

Concentration de CO2 dans l'atmosphère depuis 1958

Oriane Baussens

18/08/2020

Importation et vérification des données

Lecture des données le 18/08/2020 depuis le site du Scripps :

```
data=read.csv("https://scrippsco2.ucsd.edu/assets/data/atmospheric/stations/in_situ_co2/monthly/monthly")
```

La commande na.strings remplace les valeurs manquantes (qui apparaissent dans le fichier comme -99.99) par NA.

L'observation montre que les deux premières lignes correspondent à des commentaires sur le type de données et leur unité. Ces informations peuvent être intégrées aux titres des colonnes et enlevées des données.

```
names(data)=c("Yr", "Mn", "Date", "Date.1", "CO2 [ppm]", "seasonally (adjusted) [ppm]", "fit [ppm]", "sea")
data=data[3:758,]
```

Suppression des données manquantes :

```
ligne.na=apply(data, 1, function(x) any(is.na(x)))
data=data[!ligne.na,]
```

Vérification de la class de données :

```
class(data$`CO2 [ppm]`)
```

```
## [1] "character"
```

```
data$`CO2 [ppm]`=as.double(data$`CO2 [ppm]`)
class(data$`CO2 [ppm]`)
```

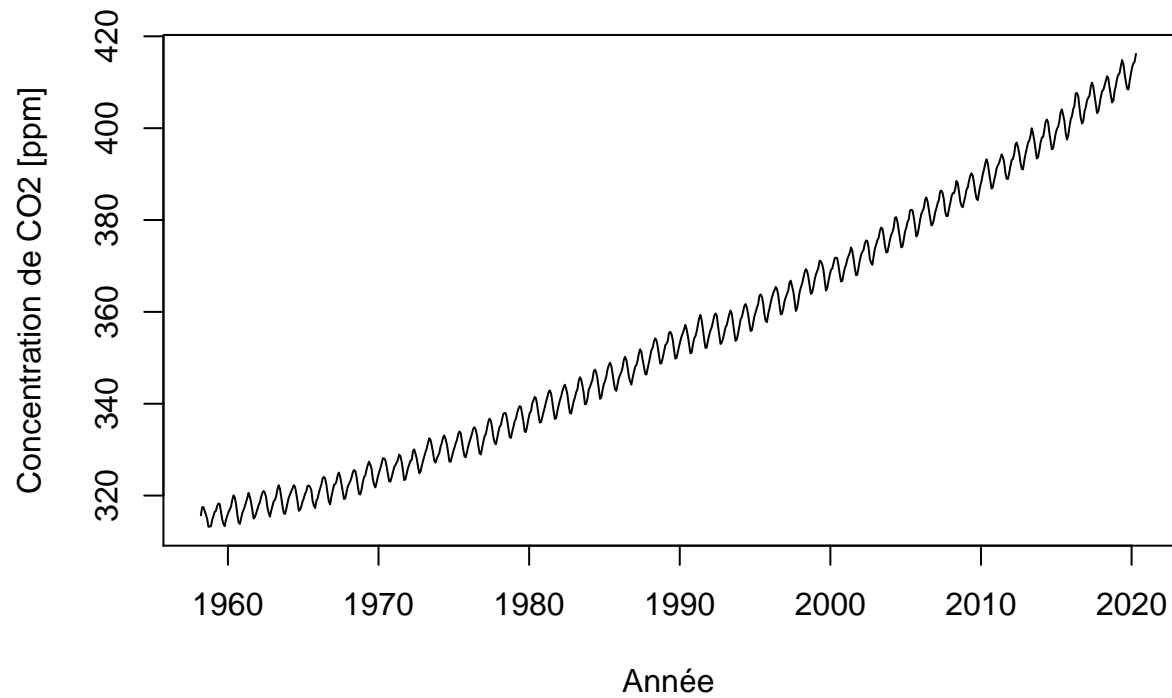
```
## [1] "numeric"
```

Les données de concentration de CO2 étaient du type caractère. Le code ci-dessus les converti en doubles.

Analyse des données

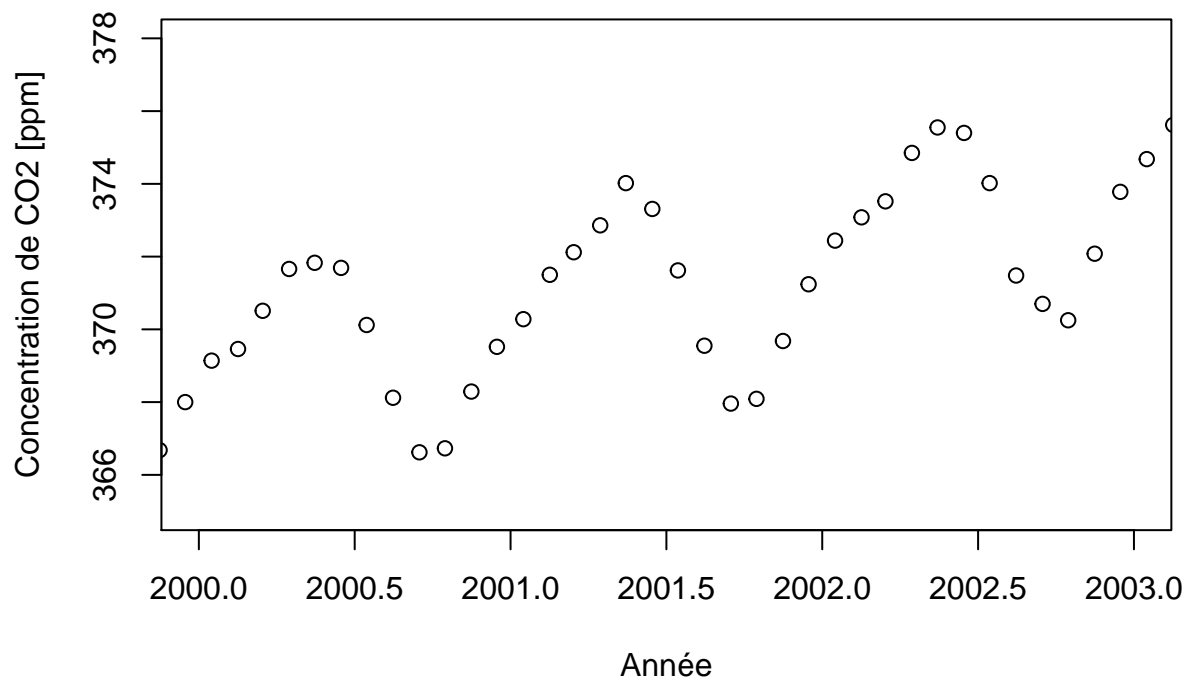
Affichage des données :

```
plot(x=data$Date.1, y=data$`C02 [ppm]`, type='l', xlab="Année", ylab="Concentration de C02 [ppm]")
```



Observation plus précise de l'oscillations périodiques :

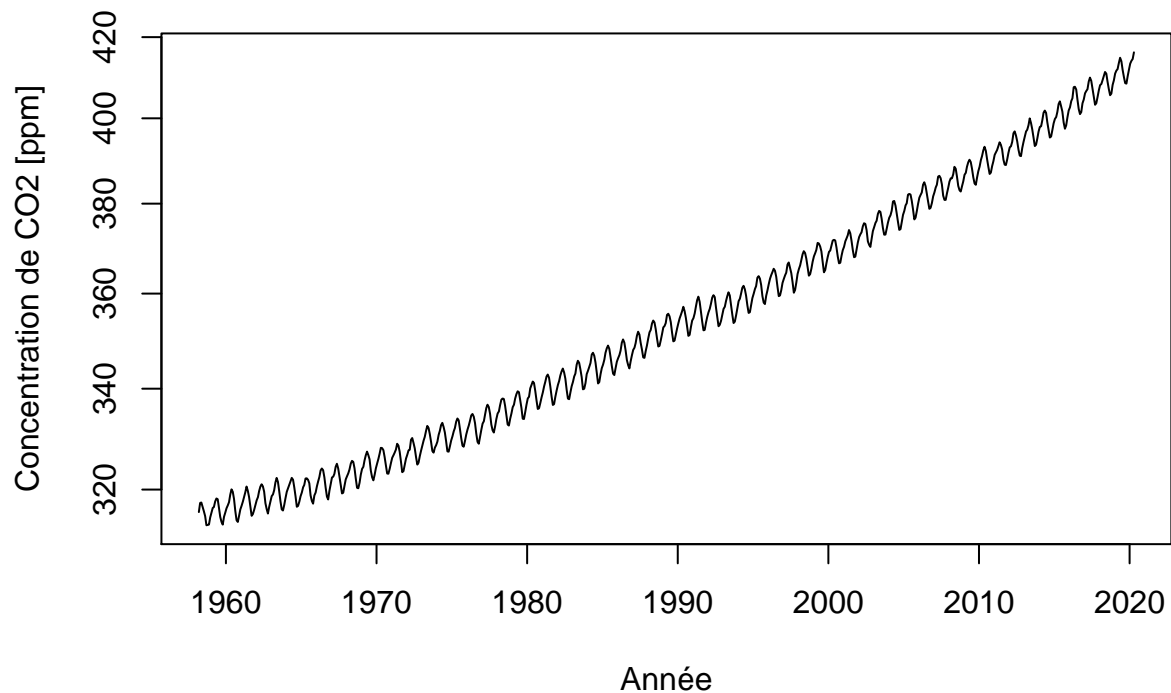
```
plot(x=data$Date.1, y=data$`C02 [ppm]`, type='p', xlab="Année", ylab="Concentration de C02 [ppm]", xlim=c(1960, 2020))
```



L'observation des oscillations périodiques montre que la période correspondant aux mois de septembre de d'octobre présente toujours la concentration de CO2 la plus faible. Celle correspondant aux mois d'avril à juin présente la concentration de CO2 la plus élevée.

L'évolution globale des données semble suivre une loi exponentielle. Pour vérifier cela, affichons la concentration de CO2 en échelle logarithmique :

```
plot(x=data$Date.1, y=data$`CO2 [ppm]`, type='l', xlab="Année", ylab="Concentration de CO2 [ppm]", log=
```



Le résultat est bien une droite. Ajustons maintenant les données :

#Equation : $Y=aX+b$ avec X les années et Y la concentration de CO2 estimée

```
linear_fit=lm(log(data$`CO2 [ppm]`)-data$Date.1, data= data)
```

```
b=linear_fit$coefficients[1]
```

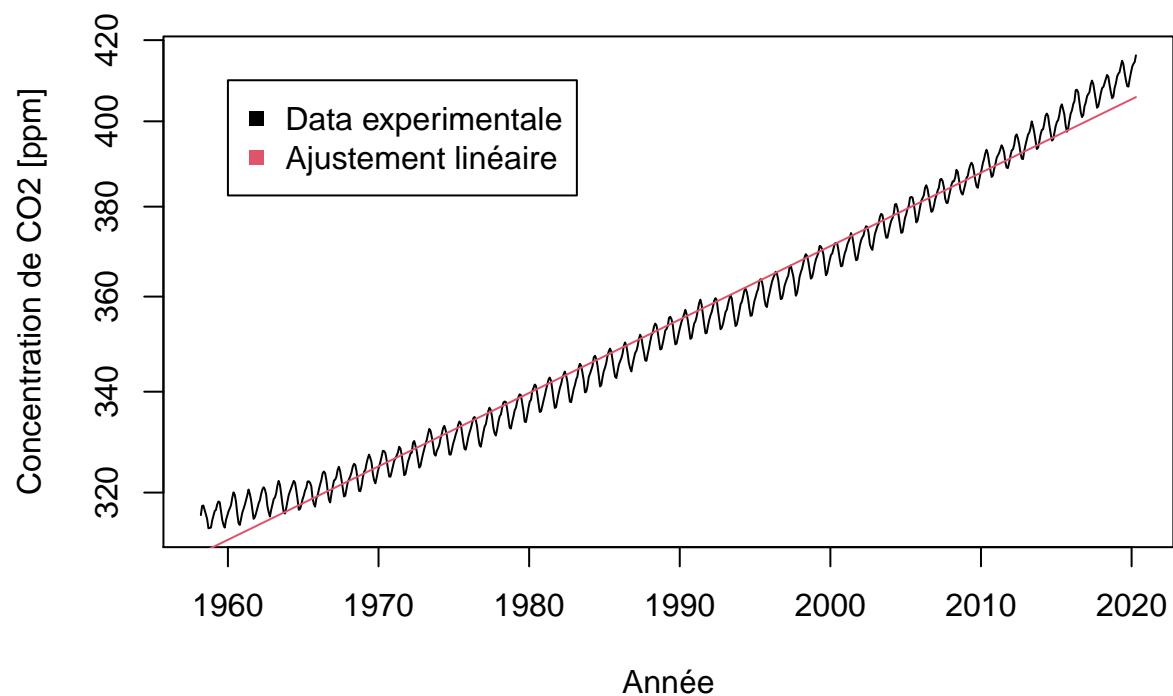
```
a=linear_fit$coefficients[2]
```

```
Y=exp(a*data$Date.1+b)
```

```
plot(x=data$Date.1, y=data$`CO2 [ppm]`, type='l', xlab="Année", ylab="Concentration de CO2 [ppm]", log=
```

```
lines(data$Date.1, Y, col=2)
```

```
legend(1960, 410, c("Data expérimentale", "Ajustement linéaire"), pch=15, col=1:2)
```



Extrapolation des données jusqu'en 2025 :

```
Co2_2025=exp(a*2025+b)
```

```
Co2_2025
```

```
## data$Date.1
```

```
## 414.3853
```

D'après la modélisation, la concentration de CO2 dans l'atmosphère devrait être de 414ppm en 2025.