concentration_co2

September 4, 2024

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

[1]:
```
%matplotlib inline
import matplotlib.pyplot as plt
import pandas as pd
```

On récupère les données. Elles ont été trouvées dans ce lien.

C. D. Keeling, S. C. Piper, R. B. Bacastow, M. Wahlen, T. P. Whorf, M. Heimann, and H. A. Meijer, Exchanges of atmospheric CO2 and 13CO2 with the terrestrial biosphere and oceans from 1978 to 2000. I. Global aspects, SIO Reference Series, No. 01-06, Scripps Institution of Oceanography, San Diego, 88 pages, 2001.

[2]:
```
data_file = 'monthly_in_situ_co2_mlo.csv'
data = pd.read_csv(data_file, skiprows=61)
data
```

[2]:

	Yr	Mn	Date	Date	CO2	seasonally adjusted	fit	\
0								
1			Excel		[ppm]	[ppm]	[ppm]	
2	1958	01	21200	1958.0411	-99.99	-99.99	-99.99	
3	1958	02	21231	1958.1260	-99.99	-99.99	-99.99	
4	1958	03	21259	1958.2027	315.71	314.43	316.21	
..	…	…	…	…	…	…	…	
801	2024	08	45519	2024.6230	-99.99	-99.99	-99.99	
802	2024	09	45550	2024.7077	-99.99	-99.99	-99.99	
803	2024	10	45580	2024.7896	-99.99	-99.99	-99.99	
804	2024	11	45611	2024.8743	-99.99	-99.99	-99.99	
805	2024	12	45641	2024.9563	-99.99	-99.99	-99.99	

	seasonally adjusted fit	CO2 filled	seasonally adjusted filled	Sta
0				NaN
1	[ppm]	[ppm]	[ppm]	NaN
2	-99.99	-99.99	-99.99	MLO
3	-99.99	-99.99	-99.99	MLO
4	314.91	315.71	314.43	MLO
..	…	…	… …	
801	-99.99	-99.99	-99.99	MLO
802	-99.99	-99.99	-99.99	MLO

```
803        -99.99        -99.99           -99.99   MLO
804        -99.99        -99.99           -99.99   MLO
805        -99.99        -99.99           -99.99   MLO

[806 rows x 11 columns]
```

Il y a un problème avec les en têtes. On concatene les deux premieres lignes et enleve celle de l'unité.

[3]:
```
columns_label = ['Yr', 'Mn', 'Date_excel', 'Date', 'CO2', 'seasonally_adjusted',
       'fit', 'seasonally_adjusted_fit', 'CO2_filled',␣
 ↪'seasonally_adjusted_filled', 'Sta']

data.columns = columns_label
data = data.drop([0, 1]).reset_index().drop('index',axis=1)
data
```

[3]:
```
        Yr   Mn Date_excel         Date       CO2 seasonally_adjusted  \
0     1958   01      21200    1958.0411    -99.99              -99.99
1     1958   02      21231    1958.1260    -99.99              -99.99
2     1958   03      21259    1958.2027    315.71              314.43
3     1958   04      21290    1958.2877    317.45              315.16
4     1958   05      21320    1958.3699    317.51              314.69
..     …   …         …           …        …                   …
799   2024   08      45519    2024.6230    -99.99              -99.99
800   2024   09      45550    2024.7077    -99.99              -99.99
801   2024   10      45580    2024.7896    -99.99              -99.99
802   2024   11      45611    2024.8743    -99.99              -99.99
803   2024   12      45641    2024.9563    -99.99              -99.99

              fit seasonally_adjusted_fit   CO2_filled  \
0         -99.99                   -99.99       -99.99
1         -99.99                   -99.99       -99.99
2         316.21                   314.91       315.71
3         317.30                   314.99       317.45
4         317.89                   315.07       317.51
..           …                        …          …
799       -99.99                   -99.99       -99.99
800       -99.99                   -99.99       -99.99
801       -99.99                   -99.99       -99.99
802       -99.99                   -99.99       -99.99
803       -99.99                   -99.99       -99.99

    seasonally_adjusted_filled   Sta
0                       -99.99   MLO
1                       -99.99   MLO
2                       314.43   MLO
3                       315.16   MLO
```

```
4                       314.69   MLO
..                         …    …
799                     -99.99   MLO
800                     -99.99   MLO
801                     -99.99   MLO
802                     -99.99   MLO
803                     -99.99   MLO

[804 rows x 11 columns]
```

Les données manquantes dans le fichier de base ont été remplacées par -99.99. Elles ne nous interessent pas. On les enlève de l'analyse.

[4]:
```
data = data[data['CO2'].astype(float)>0]
data
```

[4]:
```
        Yr   Mn Date_excel         Date       CO2 seasonally_adjusted  \
2     1958   03      21259    1958.2027    315.71               314.43
3     1958   04      21290    1958.2877    317.45               315.16
4     1958   05      21320    1958.3699    317.51               314.69
6     1958   07      21381    1958.5370    315.87               315.20
7     1958   08      21412    1958.6219    314.93               316.22
..     …   …         …           …          …                    …
794   2024   03      45366    2024.2049    425.22               423.65
795   2024   04      45397    2024.2896    426.30               423.50
796   2024   05      45427    2024.3716    426.70               423.29
797   2024   06      45458    2024.4563    426.62               424.06
798   2024   07      45488    2024.5383    425.40               424.62

              fit seasonally_adjusted_fit   CO2_filled  \
2          316.21                   314.91       315.71
3          317.30                   314.99       317.45
4          317.89                   315.07       317.51
6          315.86                   315.22       315.87
7          313.97                   315.29       314.93
..            …                        …            …
794        424.92                   423.33       425.22
795        426.42                   423.60       426.30
796        427.27                   423.86       426.70
797        426.67                   424.13       426.62
798        425.13                   424.39       425.40

    seasonally_adjusted_filled   Sta
2                       314.43   MLO
3                       315.16   MLO
4                       314.69   MLO
6                       315.20   MLO
7                       316.22   MLO
```

```
..                          …   …
794                     423.65   MLO
795                     423.50   MLO
796                     423.29   MLO
797                     424.06   MLO
798                     424.62   MLO

[792 rows x 11 columns]
```

On crée une collonne avec les dates compreensibles par pandas

[5]:
```
def convert_to_month_start(year, month):
    return pd.Timestamp(year=year, month=month, day=1)

data.loc[:, 'period'] = [convert_to_month_start(y, m) for y, m in␣
 ↪zip(data['Yr'].astype(int), data['Mn'].astype(int))]
data
```

```
/tmp/ipykernel_33641/3633217335.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data.loc[:, 'period'] = [convert_to_month_start(y, m) for y, m in
zip(data['Yr'].astype(int), data['Mn'].astype(int))]
```

[5]:
```
        Yr   Mn Date_excel         Date       CO2 seasonally_adjusted  \
2     1958   03      21259    1958.2027    315.71              314.43
3     1958   04      21290    1958.2877    317.45              315.16
4     1958   05      21320    1958.3699    317.51              314.69
6     1958   07      21381    1958.5370    315.87              315.20
7     1958   08      21412    1958.6219    314.93              316.22
..     …  …        …           …       …                   …
794   2024   03      45366    2024.2049    425.22              423.65
795   2024   04      45397    2024.2896    426.30              423.50
796   2024   05      45427    2024.3716    426.70              423.29
797   2024   06      45458    2024.4563    426.62              424.06
798   2024   07      45488    2024.5383    425.40              424.62

              fit seasonally_adjusted_fit   CO2_filled  \
2          316.21                   314.91       315.71
3          317.30                   314.99       317.45
4          317.89                   315.07       317.51
6          315.86                   315.22       315.87
7          313.97                   315.29       314.93
..            …                        …          …
794        424.92                   423.33       425.22
```

```
795       426.42                  423.60       426.30
796       427.27                  423.86       426.70
797       426.67                  424.13       426.62
798       425.13                  424.39       425.40

    seasonally_adjusted_filled   Sta     period
2                       314.43   MLO 1958-03-01
3                       315.16   MLO 1958-04-01
4                       314.69   MLO 1958-05-01
6                       315.20   MLO 1958-07-01
7                       316.22   MLO 1958-08-01
..                         …   …        …
794                     423.65   MLO 2024-03-01
795                     423.50   MLO 2024-04-01
796                     423.29   MLO 2024-05-01
797                     424.06   MLO 2024-06-01
798                     424.62   MLO 2024-07-01

[792 rows x 12 columns]
```

On fait le plot de la tendance historique

On utilise les periodes en tant qu'index et on les sorte de manière croissante

[6]:
```
sorted_data = data.set_index('period').sort_index()
sorted_data
```

[6]:
```
                Yr   Mn Date_excel         Date       CO2 seasonally_adjusted  \
period
1958-03-01  1958   03      21259    1958.2027    315.71              314.43
1958-04-01  1958   04      21290    1958.2877    317.45              315.16
1958-05-01  1958   05      21320    1958.3699    317.51              314.69
1958-07-01  1958   07      21381    1958.5370    315.87              315.20
1958-08-01  1958   08      21412    1958.6219    314.93              316.22
…           …  …        …           …          …                   …
2024-03-01  2024   03      45366    2024.2049    425.22              423.65
2024-04-01  2024   04      45397    2024.2896    426.30              423.50
2024-05-01  2024   05      45427    2024.3716    426.70              423.29
2024-06-01  2024   06      45458    2024.4563    426.62              424.06
2024-07-01  2024   07      45488    2024.5383    425.40              424.62

                    fit seasonally_adjusted_fit   CO2_filled  \
period
1958-03-01       316.21                  314.91       315.71
1958-04-01       317.30                  314.99       317.45
1958-05-01       317.89                  315.07       317.51
1958-07-01       315.86                  315.22       315.87
1958-08-01       313.97                  315.29       314.93
```

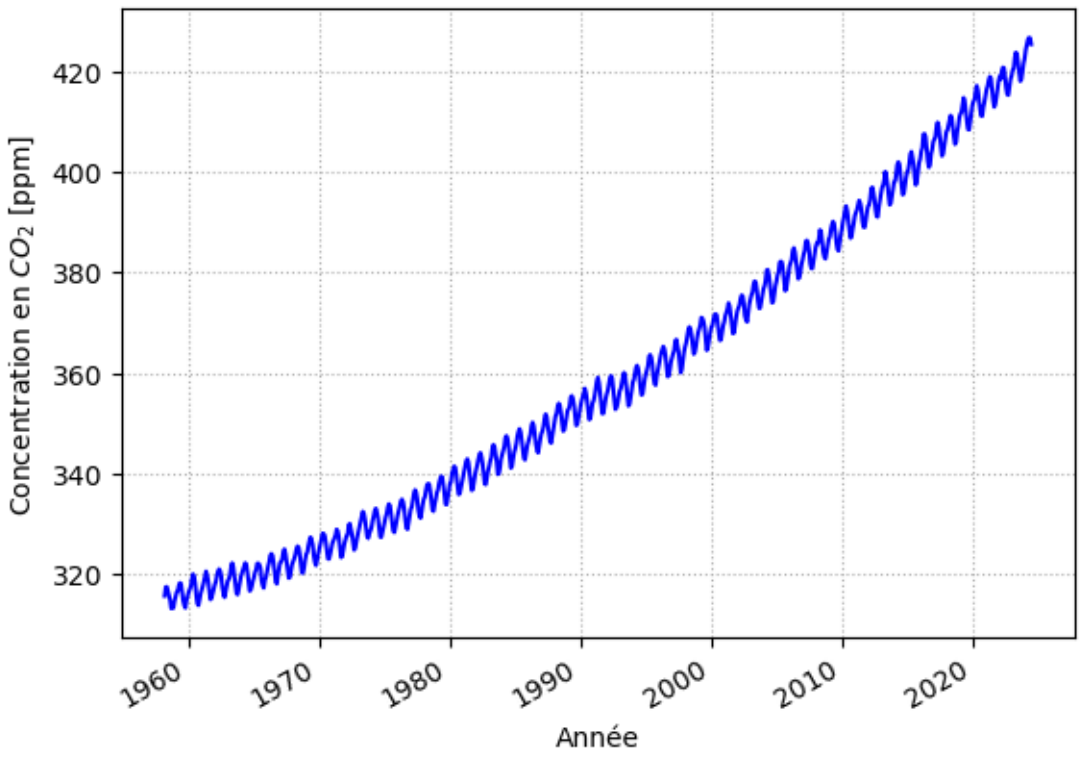
```
…                 …                    …          …
2024-03-01       424.92                423.33       425.22
2024-04-01       426.42                423.60       426.30
2024-05-01       427.27                423.86       426.70
2024-06-01       426.67                424.13       426.62
2024-07-01       425.13                424.39       425.40

           seasonally_adjusted_filled   Sta
period
1958-03-01                      314.43   MLO
1958-04-01                      315.16   MLO
1958-05-01                      314.69   MLO
1958-07-01                      315.20   MLO
1958-08-01                      316.22   MLO
…                                …   …
2024-03-01                      423.65   MLO
2024-04-01                      423.50   MLO
2024-05-01                      423.29   MLO
2024-06-01                      424.06   MLO
2024-07-01                      424.62   MLO

[792 rows x 11 columns]
```

On plot les données

[7]:
```
sorted_data['CO2'] = sorted_data['CO2'].astype(float)
sorted_data['CO2'].plot(color='blue')
plt.ylabel(r'Concentration en $CO_2$ [ppm]')
plt.xlabel('Année')
plt.grid(linestyle=':')
```



On observe deux phénomène couplés : une oscillation périodique annuelle et une contribution plus lente. On essaye de fitter cette contribution plus lente par une exponentielle croissante, selon

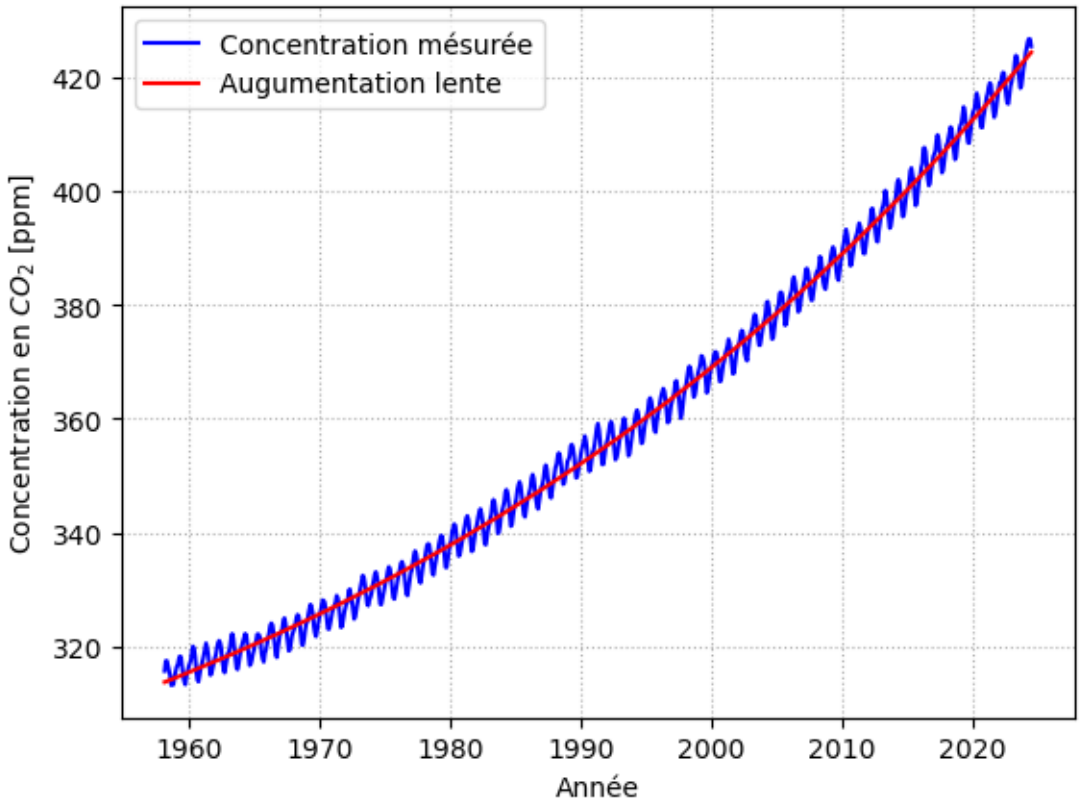
$$[CO_2] = a \cdot \exp(b \cdot t) + c$$

[8]:
```
import numpy as np
from scipy.optimize import curve_fit

def exponential_func(t, a, b, c):
    return a * np.exp(b * t) + c

start_year = sorted_data.index.min().year
sorted_data["period_fractional"] = (sorted_data.index.year - start_year) +␣
 ↪(sorted_data.index.month - 1) / 12
initial_guess = [1, 0.03, 300]
popt, pcov = curve_fit(exponential_func, sorted_data["period_fractional"],␣
 ↪sorted_data["CO2"], p0=initial_guess)
fitted_values = exponential_func(sorted_data["period_fractional"], *popt)
plt.plot(sorted_data.index, sorted_data['CO2'], color='blue',␣
 ↪label='Concentration mésurée')
plt.plot(sorted_data.index, fitted_values, color='red', label='Augumentation␣
 ↪lente')
plt.legend()
```

```
plt.ylabel(r"Concentration en $CO_2$ [ppm]")
plt.xlabel("Année")
plt.grid(linestyle=":")
plt.show()
print(f"Parametres fittés: a = {popt[0]:.2f}, b = {popt[1]:.4f}, c = {popt[2]:.
↪2f}")
```




```
Parametres fittés: a = 54.73, b = 0.0166, c = 258.87
```