

TP3 – Corrélations et régressions linéaires

- 1) Vérifier que les séries de paramètres (O3, Temp, CO et CO2) ne comportent pas de point erroné, les supprimer s'il y en a, et sauvegarder les séries corrigées dans un nouveau fichier .mat ou .txt.

Exécution du programme TP3_1.m

Il y a une valeur négative dans la série de données O3 (déjà vu dans le TP1) ; il faut la retirer.

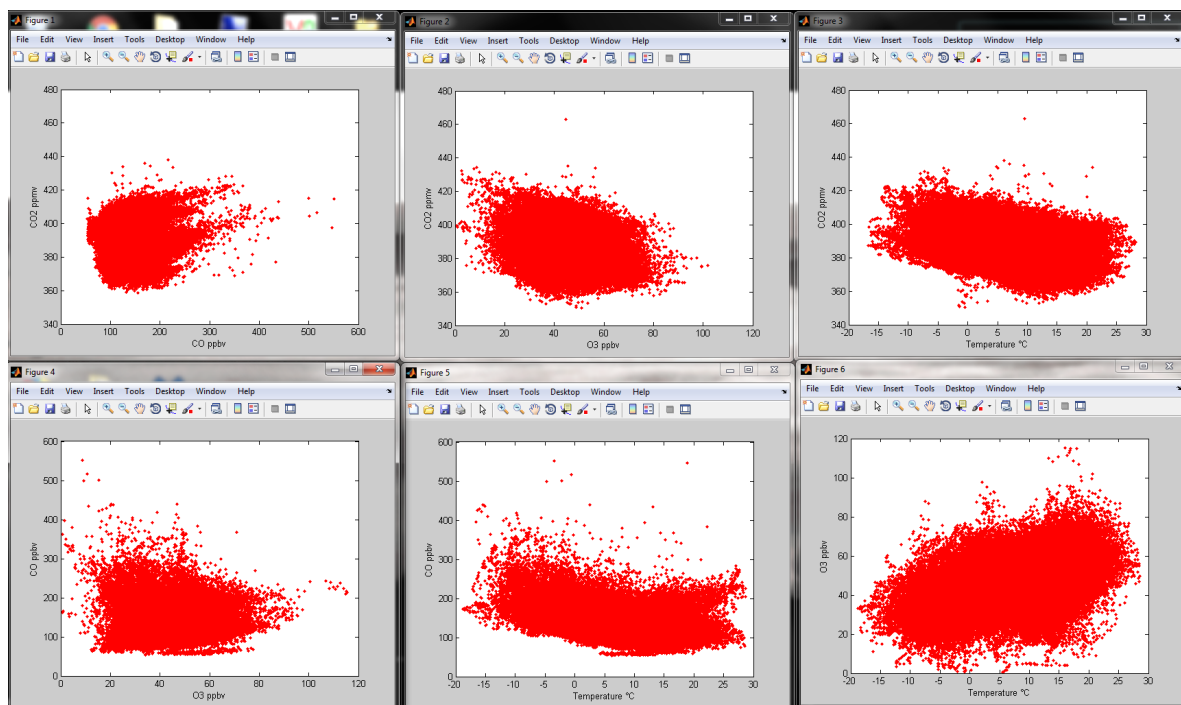
On peut faire le même nettoyage des données CO et CO2, par contre les valeurs négatives de température sont bien sûr à garder.

La série de données CO ne semble correcte qu'à partir de 2002 : seules les données post-2002 sont donc gardées.

Les séries de données Température et CO2 semblent correctes.

Sauvegarde des données « nettoyées » dans Station_PDD_horaire_1995_2017_corrige

- 2) À partir de ce nouveau fichier, tracer les paramètres (O3, Temp, CO et CO2) les uns en fonctions des autres soit 6 figures. On appelle ces types de graphique des graphes de dispersion (scatterplot en anglais). Discuter qualitativement ces résultats.



Exécution du programme TP3_2.m

L'ozone semble légèrement corrélée avec la température (production photochimique, plus de soleil en été)

Le CO₂ est anticorrélé avec la température et donc l'ozone (chauffage en hiver, plus d'absorption par les plante en été).

Le CO est anticorrélé avec la température (chauffage en hiver).

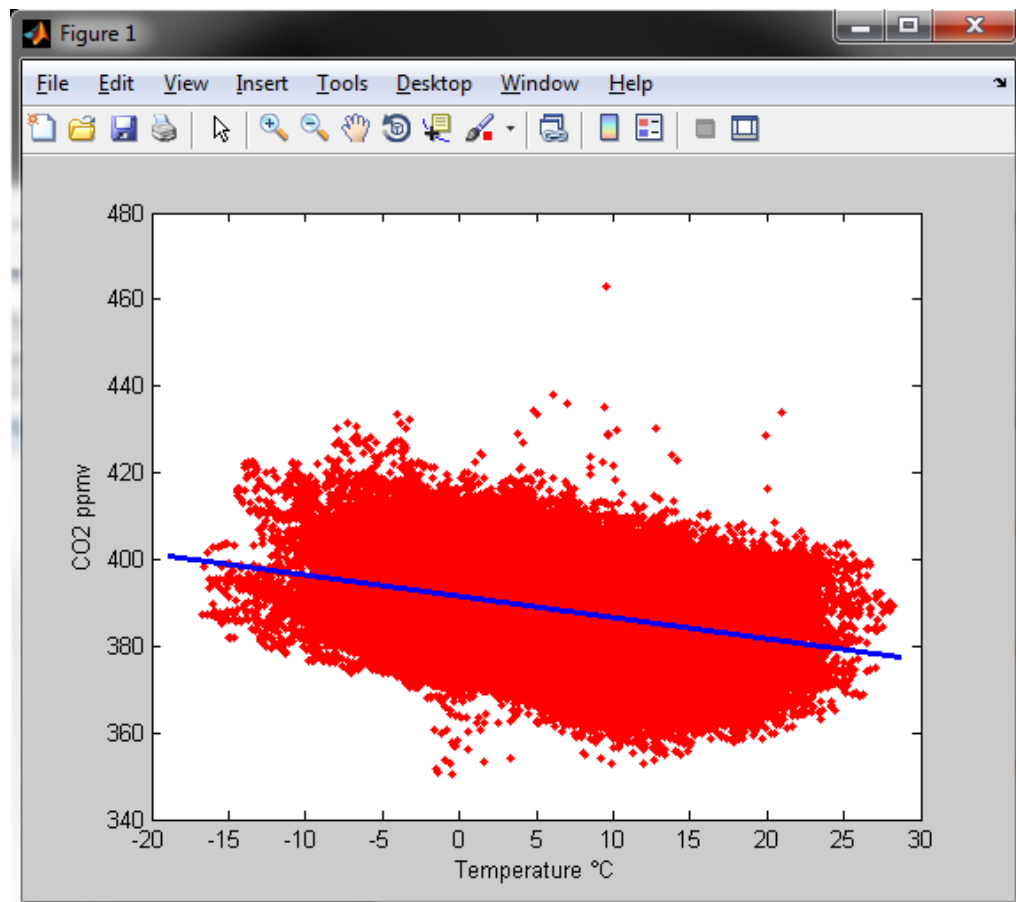
La hauteur de couche limite de l'atmosphère peut jouer un rôle également. Elle est plus haute en été qu'en hiver, donc le Puy de Dôme (PDD) a tendance à être d'avantage dans la couche limite en été, ce qui augmente donc les valeurs de CO, et CO₂ en été. Ce processus diminue probablement la netteté des corrélations.

Pas de corrélation nette entre CO et CO₂, et entre CO et ozone.

- 3) Calculer les coefficients de la droite de régression et le coefficient de corrélation dans le cas du CO₂ en fonction de la température, d'abord en programmant les équations (3), (4) et (6), puis en utilisant les fonctions `corrcoef` et `polyfit`. Tracer le graphe de dispersion du CO₂ vs température en y superposant la droite de régression. Analyser vos résultats.

Exécution du programme TP3_3.m

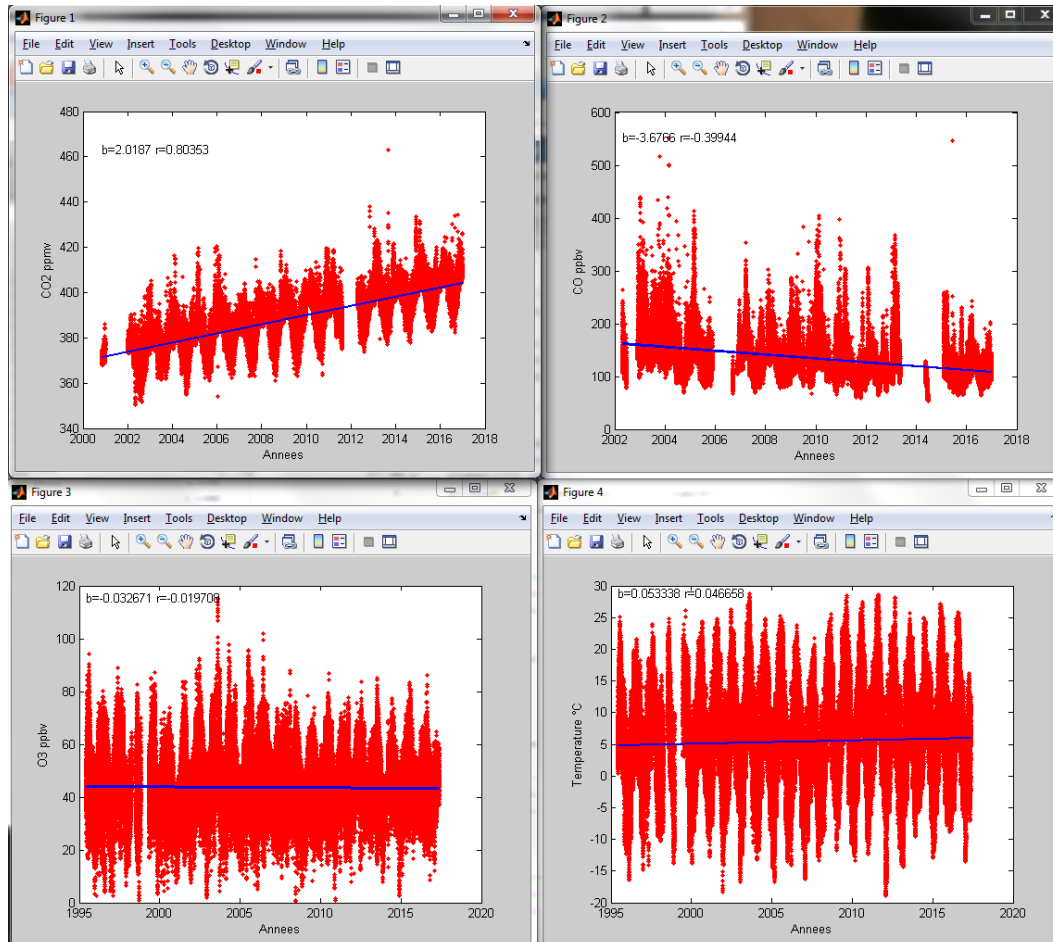
Pour le CO₂, on trouve $a = -0.4928$, $b = 391.4447$, $r = -0.3108$, et on obtient la figure suivante :



L'anticorrélation se traduit par la valeur négative de la pente de la droite (a) et du coefficient de corrélation (r). La valeur absolue de r est assez faible (-0.3).

- 4) Tracer les paramètres en fonction du temps et calculer leur tendance à long terme. Analyser vos résultats.

Exécution du programme TP3_4.m



Les tendances décennales obtenues sont de 20 ppmv de CO₂ par décade, -36,8 ppbv de CO par décade, -0,3ppbv d'O₃ par décade et +0,5°C par décade.

Les valeurs des coefficients de corrélation suggèrent que l'augmentation du CO₂ est significative, et éventuellement la diminution de CO.

Note : la diminution de CO est due à l'amélioration de la limitation des rejets industriels et des voitures.

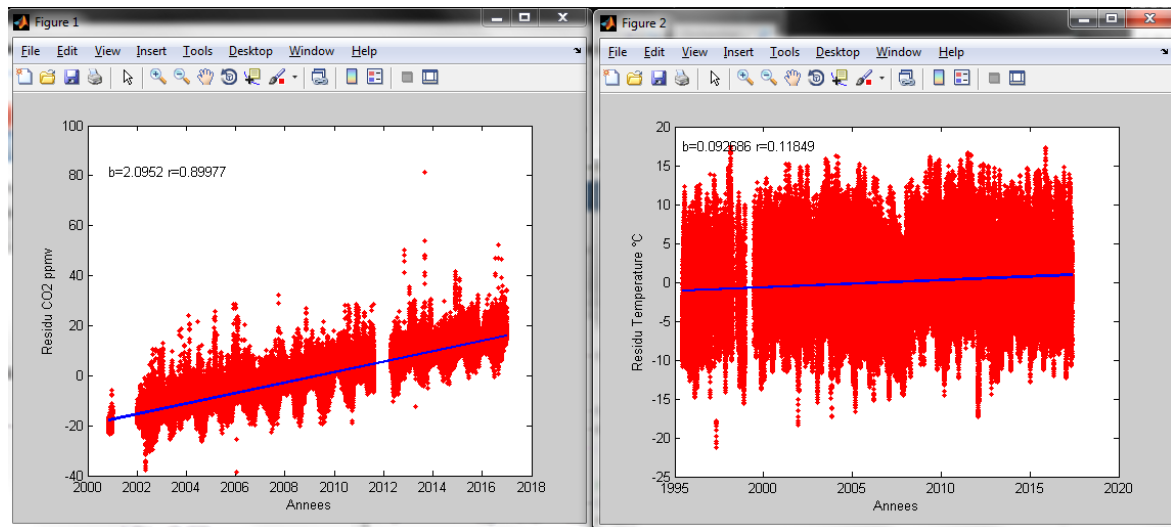
La température augmente de 0,5°C pour 10 ans, ce qui est beaucoup. C'est l'effet du changement climatique.

L'ozone est stable, sa concentration ne subit pas de tendance à long terme significative.

- 5) Désaisonnaliser les données de CO₂ et de température en retirant aux données de CO₂ les moyennes pour chaque mois de l'année. Attention : prendre *toutes* les données d'un même

mois, pour toutes les années. Calculer les coefficients de régression et de corrélation et tracer les tendances à long terme sur les résidus. Analyser vos résultats.

Exécution du programme TP3_5.m



En appliquant la désaisonnalisation avant la régression linéaire, la tendance de CO2 ne change pas significativement (~ 20 ppmv/décade) mais le coefficient de corrélation est légèrement amélioré (de 0.803 à 0.899).

L'augmentation de température passe quant à elle de $0,5^{\circ}\text{C}$ par décade à $0,9^{\circ}\text{C}$ par décade. Cependant, la corrélation n'est que légèrement améliorée (de 0.046 à 0.118) et reste faible.

L'application de cette méthode de désaisonnalisation sur les autres variables ne changerait sans doute pas grand-chose car ces variables ne semblent pas ou pas fortement corrélées.