

## **Sujet du Stage**

Notre stage a été effectué sur la plateforme QualAir qui fait partie du LERMA ainsi que du réseau de stations de AirParif qui permet de contrôler la qualité de l'air sur le site de Jussieu. Parmi les activités que proposent le LERMA, notre activité durant ce stage était une analyse de données à la frontière entre la physique et les sciences de la Terre. En effet notre sujet d'étude était la qualité de l'air et l'évolution de la pollution urbaine.

Les enjeux autour de l'environnement occupent une place de plus en plus importante dans notre société, c'est à la fois une prise de conscience individuelle, mais aussi une nécessité pour l'avenir et pour notre santé. La qualité de l'air en ville est représentative de notre mode de vie et du nombre d'industries et de machines produisant du dioxyde de carbone, mais aussi du monoxyde de carbone, des particule fine, etc .... Cependant, si l'on prend l'exemple du monoxyde de carbone, il peut être mortel, à haute concentration, pour l'être humain. C'est pourquoi il est nécessaire de surveiller le taux de ces différents polluants qui ne sont pas seulement néfastes pour l'être humain mais pour les animaux et l'environnement en général, pouvant mener à un déséquilibre naturel. AirParif, dont fait partie la plateforme QualAir, permet de suivre les concentrations de divers polluants et ainsi alerter et prévenir la population des dangers potentiels. L'installation sur laquelle il nous a été donné de travailler est l'analyseur CO11M qui relève la concentration de monoxyde de carbone en continu.

Notre travail pendant ce stage a été de traiter convenablement un ensemble de données afin de proposer des hypothèses vérifiables et d'aboutir à des conclusions rationnelles. Les données traitées ici représentaient l'évolution de la concentration en CO dans l'atmosphère au niveau de la station QualAir à Jussieu, ainsi que d'autres variables permettant d'obtenir des précisions sur les conditions d'acquisition de ces valeurs. En outre on disposait de 12 fichiers de données, sachant que chaque fichier couvrait un mois, on a ainsi pu étudier l'évolution de la concentration en CO dans l'air autour de Jussieu du 01/12/17 au 31/11/18.

# **Travail Réalisé**

## **Introduction**

Pour notre stage nous avons pu travailler grâce à des données fournies par un analyseur CO11M qui analyse la concentration en CO dans l'atmosphère par spectroscopie. Cette analyse de l'atmosphère est une analyse locale, ce qui signifie que l'appareil prélève localement l'air du site de Jussieu, qui passe ensuite dans la chambre de mesure, où un filament chauffé va émettre un rayonnement infrarouge d'environ  $4,7\mu\text{m}$  (car le spectre d'absorption du monoxyde de carbone est maximum à  $4,67\mu\text{m}$ ). Ce rayonnement va traverser une roue de corrélation, entraînée par un moteur, qui possède 3 secteurs :

- Un secteur rempli de CO d'une concentration connue permettant d'étalonner la machine
- Un secteur opaque permettant à la machine de "recaler" son 0 à chaque fois. En effet en passant à ce secteur la machine sait qu'elle doit obtenir une concentration de CO nulle.
- Un secteur transparent permettant au rayonnement d'accéder à la chambre de mesure

Le rayonnement qui est passé par la chambre de mesure va être récupéré grâce à un détecteur infrarouge semi-conducteur Pb-Se et les données seront relevées sur un fichier par mois.

## **Tri des données**

Etant donné le grand nombre de données, étalées sur un an, dans des fichiers séparés, il nous a semblé judicieux de concaténer les données dans un seul et même fichier à l'aide de MATHLAB.

Une fois cela fait, nous avons tracé la concentration de CO en fonction du temps (figure 1). Cependant le résultat n'était pas très concluant, il était difficile d'en tirer des conclusions. Pire encore, certaines valeurs (par exemple les pics abrupts et extrêmement élevés) nous semblaient étranges, nous en avons donc conclu qu'avant toute analyse des données il faudrait trier les valeurs.

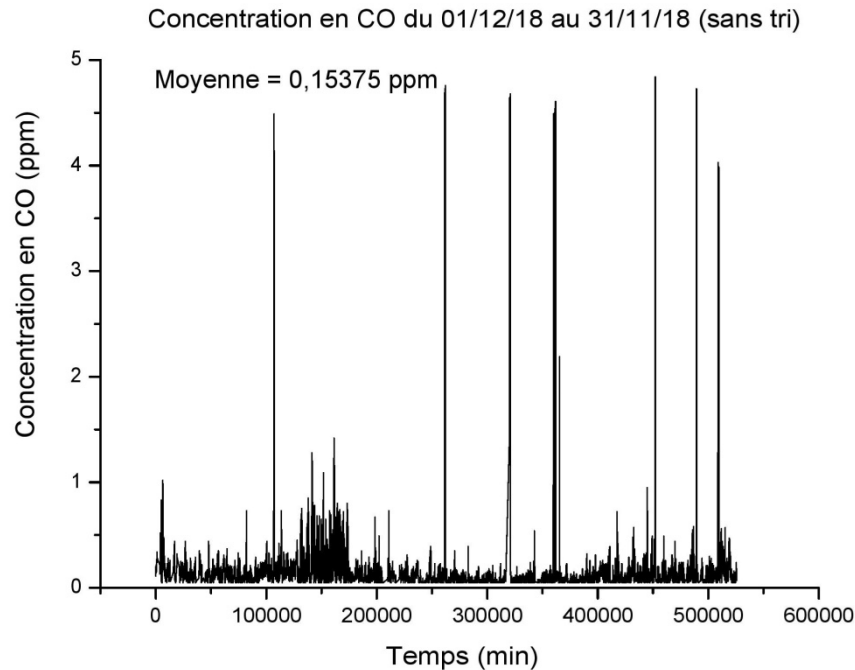


Figure 1 : Graphique de la concentration en CO en fonction du temps sur un an (sans tri)

Les fichiers à trier se présentaient sous la forme d'un tableau de 7 colonnes contenant : la date, l'heure, l'incrémentation par quart d'heure (0min, 15min, 30min, 45min, 60min, 75min, ...), un nombre représentant de potentiels défauts techniques, mesure du CO en ppm, pression en hPa et la tension d'alimentation en Volt. Les valeurs de CO et de pression obtenues ne sont cependant pas des mesures ponctuelles, ce sont des moyennes coulissées obtenues à partir d'un grand nombre de valeurs, le CO11M effectuant des mesures plusieurs fois par minute.

## Premier tri

On a effectué un premier triage sur les erreurs de mesure évidentes de l'appareil en fonction de 3 paramètres : la tension, la pression et les messages d'erreur de l'appareil.

Le triage sur la tension consistait à garder uniquement les valeurs du tableau pour lesquelles la tension est de 4,94 en considérant que si la tension est différente alors l'appareil ne peut pas fonctionner correctement.

On a ensuite choisi un intervalle de  $\pm 20$  hPa autour d'une pression de référence de 1007 hPa (qui représente la moyenne annuelle de la pression). Ce triage permet de toujours prendre la pression dans un certain intervalle de confiance et que les mesures de CO associées soient cohérentes entre-elles. L'intervalle choisi de  $\pm 20$  hPa est l'intervalle pour lequel on a estimé les valeurs de pression plausibles et même si certaines valeurs de pressions peuvent parfaitement être correctes au-delà de cet intervalle (ce qui revient à supprimer des valeurs pouvant être utilisées pour l'analyse), notre but est d'uniquement conserver des valeurs justes mais au vu de la grande quantité de données on peut se permettre de supprimer quelques bonnes données si cela signifie se débarrasser de toutes les données erronées.

Pour expliquer le tri sur le message d'erreur nous nous sommes référés au manuel d'utilisation et résumé dans le tableau suivant. Il faut noter que les alertes sont rangées par ordre d'importance (plus le code est élevé plus le problème est grave) et que si plusieurs alertes sont présentes alors celles-ci

vont se sommer (par exemple si la machine a un problème à la fois : de chopper, de débit et de vitesse moteur on aura message d’alerte de 0090).

Code	CO11M
---01	Générale
---02	Vitesse moteur
---04	Température optique
---08	Débit
---10	Signal
---20	Pression
---40	Température interne
---80	Chopper

Tableau des différents messages d’erreur de la machine CO11M d’après le manuel d’utilisation.

Dans les données fournies, nous avons remarqué qu’un message d’erreur revenait tout le temps : 0002. Ce qui signifie que durant toute l’année la machine a eu un problème sur la vitesse du moteur, cependant cela n’influe pas sur les mesures car une donnée est une moyenne glissée qui permet ainsi de “lisser” les potentielles erreurs dont celle-ci. Nous avons donc uniquement gardé les valeurs ayant un message d’erreur de 0002 ou en dessous.

Ce premier tri a retiré 1969 mesures soit environs 5.7% des données d’origine.

Il est important de noter qu’une grande partie du travail de ce premier tri était de bien vérifier après chaque étape du programme utilisé sur nos fichiers pour les trier qu’il effectuait bien l’action demandée. Pour donner un exemple : la première étape de notre programme était de rassembler les 12 fichiers en un unique grand fichier pour simplifier le tri, il nous a donc fallu vérifier à la main si toutes les valeurs étaient bien présentes et qu’il n’y avait pas de décalage. On a choisi une valeur au hasard dans les données brutes, puis, grâce, à la colonne de la date et de l’heure on a confirmé que les valeurs de la pression, message d’erreur, concentration en CO et tension coïncidaient bien entre elles. On a répété le même processus à chaque étape de notre tri.

## Second tri

Il s’est avéré qu’après le premier tri en traçant le graphique de la concentration de CO sur un an on obtenait encore des zones absurdes, comme par exemple une concentration de CO nulle sur plusieurs jours ou alors un pic important autour de 4 ppm formant une sorte de plateau comme si l’appareil était bloqué à cette valeur. Il a donc été nécessaire d’effectuer un deuxième tri plus pointu et effectué à la main.

En traquant graphiquement, sur Origin61 :

- Les pics abrupts (passant de 0,1 ppm à 4ppm en moins d’une heure) et ne variant quasiment pas puis s’arrêtant brutalement aux alentours de 23h au moment du réétalonnage de l’appareil
- Les “lignes de 0” (une succession de 0 dans la colonne de concentration de CO) pendant une période de 5h ou plus et s’arrêtant vers 23h (réétalonnage de l’appareil)

On relève toutes ces valeurs puis on les supprime à l’aide de MATHLAB, car il est absurde d’avoir pendant, parfois plus d’une journée, une concentration de CO quasi nulle en plein paris ou alors un grand pic brutalement pendant plusieurs heures ne variant quasiment pas (comme si soit quelqu’un fumait à côté, soit la machine est restée bloquée à cette valeur pour une raison inconnue). Ces valeurs ne correspondaient pas à des situations physiques réalistes.

Ce deuxième tri a retiré 1027 mesures soit environ 3% des données d'origines. Notre analyse se portera donc sur 31565 mesures sur 34561, on a perdu environs 8.7% de nos données d'origine.

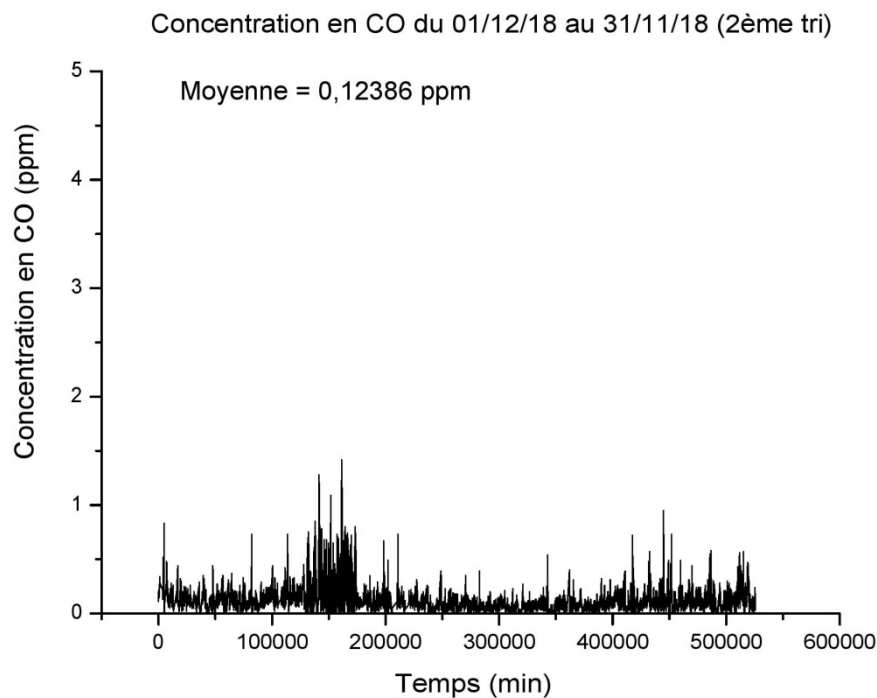


Figure 2 : Graphique de la concentration de CO en fonction du temps sur un an (2ème tri)

## **Analyse des données**

A partir de multiples recherches bibliographiques, on a obtenu les trois causes principales de la variabilité en CO dans l'atmosphère :

- **Couche limite atmosphérique** : La couche limite atmosphérique est définie comme la partie de la troposphère directement soumise à l'influence de la surface terrestre à l'échelle de la journée. Elle se situe en moyenne à 1500m d'altitude, avec une grande variabilité entre le jour et la nuit, et l'été et l'hiver. Les principaux facteurs de sa variabilité sont dynamiques et thermiques, c'est-à-dire le vent et la température du sol. En hiver, lorsque la température baisse, la couche limite se rétracte et donc le CO a un volume de dilution moins grand et à l'inverse en été, la couche limite se dilate et la concentration en CO diminue alors car il y a un volume de dilution plus grand.
- **Photochimie** : Un autre phénomène thermique entre en jeu, cette fois directement liée à l'énergie rayonnée par le Soleil. De nombreux phénomènes photochimiques se produisent dans l'atmosphère, comme la combinaison des molécules d'OH avec les molécules de CO en de nouvelles molécules, permettant donc une baisse de la concentration en CO dans l'atmosphère. Or les molécules d'OH apparaissent plus fréquemment lorsque la température est plus élevée, renforçant le rôle de la température comme inhibiteur de l'augmentation de la concentration en CO de l'atmosphère.

- Pollution humaine : Enfin dernier phénomène majeur influençant la concentration en CO dans l'atmosphère, les émissions de CO humaines. Une analyse sommaire des pics et creux de l'évolution de la concentration en CO permet de d'obtenir une certaine corrélation entre la variabilité de l'activité humaine et la part de CO dans l'atmosphère. En effet, le creux en été peut être lié aux départs en vacances et les pics en hiver à l'utilisation croissante de chauffage.

Le graphique de la concentration de CO en fonction du temps à partir des données brutes triées (figure 2) nous permet d'y voir un peu plus clair : il semble y avoir une corrélation entre saison et concentration de CO, au milieu du graphique il semble y avoir un creux qui correspondrait à la période de l'été. Cependant la base horaire choisie (minute) ne permet pas d'affirmer clairement cela du fait des nombreuses variations. On va donc recourir à de nouvelles bases horaires.

## Evolution annuelle de la concentration en CO

L'évolution annuelle de la concentration en CO présente la particularité de prendre des valeurs minimales en été et maximales en hiver. On décide donc grâce au logiciel MATHLAB de faire la moyenne de la concentration en CO sur chaque semaine de l'année puis ensuite tracer le graphique associé.

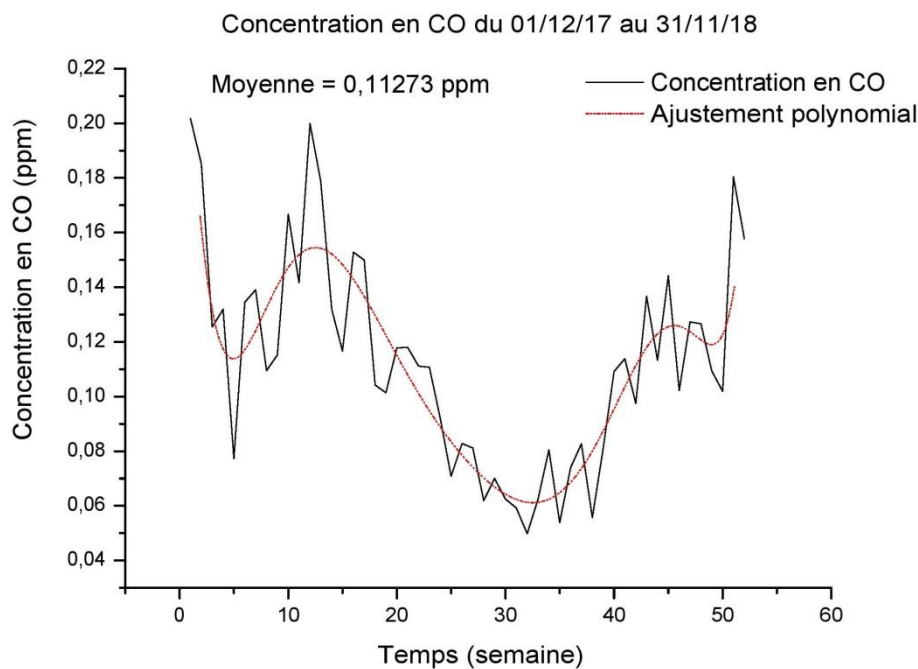


Figure 3 : Graphique de la concentration de CO en fonction du temps sur un an (base hebdomadaire)

On observe bien, comme on s'y attendait, un creux marqué de la semaine 20 à la semaine 43, ce qui correspondrait à la date du 18 Avril 2018 au 26 Septembre 2018 qui englobe toute la période estivale. Cela est aussi en accord avec la théorie et les sources bibliographiques, qui affirment qu'avec la hausse des températures, la concentration en CO doit diminuer que ce soit dû à un effet photochimique ou au plus grand volume de dilution du gaz. Mais aussi avec l'activité humaine, qui, sur Paris, est moins intense due aux périodes de vacances et à l'utilisation réduite de chauffage par exemple.

## Evolution journalière de la concentration en CO

Avec la base hebdomadaire l'influence de la température et des saisons sur la concentration en CO est plutôt évidente, cependant, dû aux nombreux facteurs influençant l'activité humaine, il est difficile d'affirmer précisément quelle est son poids sur la modification de la concentration de CO.

On a décidé que pour mettre en évidence le facteur de la pollution humaine il serait astucieux de s'intéresser à la variation journalière de la concentration de CO. Nous avons donc tracé plusieurs jours au hasard et nous avons remarqué que durant les jours de la semaine il y avait globalement 2 pics : un le matin et un le soir tandis que le week-end les pics étaient très dilués, voir quasi-inexistants. Nous avons avancé que les 2 pics des jours de la semaine étaient sûrement dû aux gens partants et revenants du travail.

Nous avons donc dû compléter le programme pour qu'il puisse créer de nouvelles bases de données, cette fois-ci où les valeurs sont moyennées sur une base horaire selon le jour de la semaine auxquelles elles appartiennent, créant ainsi des modèles de journées moyennes. Nous avons donc créé ce nouveau fichier avec la totalité de nos données, puis en séparant les intervalles du 01/12/17 au 31/03/18 et du 27/10/18 au 31/11/18, et du 31/03/18 au 27/10/18, correspondants respectivement aux heures d'hiver et d'été.

En étudiant l'évolution de la concentration en CO d'une journée moyenne, particulièrement en dehors du week-end, on observe un pic bien défini le matin et un autre le soir un peu plus diffus, correspondants aux pics de densité du trafic routier.

Pour vérifier cela, nous avons utilisé le critère de changement d'heure, qui n'a d'influence que sur l'activité humaine et non sur le climat.

En effet, le changement d'heure oblige les gens qui travaillent à partir une heure plus tôt au travail en hiver. Ce qui signifie que si on trace le même graphique journalier en hiver et en été et que le pic du graphique "heure d'hiver" se trouve vers 9h (intensité maximale de la concentration en CO), on s'attend donc à ce que le pic du graphique "heure d'été" se trouve vers 10h, soit décalé d'une heure. Si l'activité humaine a une influence prépondérante sur la concentration en CO, alors un décalage des pics d'activités humaines devrait entraîner un décalage des pics de concentration en CO.

On effectue un ajustement polynomial afin de voir plus clairement le centre du pic et ainsi comparer les graphiques avec plus de précision.

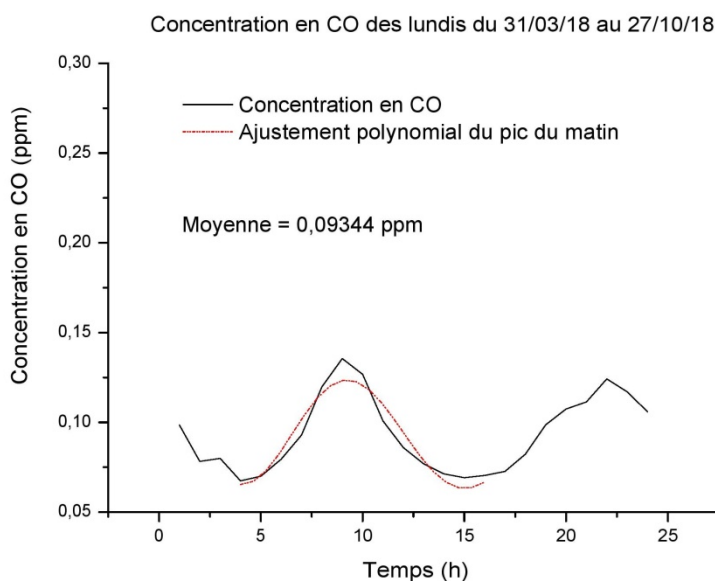


Figure 4: Graphique de la concentration de Co en fonction du temps (moyenne des lundis à l'heure d'été)

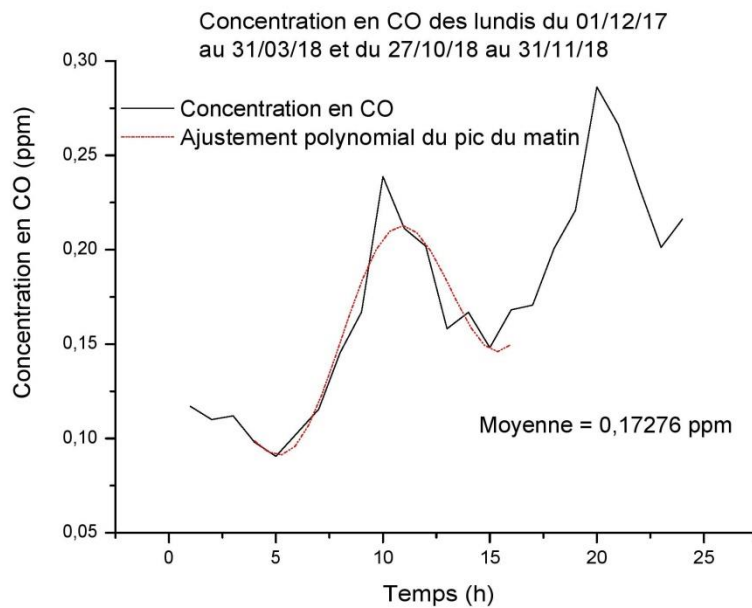


Figure 5 : Graphique de la concentration de Co en fonction du temps (moyenne des lundis à l'heure d'hiver)

On observe que les pics du matin sont très clairement décalés d'une heure, or le rythme de travail reste le même en été comme en hiver, on peut donc en déduire que ce décalage est dû à l'activité humaine qui a été décalée à cause du décalage horaire. Les variabilités journalières de CO sont donc en grande partie causée par l'activité humaine.

## **Conclusion**

Durant ce stage on a mis en évidence, à partir d'un relevé de données sur un an, plusieurs facteurs influençant la concentration du monoxyde de carbone dans l'atmosphère comme l'activité humaine ou la variabilité saisonnière. Cependant dû aux grands nombres de facteurs influençant l'environnement, on pourrait mettre en place de meilleurs moyens de déduire les différents facteurs mais aussi leur degré d'influence. Une étude sur une plus grande période de temps de la variabilité saisonnière nous aurait permis d'observer la tendance globale de l'évolution en CO, ce qui aurait pu être particulièrement intéressant afin d'analyser l'efficacité des récentes politiques de régulations le concernant. Enfin on aurait aussi pu étudier plus en détail les grands pics de pollution, ce qui permettrait d'émettre des conjectures quant à leurs conditions d'apparition, et si des conditions propices sont reconnues quel protocole mettre en œuvre pour éviter l'apparition de tels phénomènes.