

Sujet 1 : Concentration de CO2 dans l'atmosphère depuis 1958

Alicia CULOT

2025-10-15

Introduction

Depuis 1958, l'observatoire de Mauna Loa, à Hawaï, mesure en continu la concentration de dioxyde de carbone (CO₂) dans l'atmosphère. Ces mesures, initiées par Charles David Keeling, ont permis de mettre en évidence une augmentation continue de la concentration atmosphérique de CO₂, connue sous le nom de **courbe de Keeling**. Cette série de données constitue une référence mondiale pour l'étude du changement climatique.

L'objectif de cette étude est de caractériser la **dynamique temporelle** de la concentration de CO₂ depuis 1958 jusqu'à aujourd'hui, à partir des mesures hebdomadaires fournies par le *Scripps Institution of Oceanography*.

Préparation des données

Chargement du package ggplot2

```
library(ggplot2)
```

Chargement URL

Chargement des données hebdomadaires de mesures de CO2 à la station d'observation de Mauna Loa.

Date de téléchargement: 15 Octobre 2025

```
data_url =  
"https://scrippsco2.ucsd.edu/assets/data/atmospheric/stations/in_situ_co2/weekly/weekly_in_situ_co2_mlo.csv"
```

Téléchargement des données

Sélection des données d'intérêt

Le code suivant permet de sauter les métadonnées et d'extraire seulement les dates (première colonne) et les concentrations en CO2 (ppm; deuxième colonne):

```
data_raw <- read.csv(  
  data_url,
```

```

skip = 44,
header = FALSE,
sep = ",",
stringsAsFactors = FALSE,
na.strings = c("-99.99", "-1", "NA", "")
)

```

Création dataframe

Les noms sont ajoutés aux colonnes:

```

data <- data_raw[, 1:2]
names(data) <- c("Date", "CO2")

```

Conversion

Modification du format: yyyy-mm-dd pour les dates et numérique pour les concentrations en CO2

```

data$Date <- as.Date(trimws(data$Date), format = "%Y-%m-%d")
data$CO2 <- as.numeric(data$CO2)

```

Affichage et vérifications des données

```
str(data)
```

```

## 'data.frame':    3440 obs. of  2 variables:
## $ Date: Date, format: "1958-03-29" "1958-04-05" ...
## $ CO2 : num  316 317 318 318 316 ...

```

```
head(data)
```

```

##      Date      CO2
## 1 1958-03-29 316.19
## 2 1958-04-05 317.31
## 3 1958-04-12 317.69
## 4 1958-04-19 317.58
## 5 1958-04-26 316.48
## 6 1958-05-03 316.95

```

Graphique

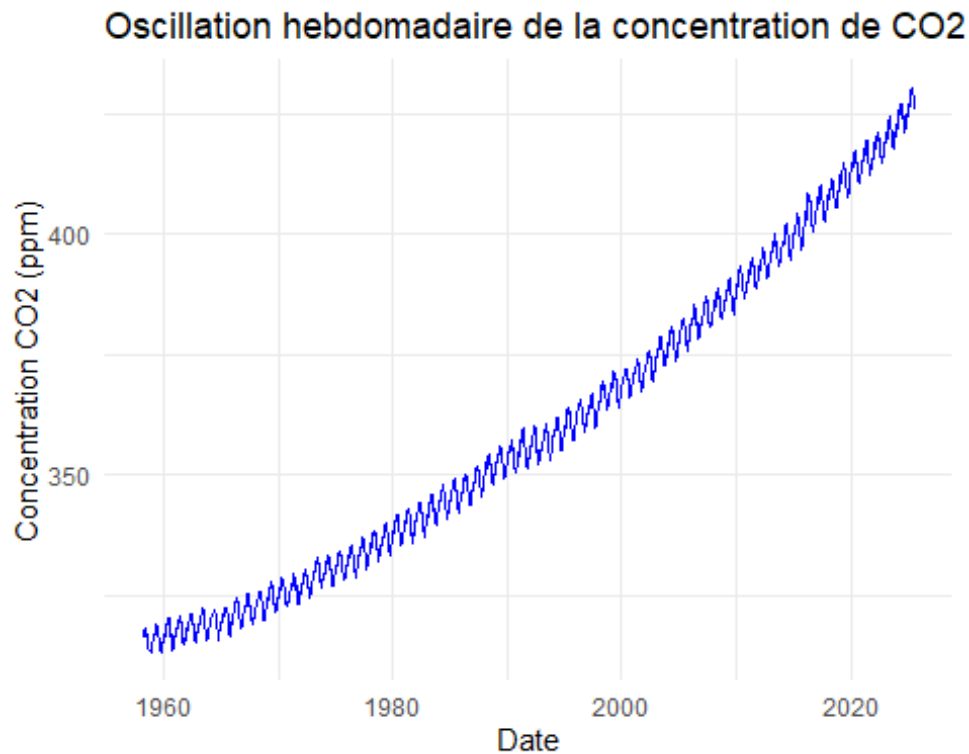
Création du graphique de l'oscillation hebdomadaire de la concentration de CO₂ à Mauna Loa

```

ggplot(data, aes(x = Date, y = CO2)) +
  geom_line(color = "blue") +
  labs(
    title = "Oscillation hebdomadaire de la concentration de CO2 à Mauna Loa
à partir de 1958",
    x = "Date",

```

```
y = "Concentration CO2 (ppm)"  
) +  
theme_minimal()
```



Le graphe montre une séparation de deux phénomènes:

- **Oscillation périodique:** les variations rapides et régulières de CO₂ dues aux saisons, visibles comme les « dents de scie » sur la courbe.
- **Tendance lente:** l'augmentation générale du CO₂ au fil des années, plus régulière et progressive.

Séparation des deux phénomènes

Séparer les deux phénomènes permettra de:

- caractériser la période et l'amplitude de l'oscillation annuelle
- modéliser correctement la tendance à long terme
- faire des prévisions futures.

Conversion de la série en série temporelle hebdomadaire

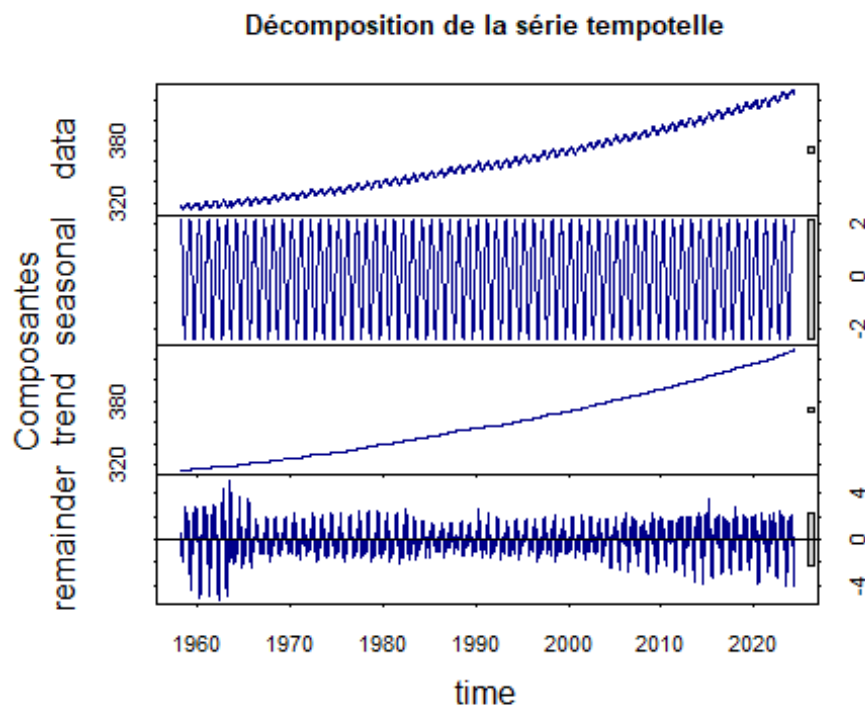
Une fréquence = 52 semaines

```
co2_ts <- ts(data$CO2, start = c(1958, 12), frequency = 52)
```

Décomposition de série temporelle

Utilisation de la fonction `stl()` pour séparer l'oscillation périodique de la tendance lente:

```
co2_stl <- stl(co2_ts, s.window = "periodic")
plot(co2_stl, main = "Décomposition de la série temporelle", col = "dark blue")
title(ylab = "Composantes")
```



Les trois composantes: la tendance lente (*trend*), l'oscillation saisonnière (*seasonal*) et les résidus (*remainder*) vont permettre de caractériser l'oscillation et modéliser la tendance.

Extraction des 3 composantes

```
co2_trend <- co2_stl$time.series[, "trend"]
co2_seasonal <- co2_stl$time.series[, "seasonal"]
co2_remainder <- co2_stl$time.series[, "remainder"]
```

Caractérisation de l'oscillation périodique

Amplitude : différence max - min de la saisonnalité:

```
amplitude <- max(co2_seasonal) - min(co2_seasonal)
cat("Amplitude de l'oscillation annuelle (ppm) :", round(amplitude,2), "\n")
## Amplitude de l'oscillation annuelle (ppm) : 4.63
```

Proposition de modèle simple de la contribution lente

Le modèle linéaire simple est proposé:

$$CO2 = a + bT$$

Avec $CO2$ = concentration en CO_2 (ppm) et T = temps (semaines)

Mise en place du modèle:

```
time <- time(co2_ts)
trend_model <- lm(co2_trend ~ time)
```

Estimation des paramètres:

```
summary(trend_model)

##
## Call:
## lm(formula = co2_trend ~ time)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.897 -3.527 -1.861  3.426 12.288
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.971e+03  8.042e+00  -369.4   <2e-16 ***
## time         1.673e+00  4.039e-03   414.3   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.523 on 3438 degrees of freedom
## Multiple R-squared:  0.9804, Adjusted R-squared:  0.9804
## F-statistic: 1.717e+05 on 1 and 3438 DF, p-value: < 2.2e-16
```

Intercept(a) = Valeur théorique de CO_2 au quand $T = 0$

penste b = 1.673 ppm/semaine

p-value < 2.2e-16 : très significatif

Multiple R-squared = 98.04% de la variance de la tendance est expliquée par ce modèle linéaire.

Extrapolation jusqu'en 2026

Le code suivant permet de faire une extrapolation du modèle linéaire de tendance jusqu'en 2026:

```
time_future <- seq(start(co2_ts)[1], 2026, by = 1/52) # Vecteur de temps
allant du début de ta série (1958) jusqu'à 2026, avec un pas de 1/52
```

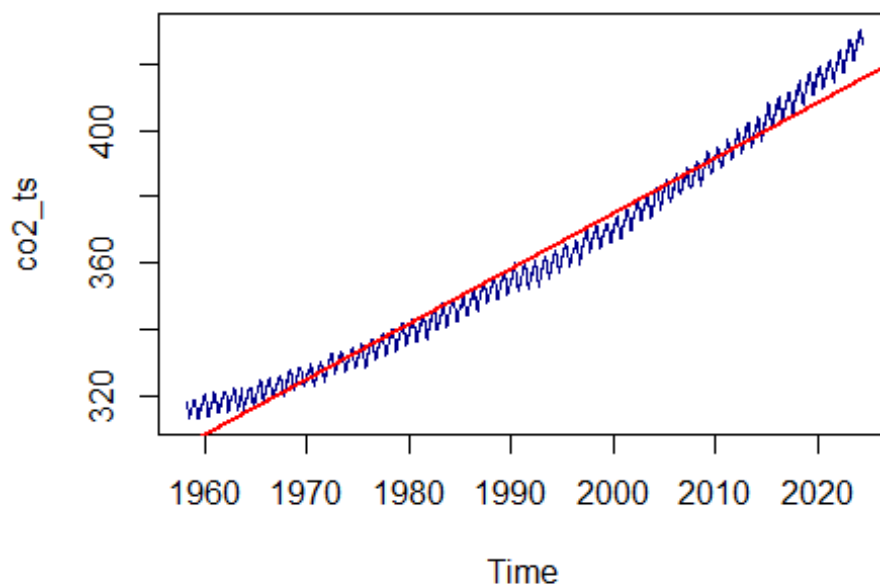
```
trend_pred <- predict(trend_model, newdata = data.frame(time = time_future))
# Utilise le modèle linéaire pour prédire la concentration de CO2 à chaque
valeur de time_future.

trend_pred_ts <- ts(trend_pred, start = start(co2_ts), frequency = 52) #
Convertit ces valeurs extrapolées en une série temporelle complète (ts)
compatible avec la série d'origine
```

Cela crée une série temporelle exploitable pour les graphiques ou l'analyse et de tracer la tendance extrapolée:

```
plot(co2_ts, col = "darkblue", main = "Modèle linéaire de la concentration en
CO2 en fonction du temps")
lines(trend_pred_ts, col = "red", lwd = 2)
```

Modèle linéaire de la concentration en CO2 en fonction du temps



Exemple d'extrapolation

Par exemple, si l'on souhaite estimer la concentration (ppm) à une semaine donnée, on peut réaliser ce code en changeant le mois:

```
time_feb2026 <- 2026.08
CO2_feb2026 <- predict(trend_model, newdata = data.frame(time =
time_feb2026))
cat("Estimation de la concentration en CO2 en février 2026 (ppm) :",
CO2_feb2026, "\n")

## Estimation de la concentration en CO2 en février 2026 (ppm) : 418.8923
```

Il est donc possible, avec ce modèle linéaire simple, de pouvoir estimer les concentrations futures.