et le mois m de l'année y dont on souhaite connaître les températures est tel que ($numligne$ est le numéro de la ligne recherchée) :

$$numligne = (y-1850)*12*37 + (m-1)*37 + 1$$

Ajouter au programme une fonction `def get_data(y, m)` retournant une liste de réels, permettant d'obtenir, pour l'année y et le mois m (via $numligne$), une liste (qu'on appellera « data ») contenant les valeurs réelles (pas de string !) des anomalies de température pour le globe complet (bloc de 36 lignes de données).

Question 12.

Ajoutez à votre programme une fonction permettant de calculer la moyenne des valeurs `def moyenne(data)` contenues dans la liste construite dans la question 11. Vous prendrez garde, évidemment, à ne pas prendre en compte les valeurs correspondant à une information manquante ("-1e+30"). Nota : cette moyenne sera un réel.

Question 13.

Ajoutez à votre programme les instructions permettant d'obtenir les 3 listes suivantes :

- la liste des températures mensuelles, du premier au dernier mois ;
- la liste des températures annuelles, de la première à la dernière année ;
- la liste des températures dans la décennie précédent une année donnée. Cette liste contiendra autant d'éléments que d'années (par exemple, pour l'année 1904, on calculera la moyenne des années 1894-1904, pour l'année 1905 la moyenne 1895-1905 etc.). Pour information, on parle de "moyenne mobile" sur 10 ans.

Les résultats correspondant seront stockés dans les listes `temp_mensuelle`, `temp_annuelle`, `temp_decennie`.



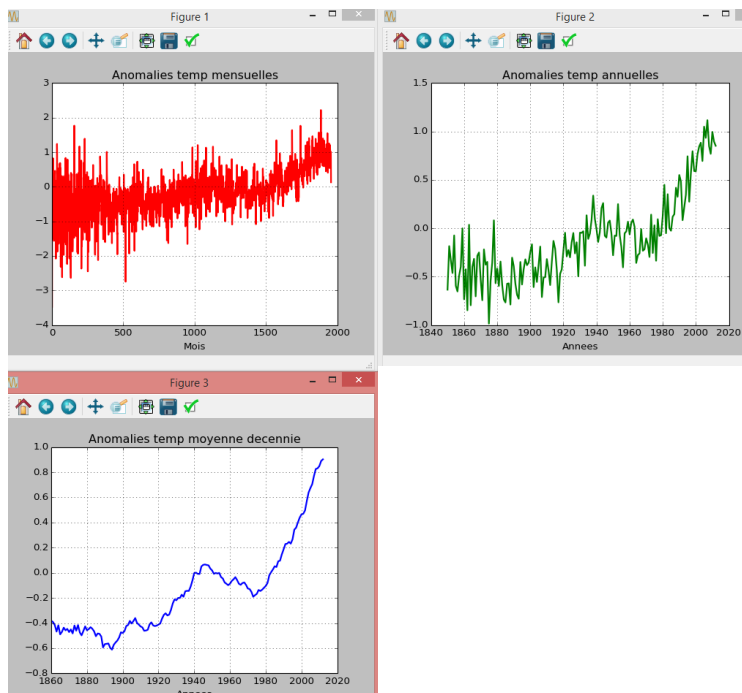
TD– analyse de température

Question 14.

En utilisant la bibliothèque de tracé de courbes matplotlib, représenter ces 3 listes sur 3 graphiques distincts en créant les fichiers PDF.

Que concluez-vous quant à l'évolution des températures ? Que pouvez-vous conclure de cette partie ?

Voilà les courbes que vous devriez obtenir si votre programme est bon



Caractéristiques Techniques de la machine de traitement

Processeur

Processeur Intel® Core™ i7-4770U de 4e génération (8 Mo cache, jusqu'à 3.9 GHz)

Système d'exploitation

Windows 8.1 64bit , Français

Basique

Base standard XPS 8700

Logiciels de productivité

Microsoft Office 2013 Version d'évaluation, interface utilisateur multilingue

Logiciels de sécurité

McAfee® LiveSafe, abonnement de 12 mois - Français

Moniteurs

Écran non inclus

Mémoire^L

16 Go de mémoire bicanale DDR3 à 1 600 MHz, 4 emplacements DIMM

Disque dur

Disque dur SATA de 2 To à 7 200 tr/min+ 32Go SSD (Intel SRT)

Carte vidéo

Carte graphique NVIDIA® GeForce GTX 660 avec 1.5 Go de mémoire DDR5

Lecteur optique

Lecteur combo Blu-ray (lecture BD et gravure DVD/CD)

Logiciels à télécharger

Dell Backup and Recovery Professional for CSMB, Digitally Delivered

Clavier

Clavier multimédia USB Dell KB213, français

Souris

Souris laser Dell