



# Statistiques Appliquées à l'Océanographie

## Projet de Data Analysis

Présenté par :  
Julio Leonel DONFACK

Sous la direction de :  
**Prof. Frédéric BONOU**

# Statistiques Appliquées à l’Océanographie

## Projet de Data Analysis

Julio, Leonel DONFACK

25 février 2024

---

---

# Table des matières

---

<b>1 Projet de Data Analysis</b>	<b>1</b>
1.1 Zone d'étude et caractéristiques des grandeurs considérées . . . . .	1
1.2 Analyse statistique unidimensionnelle . . . . .	1
1.2.1 Statistique descriptive : Paramètres de tendance centrale et de dispersion	2
1.2.2 Tests statistiques et estimation de la moyenne de SST dans le Pacifique tropical . . . . .	4
1.3 Statistique bidimensionnelle . . . . .	7
1.3.1 Statistique bidimensionnelle 1 . . . . .	7
1.3.2 Statistique bidimensionnelle 2 . . . . .	11
1.4 Statistique tridimensionnelle . . . . .	14
1.4.1 Statistique descriptive : Paramètres de tendance centrale et de dispersion	15
1.4.2 corrélation multiple . . . . .	16
1.5 Analyse en composantes principales . . . . .	18
1.5.1 Anomalie saisonnière . . . . .	19
1.6 Conclusion . . . . .	24
<b>Bibliographie</b>	<b>25</b>

---

# Table des figures

---

1.1	Pluviométrie et SST dans le Pacifique tropical en Janvier 1998. . . . .	2
1.2	Série temporelle de la SST annuelle dans le Pacifique équatorial de 1958 à 2023.	3
1.3	Série temporelle de la SST mensuelle dans le Pacifique équatorial de 1958 à 2023.	4
1.4	Histogrammes des deux séries de valeurs moyennes mensuelles de SST. . . . .	6
1.5	Séries temporelles d'anomalies de SST et de précipitation pour la période de janvier 1958 à décembre 2023. . . . .	8
1.6	Séries temporelles normalisées, Nuage de points de SST et d'anomalie de précipitation. . . . .	9
1.7	Fonction de corrélalion de la SST et de l'anomalie de précipitation. . . . .	10
1.8	Séries temporelles d'anomalies de SST et de CO2 pour la période de janvier 1958 à décembre 2023. . . . .	11
1.9	Séries temporelles normalisées, Nuage de points de SST et d'anomalie de CO2. .	12
1.10	Fonction de corrélalion de la SST et de l'anomalie de CO2. . . . .	13
1.11	Séries temporelles d'anomalies de SST, de CO2 et de précipitation pour la période de janvier 1958 à décembre 2023. . . . .	15
1.12	Séries temporelles normalisées, Nuage de points de CO2 et d'anomalie de précipitation. . . . .	17
1.13	Distribution du CO2 et des précipitations selon les températures. . . . .	18
1.14	Climatologie de la SST. . . . .	19
1.15	Climatologie de l'anomalie de précipitation. . . . .	19
1.16	Anomalie saisonniere de temperature. . . . .	20

1.17 Illustration des valeurs propres les plus importantes. . . . .	21
1.18 Pourcentages relatifs des variances expliquées. . . . .	21
1.19 (a) Structure spatiale de l'anomalie de SST et (b) série temporelle du premier mode de la décomposition en composantes principales. . . . .	22
1.20 Structure spatiale et série temporelle du deuxième ((a) et (b)), troisième ((c) et (d)) et quatrième ((e) et (f)) mode de l'analyse en EOF. . . . .	23

# PROJET DE DATA ANALYSIS

---

## 1.1 Zone d'étude et caractéristiques des grandeurs considérées

Cette étude statistique porte sur le Pacifique tropical qui est le théâtre des événements El-Niño. Pour la mener à bien, nous considérons la température (SST) [1], en degré Celsius, de surface de l'océan Pacifique tropical, la concentration de CO<sub>2</sub> [2] en micro mol/mol et les anomalies mensuelles de précipitations [3] données en millimètre par jour dans la même zone d'étude définie par les coordonnées (140°E-75°W, 30°S-30°N). La période d'étude s'étend de janvier 1958 à décembre 2023 soit 66 ans.

## 1.2 Analyse statistique unidimensionnelle

Nous étudions dans ce paragraphe les statistiques unidimensionnelles des séries temporelles de SST et précipitation moyennes dans le Pacifique tropical. La carte de la figure 1 est une bonne illustration de la zone d'étude. En outre elle nous illustre les variations spatiales de SST et de précipitation (anomalie mensuelle de précipitation plus précisément) au mois de janvier 1998 dans la région d'intérêt. On peut remarquer qu'il y'a un maximum de précipitation dans le Pacifique équatorial central, dû au déplacement de la 'warm pool' du Pacifique ouest au centre du bassin pendant le phénomène d'El-Niño des années 1997-1998.

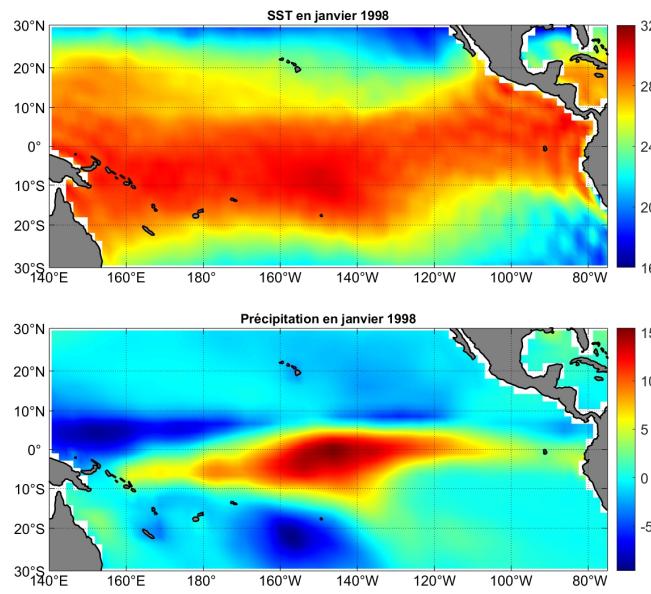


FIGURE 1.1 – Pluviométrie et SST dans le Pacifique tropical en Janvier 1998.

### 1.2.1 Statistique descriptive : Paramètres de tendance centrale et de dispersion

Intéressons nous à présent à la statistique descriptive de la SST dans le Pacifique tropical. Pour cela, pour un mois donné, nous calculons la SST représentative de la région d'étude en faisant une moyenne spatiale sur toute la région. On obtient ainsi une série temporelle de 792 valeurs qui représentent la SST aux différents mois de la période allant de janvier 1958 à décembre 2023. En prenant ensuite les moyennes annuelles on a une série statistique représentative des 66 années. Cette série est illustrée à la figure 2. Quelques paramètres quantitatifs de cette série de SST sont :

- La moyenne de la SST dans la région au cours de la période considérée est 25,83 °C. Elle est représentée par la ligne horizontale épaisse noire de la figure 2.
- La médiane est 25,84 °C. Sur la période d'étude qui est de 66 ans, pendant au moins 33 années la SST est inférieure à la médiane (25,84°C) ; et pendant au moins 33 autres années, elle est supérieure à 25,84 °C.
- Le mode de la série est 2015, c'est-à-dire qu'en 2015 il a fait le plus chaud dans le

- Pacifique tropical. La température a atteint la valeur de 26,39 °C.
- L'écart-type est 0.29 °C. Elle caractérise la dispersion des valeurs de la SST autour de la moyenne. La plupart des valeurs de SST dans cette région se situe dans l'intervalle [25,83 – 0,29 °C ; 25,83 + 0,29 °C].

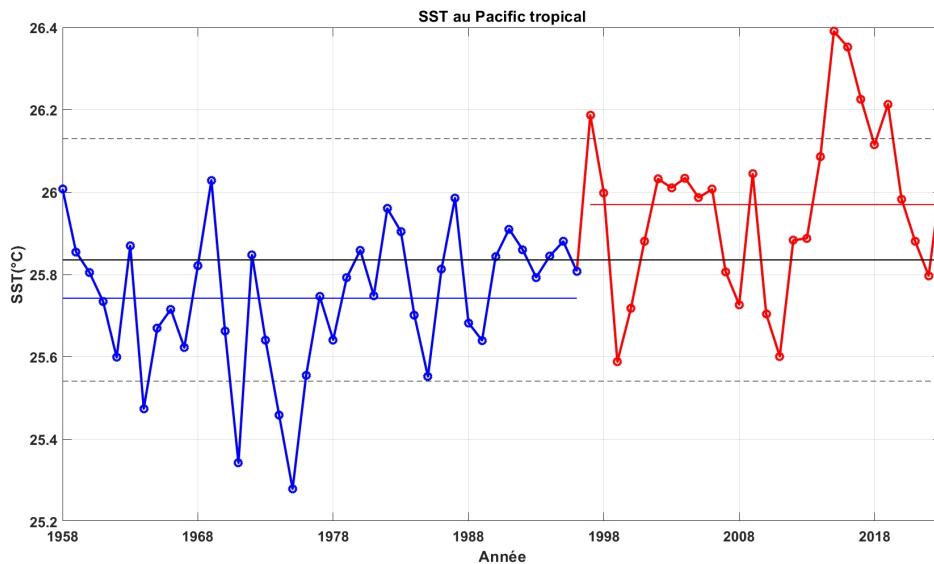


FIGURE 1.2 – Série temporelle de la SST annuelle dans le Pacifique équatorial de 1958 à 2023.

L'observation de la série temporelle de la figure 2 nous permet de faire une remarque notable. En effet, la plupart des valeurs de SST sont comprises dans l'intervalle [25,54 °C ; 26,12 °C] (voir les lignes horizontales interrompues de la figure 2) comme mentionné plus haut. Cependant jusqu'en 1996, toutes les valeurs de SST sont inférieures à 26,12 °C. En plus, au delà de 1997, toutes les valeurs de SST sont supérieures à 25,54 °C. Ainsi, en 1997 la SST est supérieure à celle de toutes les années précédentes. En d'autres termes, la remarque frappante est qu'au delà de 1997 les valeurs de la SST sont pour la plupart largement supérieures à la moyenne, contrairement aux valeurs de SST des années avant 1997. Cette remarque nous suggère de scinder la série de SST en deux séries représentative des années 1958 à 1996 et 1997 à 2023. Sur la figure 1.2, la première série est en bleu et la seconde série est en rouge. Les moyennes des deux séries ont été également évaluées et valent 25,68 °C pour la première et 25,86 °C pour la seconde. Elles sont représentées respectivement par les lignes horizontales bleu et rouge sur

la figure 1.2.

### 1.2.2 Tests statistiques et estimation de la moyenne de SST dans le Pacifique tropical

Dans le but de limiter la perte d'information, nous reprenons la série temporelle de 792 valeurs qui représentent la SST aux différents mois de la période allant de janvier 1958 à décembre 2023. Comme suggéré plus haut, nous scindons cette série en deux, une première constituée de 468 valeurs représentatives des années 1958 à 1996 et une deuxième constituée de 324 valeurs. La série temporelle initiale est représentée à la figure 1.3.

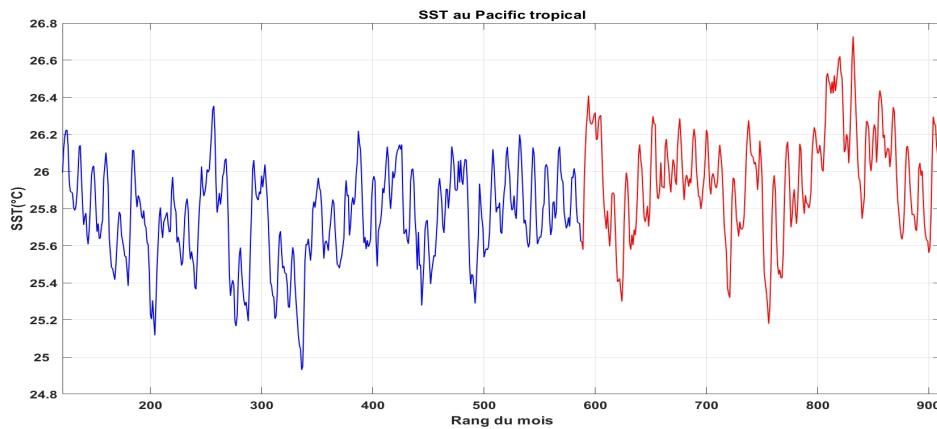


FIGURE 1.3 – Série temporelle de la SST mensuelle dans le Pacifique équatorial de 1958 à 2023.

#### 1.2.2.1 Test d'homogénéité

Vérifions que chacune de ces séries suit une loi normale. Pour cela, nous regroupons chacune de ces séries en des classes d'amplitude 0.05. L'histogramme de chacune des séries est représenté sur la figure 4. Notons que le nombre de classes des séries est  $m_1 = 37$  pour la série n°1 et  $m_2 = 41$  pour la série n°2. Effectuons le test d'homogénéité du  $\chi^2$  à chacune de ces séries afin de conclure. On a par définition :

$$\chi^2 = \sum_{i=1}^m \frac{(SST_i - E_i)^2}{E_i}, \quad (1.1)$$

où  $E$  est la densité de probabilité théorique d'occurrence des valeurs mensuelles de SST si on admettait que ces valeurs suivent la distribution normale de moyenne  $\mu$  et d'écart type  $\sigma$  espérée. On a

$$E_i = \frac{m}{\sigma\sqrt{2\pi}\text{Sum}(\text{SST})} \exp\left(-\frac{(SST_i - \mu)^2}{2\sigma^2}\right), \quad (1.2)$$

où  $m$  est bien évidemment le nombre de classes de la série temporelle de SST et  $\text{Sum}()$  est la fonction somme. Pour chacune des séries ci-dessus considérées, rappelons les valeurs des moyennes  $\mu$ , des écarts-type  $\sigma$  et des nombres de classes  $m$  :

$$\begin{aligned} \mu_1 &= 25.68 \text{ } ^\circ\text{C}, & \sigma_1 &= 0.27 \text{ } ^\circ\text{C}, & m_1 &= 37, \\ \mu_2 &= 25.86 \text{ } ^\circ\text{C}, & \sigma_2 &= 0.24 \text{ } ^\circ\text{C}, & m_2 &= 41. \end{aligned} \quad (1.3)$$

Connaissant ces différentes valeurs, la densité de probabilité théorique que suivraient les valeurs de SST si leur distribution était normale est évaluée grâce à l'équation 1.3 et représentée dans la figure 1.4. On observe un accord plus ou moins acceptable entre l'histogramme et la densité de probabilité théorique. Pour nous en convaincre nous suivons la démarche ci-dessous :

- Nous admettons l'hypothèse 'nulle' selon laquelle la distribution des valeurs de SST est conforme à la distribution normale avec un seuil de significativité de  $\alpha = 5\%$  pour chacune des séries considérées.
- Nous évaluons ensuite le paramètre  $\chi_2$  grâce à l'équation 1.1 pour chacune des séries.
- Ce paramètre est comparé à la valeur standard  $\chi_2^\alpha$ , évaluée ici grâce au logiciel MATLAB.

Pour la série 1, on a  $\chi_2 = 39.25$  et  $\chi_2^\alpha = 51.00$  et pour la série 2 on a  $\chi_2 = 35.11$  et  $\chi_2^\alpha = 55.76$ . On remarque bien que  $\chi_2 < \chi_2^\alpha$  dans chacun des deux cas. Par conséquent on conclut que les deux séries de SST sont conformes aux distributions normales, bien-sûr avec une limite de significativité de 95%.

### 1.2.2.2 Test de Fisher sur la variance

D'après ce qui précède, les deux (2) séries des valeurs mensuelles de SST dans le Pacifique tropical sont distribuées selon la loi de probabilité gaussienne. Il est nécessaire de vérifier que leurs variances sont significativement identiques. Pour cela nous calculons le paramètre de Fisher

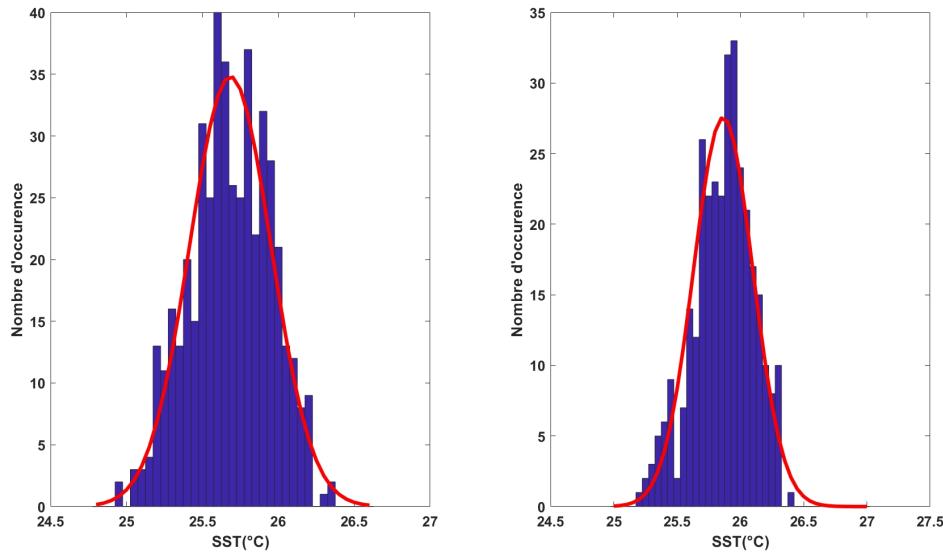


FIGURE 1.4 – Histogrammes des deux séries de valeurs moyennes mensuelles de SST.

défini par :

$$F = \frac{\sigma_2^2}{\sigma_1^2}. \quad (1.4)$$

Ce paramètre vaut  $F = 0.7688$ . Le paramètre critique de Fisher, évalué sur MATLAB vaut  $F_c = 1.1817$ . On a manifestement  $F < F_c$ . Il vient que les deux séries ont la même variabilité avec une certitude de 95%.

### 1.2.2.3 Test de Student sur la valeur moyenne

Nous venons de montrer ci-dessus que les deux séries considérées sont distribuées normalement et ont des variances significativement identiques avec une certitude de 95%. Il est à présent important de vérifier si la différence des valeurs de leurs moyennes est simplement la conséquence de la différence d'échantillonnage ou non. On invoque pour ce faire la statistique de Student définie par :

$$t = \frac{\mu_1 - \mu_2}{S_{12}}, \quad (1.5)$$

où

$$S_{12} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}. \quad (1.6)$$

On a d'après l'équation 1.5  $t = -10.0405$  et le paramètre critique vaut  $t_c = -1.963$ . On constate que  $|t| > |t_c|$ . On conclut que les deux séries précédentes ont des valeurs de SST dans le Pacifique tropical ont des moyennes significativement différentes.

L'intervalle de confiance à 95% de la moyenne de SST est donné par

$$\left[ \mu - 2,26 \frac{\sigma}{\sqrt{n-1}} ; \mu + 2,26 \frac{\sigma}{\sqrt{n-1}} \right]. \quad (1.7)$$

Pour les années allant 1958 à 1996 l'intervalle de confiance à 95% de la moyenne de la SST moyennée sur le Pacifique tropical est  $[25,66^\circ\text{C}; 25,71^\circ\text{C}]$ . Pour les années allant de 1997 à 2023, cet intervalle vaut plutôt  $[25,83^\circ\text{C}; 25,89^\circ\text{C}]$ . On peut remarquer que ces deux intervalles sont disjoints, le second ayant des valeurs supérieures au premier. Ce résultat rend compte d'un réchauffement du Pacifique tropical ces dernières 27 années.

## 1.3 Statistique bidimensionnelle

### 1.3.1 Statistique bidimensionnelle 1

Nous étudions dans ce paragraphe les éventuelles relations qui existent entre la SST et la pluviométrie dans le Pacifique tropical. Nous rappelons que nous avons précisément les données d'anomalie de la précipitation. La figure 1.5 illustre les séries temporelles d'anomalies de SST et de précipitation moyennées sur la zone d'étude.

Nous considérons dans la suite les séries temporelles des moyennes annuelles (sur 66 ans) de SST et d'anomalie de précipitation et étudions les corrélations pouvant exister entre les deux quantités. Pour ce faire, les séries normalisées de ces deux grandeurs sont illustrées à la figure 1.6. Elles sont évaluées grâce à la relation

$$x_{norm} = \frac{x - \mu}{\sigma}, \quad (1.8)$$

où  $\mu$  et  $\sigma$  sont respectivement la moyenne et l'écart-type. Notons que les valeurs d'écart-type sont  $\sigma_1 = 0.22$  et  $\sigma_2 = 0.13$  pour la SST et l'anomalie de précipitation respectivement. L'inspection de la figure 1.6 nous révèle une certaine corrélation entre la SST et la pluviométrie notamment pour certaine période. En effet, entre 1958 et 1962 par exemple la SST et l'anomalie de précipitation ont des sens de variation synchronisés. Au cours des années 1956, 1969-1970,

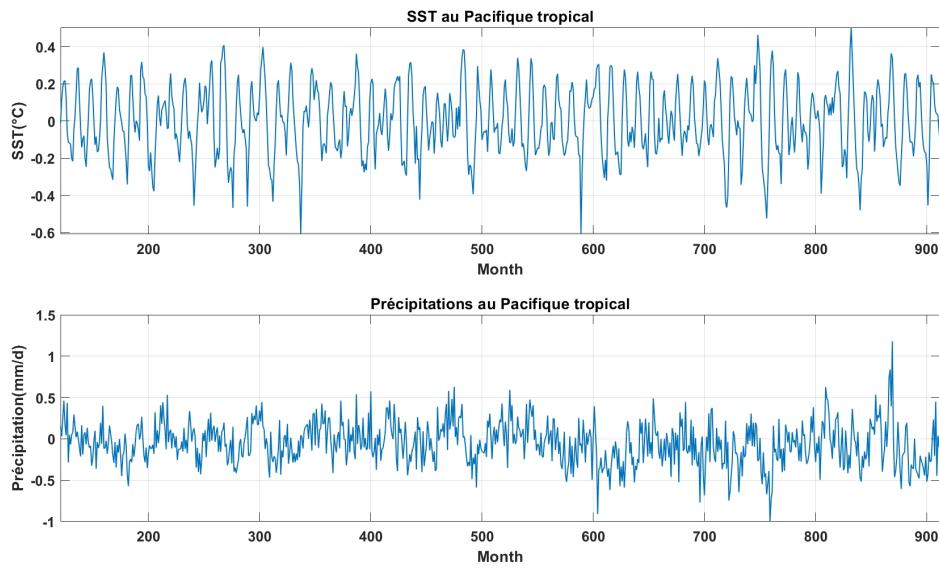


FIGURE 1.5 – Séries temporelles d'anomalies de SST et de précipitation pour la période de janvier 1958 à décembre 2023.

1982-1983 et 1988 en particulier, on observe une bonne corrélation. Enfin, il en ressort de l'analyse de la figure 1.6 qu'au delà de 1994 les valeurs de SST sont dans l'ensemble plus importantes que celles des années avant 1994. Tandis que le contraire s'observe pour la pluviométrie, en effet, après l'année 1994, les valeurs d'anomalie de précipitation baissent dans l'ensemble.

### 1.3.1.1 Etude de la corrélation linéaire

Le coéfficient de corrélation linéaire des deux séries est évalué sur le logiciel MATLAB par une fonction implémentant la relation

$$r = \frac{1}{N-1} \sum_{i=1}^N \frac{(x_i - \mu_1)}{\sigma_1} \frac{(y_i - \mu_2)}{\sigma_2}. \quad (1.9)$$

Sa valeur vaut  $r = 0.31$ . Sur la figure 1.6 nous représentons le nuage de points relatif aux deux séries. Nous y représentons également la droite de regression linéaire dont les coefficients sont calculés sur le logiciel MATLAB. Au regard de cette figure, il existerait une corrélation linéaire entre la SST et les précipitations dans le Pacifique tropical. Pour nous en convaincre nous effectuons ci-dessous le test de corrélation linéaire.

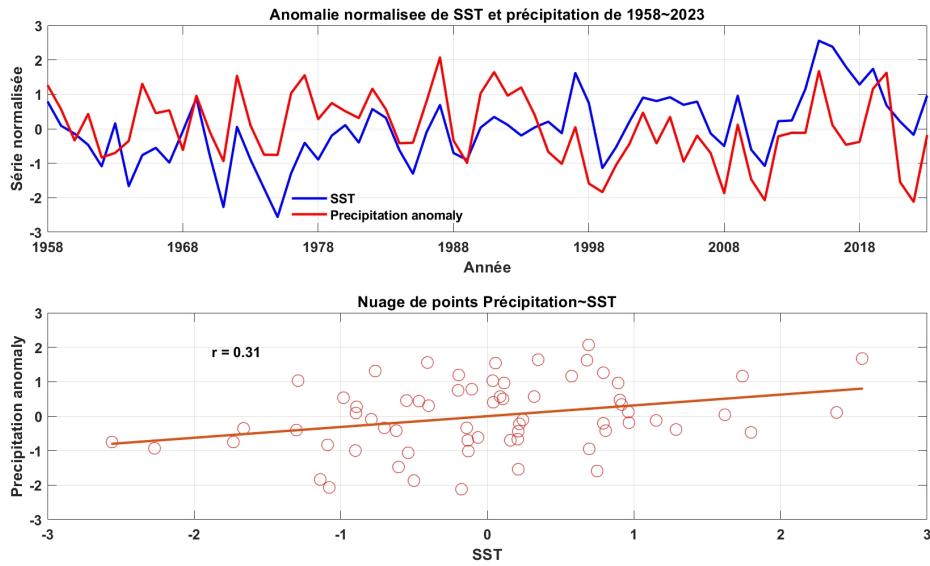


FIGURE 1.6 – Séries temporelles normalisées, Nuage de points de SST et d'anomalie de précipitation.

### 1.3.1.2 Test de Student sur la corrélation

Pour y parvenir, nous admettons l'hypothèse 'nulle' selon laquelle il n'y a aucune corrélation entre la SST et la pluviométrie. Nous admettons également que les valeurs d'anomalie de précipitation sont normalement distribuées, alors la quantité  $t$  exprimée comme

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (1.10)$$

suit une loi de Student avec  $n - 2$  degré de liberté. L'évaluation de  $t$  nous donne  $t = 2.62$ . Le paramètre critique  $t_c$  de Student avec un niveau de confiance de 95% vaut  $t_c = 2$ . On rejette l'hypothèse nulle puisque  $|t| > |t_c|$ . Par conséquent la corrélation linéaire entre la SST et l'anomalie de précipitation est significativement non nulle avec un niveau de confiance de 95%. Intéressons nous à présent à l'auto-corrélation des grandeurs considérées dans cette étude.

### 1.3.1.3 Corrélation avec décalage temporel

Dans le but d'examiner l'effet "mémoire" qu'aurait la série temporelle de la SST à l'égard de celle des précipitations, nous calculons grâce au logiciel et représentons dans la figure 1.7 la

fonction de corrélation  $r(\tau)$  des grandeurs en considération dans ce rapport. Déterminons pour

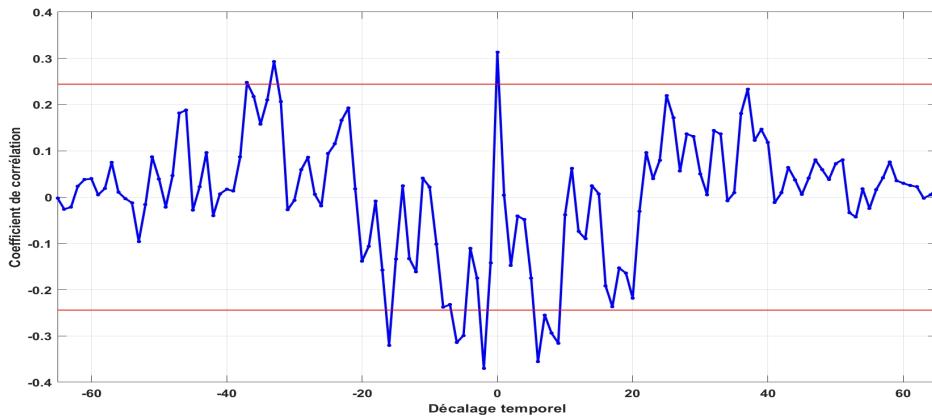


FIGURE 1.7 – Fonction de corrélalion de la SST et de l'anomalie de précipitation.

quelles valeurs  $r$  de la fonction de corrélation, la corrélation linéaire entre la SST et l'anomalie de précipitation est significativement non nulle avec un niveau de confiance de 95%. Pour cela rappelons que l'expression du paramètre de Student dans ce cas est donnée par l'équation (1.10). Partant de cette équation, on obtient l'expression de  $r$  donnée par

$$r = \pm \sqrt{\frac{t^2}{n - 2 + t^2}}. \quad (1.11)$$

Sachant que la corrélation linéaire entre les grandeurs considérées est significativement non nulle lorsque  $|t| > |t_c|$ , trouvons un encadrement de  $r$ . On montre évidemment que, dans le cas  $n > 2$ , la corrélation linéaire est significativement non nulle pour

$$\begin{cases} r > \sqrt{\frac{t_c^2}{n - 2 + t_c^2}} & \text{si } r > 0 \\ r < -\sqrt{\frac{t_c^2}{n - 2 + t_c^2}} & \text{si } r < 0 \end{cases} \quad (1.12)$$

Nous prenons ici un seuil de significativité de 95% et évaluons le second membre de la relation (1.12) grâce au logiciel MATLAB. Il vient que :

$$\begin{cases} r > 0.244 & \text{si } r > 0 \\ r < -0.244 & \text{si } r < 0 \end{cases} \quad (1.13)$$

Sur la figure 1.7, deux lignes horizontales correspondant aux valeurs seuils de  $r$  sont représentées. La lecture de ce graphique nous permet de déterminer les 'lags' pour lesquels la corrélation est significativement non nulle, avec un seuil de significativité de 95%. Il s'agit des 'lags' -16,-5, -3, -2, 0, 6-9 et 25. Ce résultat signifie qu'il y'a une corrélation linéaire non nulle entre la SST à la date  $t + 25$  ans et la pluviométrie à la date  $t$ .

### 1.3.2 Statistique bidimensionnelle 2

Nous étudions dans ce paragraphe les éventuelles relations qui existent entre la SST et la distribution des concentration de CO<sub>2</sub> dans le Pacifique tropical. En suivant les mêmes étapes que précédemment, nous commençons par rappelé ici que nous avons précisément les données d'anomalie de CO<sub>2</sub>. La figure 1.8 illustre les séries temporelles d'anomalies de SST et de CO<sub>2</sub> moyennées sur la zone d'étude.

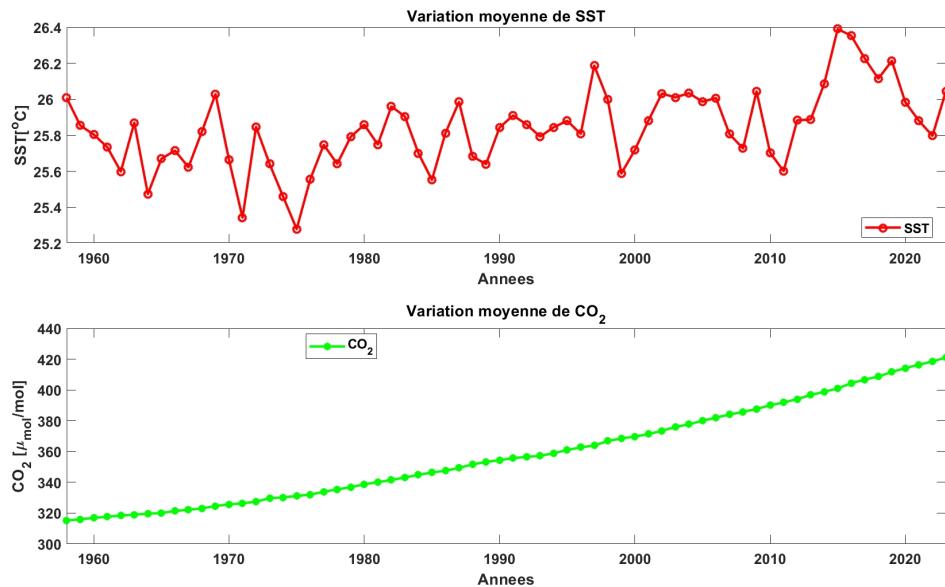


FIGURE 1.8 – Séries temporelles d'anomalies de SST et de CO<sub>2</sub> pour la période de janvier 1958 à décembre 2023.

Nous considérons dans la suite les séries temporelles des moyennes annuelles (sur 66 ans) de SST et d'anomalie de CO<sub>2</sub> et étudions les corrélations pouvant exister entre les deux quantités.

Pour ce faire, les séries normalisées de ces deux grandeurs sont illustrées à la figure 1.9. Elles sont évaluées grâce à la relation (1.8).

Notons que les valeurs d'écart-type sont  $\sigma_1 = 0.22$  et  $\sigma_3 = 31.55$  pour la SST et l'anomalie de CO<sub>2</sub> respectivement. L'inspection de la figure 1.9 nous révèle une certaine corrélation entre la SST et CO<sub>2</sub> notamment pour certaine période. En effet, nous observons une croissance de la SST et du CO<sub>2</sub> sur la période d'étude. La SST présente des minima et maxima de plus en plus hauts. Quant au CO<sub>2</sub> nous observons une croissance presque affine sur la période d'étude.

### 1.3.2.1 Etude de la corrélation linéaire

Le coefficient de corrélation linéaire des deux séries est évalué sur le logiciel MATLAB par une fonction implémentant la relation (1.9).

Sa valeur vaut  $r = 0.5540$ . Sur la figure 1.9 nous représentons le nuage de points relatif aux deux séries. Nous y représentons également la droite de régression linéaire dont les coefficients sont calculés sur le logiciel MATLAB. Cette figure, confirme bien l'existence d'une corrélation linéaire entre la SST et le CO<sub>2</sub> dans le Pacifique tropical.

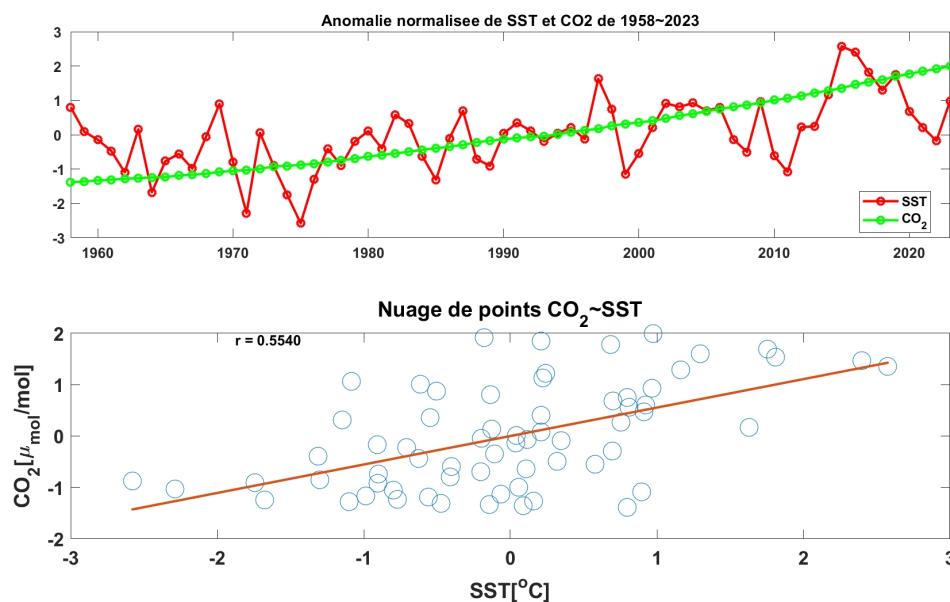


FIGURE 1.9 – Séries temporelles normalisées, Nuage de points de SST et d'anomalie de CO<sub>2</sub>.

Pour nous en convaincre nous effectuons ci-dessous le test de corrélation linéaire.

### 1.3.2.2 Test de Student sur la corrélation

Pour y parvenir, nous admettons l'hypothèse 'nulle' selon laquelle il n'y a aucune corrélation entre la SST et le CO<sub>2</sub>. Nous admettons également que les valeurs d'anomalie de CO<sub>2</sub> sont normalement distribuées, alors la quantité  $t$  exprimée d'après la relation (1.10) suit une loi de Student avec  $n - 2$  degrés de liberté. L'évaluation de  $t$  nous donne  $t = 5.2817$ . Le paramètre critique  $t_c$  de Student avec un niveau de confiance de 95% vaut  $t_c = 2$ . On rejette l'hypothèse nulle puisque  $|t| > |t_c|$ . Par conséquent la corrélation linéaire entre la SST et l'anomalie de CO<sub>2</sub> est significativement non nulle avec un niveau de confiance de 95%. Intéressons-nous à présent à l'auto-corrélation des grandeurs considérées dans cette étude.

### 1.3.2.3 Corrélation avec décalage temporel

Dans le but d'examiner l'effet "mémoire" qu'aurait la série temporelle de la SST à l'égard de celle du CO<sub>2</sub>, nous calculons grâce au logiciel et représentons dans la figure 1.10 la fonction de corrélation  $r(\tau)$  des grandeurs en considération dans ce rapport. Déterminons pour quelles

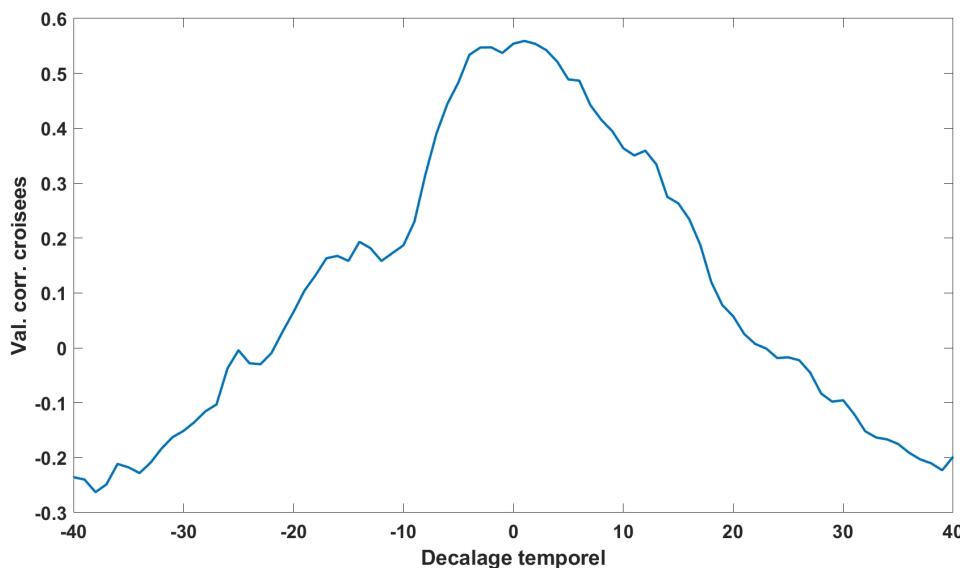


FIGURE 1.10 – Fonction de corrélation de la SST et de l'anomalie de CO<sub>2</sub>.

valeurs  $r$  de la fonction de corrélation, la corrélation linéaire entre la SST et l'anomalie de CO<sub>2</sub> est significativement non nulle avec un niveau de confiance de 95%. Pour cela rappelons que l'expression du paramètre de Student dans ce cas est donnée par l'équation (1.11). Partant de cette équation, on obtient l'expression de  $r$  donnée par (1.12). Sachant que la corrélation linéaire entre les grandeurs considérées est significativement non nulle lorsque  $|t| > |t_c|$ , trouvons un encadrement de  $r$ . On montre évidemment que, dans le cas  $n > 2$ , la corrélation linéaire est significativement non nulle pour (1.12).

Nous prenons ici un seuil de significativité de 95% et évaluons le second membre de la relation (1.12) grâce au logiciel MATLAB. Il vient que :

$$\begin{cases} r > 0.2441 & \text{si } r > 0 \\ r < -0.2441 & \text{si } r < 0 \end{cases} \quad (1.14)$$

La figure 1.10 montre une corrélation positive et maximale ( $r = 0,5540$ ) en l'absence de décalage temporel. En d'autres termes une variation de la concentration de CO<sub>2</sub> au cours d'une année pourrait s'expliquée par une variation de la SST, à la surface de l'océan et vice-versa. La courbe d'anomalie normalisée et le nuage de points (1.9) montrent une tendance de croissance commune aux deux courbes. L'évolution du CO<sub>2</sub> en fonction de la SST suit une droite affine dont les paramètres sont  $a = r, b = 0$  puisque nous sommes en représentation normalisée de ces deux variables. Étant donné que nous observons un nuage de points(1.9) peu concentré sur la droite de régression : il vient que la variation totale (croissante) du (CO<sub>2</sub>, SST) ne peut pas être expliquée par la variance expliquée de la droite de régression (CO<sub>2</sub>, SST). Nous dirons que d'autres sources ont une contribution non négligeable sur la période d'étude.

## 1.4 Statistique tridimensionnelle

Nous étudions dans ce paragraphe les éventuelles relations qui existent entre la SST, la distribution des concentrations de CO<sub>2</sub> et la pluviométrie dans le Pacifique tropical. Nous commençons par rappelé ici que nous avons précisément les données d'anomalie de CO<sub>2</sub> et de précipitation. La figure 1.11 illustre les séries temporelles d'anomalies de SST, de CO<sub>2</sub> et de pluviométrie moyennées sur la zone d'étude. Nous considérons dans la suite les séries temporelles

des moyennes annuelles (sur 66 ans) de SST, d'anomalie de CO<sub>2</sub> ainsi que de pluviométrie et étudions les corrélations pouvant exister entre les trois quantités. Pour ce faire, nous commençons par faire une statistique descriptive rapide des trois variables . L'inspection de la figure 1.11 nous révèle une certaine corrélation entre la SST, CO<sub>2</sub> et pluviométrie notamment pour certaine période. En effet, nous observons une croissance de la SST et du CO<sub>2</sub> sur la période d'étude. La SST présente des minima et maxima de plus en plus hauts. Quant au CO<sub>2</sub> nous observons une croissance presque affine sur la période d'étude.

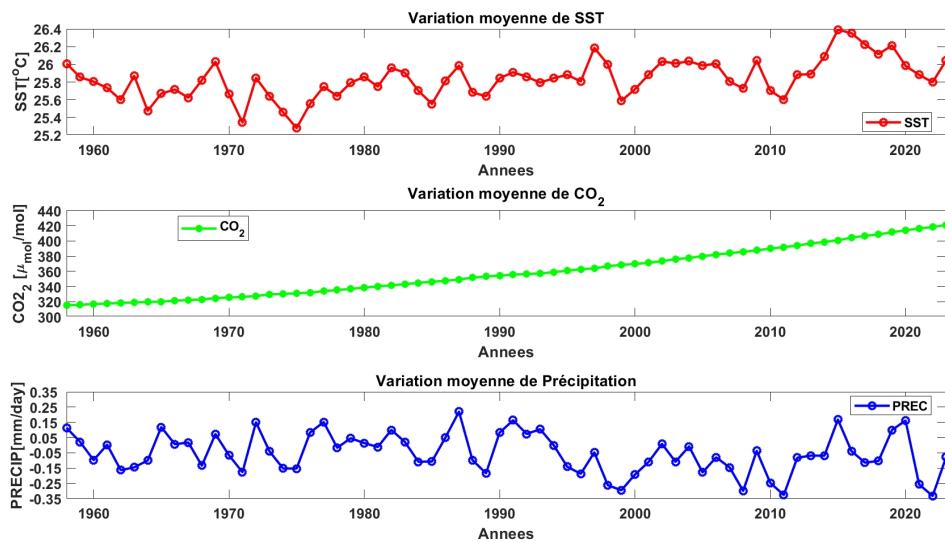


FIGURE 1.11 – Séries temporelles d'anomalies de SST, de CO<sub>2</sub> et de précipitation pour la période de janvier 1958 à décembre 2023.

#### 1.4.1 Statistique descriptive : Paramètres de tendance centrale et de dispersion

Intéressons nous à présent à la statistique descriptive de la SST, du CO<sub>2</sub> et de la pluviométrie dans le Pacifique tropical. En utilisant les fonctions `mean(.)`, `median(.)`, `mode(.)`, `max(.)`, `min(.)`, `quantile(.)` et `std` nous obtenons les principaux paramètres de tendances centrales et de dispersion. Le tableau 1.1 nous résume ces différents paramètres de tendance des trois variables :

	Min	1st Qu	Median	Mean	Mode	3rd Qu	Max	std
SST	25.28	25.70	25.84	25.83	25.28	25.99	26.39	0.22
CO2	315.23	330.19	355.07	358.59	315.23	384.02	421.08	31.55
Précipitation	-0.33	-0.15	-0.07	-0.05	-0.33	0.05	0.22	0.13

Tableau 1.1 -Paramètres de tendance centrale et de dispersion

### 1.4.2 corrélation multiple

Intéressons nous à présent à la corrélation multiple de la SST, du CO2 et de la pluviométrie dans le Pacifique tropical. En utilisant la fonction **corr(.)**, nous obtenons les degrés de liaisons entre les trois variables pris deux à deux. Le tableau 1.2 nous résume ces différents degrés de liaisons entre les trois variables :

corr	SST	CO2	Précipitation
SST	1.00	0.55	0.31
CO2	0.55	1.00	-0.27
Précipitation	0.31	-0.27	1.00

Tableau 1.2 - Corrélation multiple

D'après ce tableau nous pouvons endéduire une corrélation positive entre les variables de SST et CO2, positive également entre SST et précipitation. Il en ressort également de ce tableau une corrélation négative entre les variables de CO2 et précipitation. Il est à noté que nous avons une très grande corrélation entre la SST et le CO2 par rapport au précipitation. Ce qui confirme bien nos travaux précédent sur la statistique bidimensionnelle. C'est-à-dire l'augmentation de la SST entraîne de manière simultanée une augmentation du CO2 et des précipitations dans le Pacifique Tropical. Cependant, comme nous avons une corrélation négative entre les variables de CO2 et de pluviométrie, cela signifie que les deux variables évolue en sens opposé. Cela peut se justifier à travers la figure (12). Nous pouvons déduire de cette courbe que avant les années 1986, nous enregistrons plus de précipitation que d'émission de CO2 et après cette date, la situation c'est inversée. Qu'est-ce qui pourrait expliquer un tel retournement de situation ?

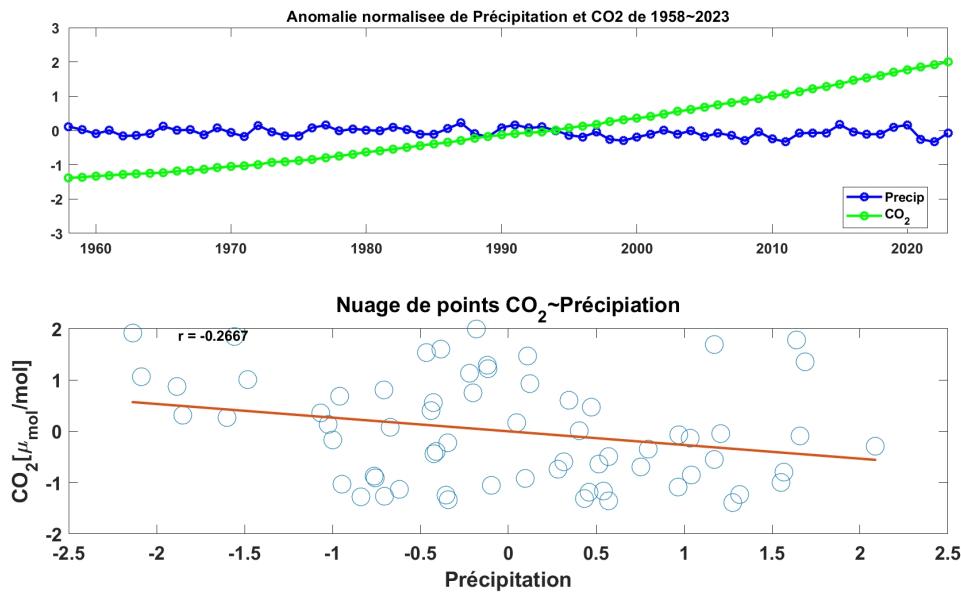


FIGURE 1.12 – Séries temporelles normalisées, Nuage de points de CO<sub>2</sub> et d'anomalie de précipitation.

Ceci pourrait s'expliquer par un certains nombre d'événements tels que la révolution industrielle débuté dans les années 1750 en Grande Bretagne selon le **GIEC** (IPCC en Anglais) et qui s'est de plus en plus accélérée et même généralisée à travers les différents coins du globe. Nous pouvons cité l'impact qu'auraient eu les essais des bombes nucléaires des années 1963. Bien évidemment une décroissance du carbone 14 dans l'atmosphère. Nous pouvons aussi cité les accidents nucléaires tels que celui du 26 avril 1986 (Tchernobyl) qui aurait contribué à cette décroissance du carbone 14 et donc une augmentation du CO<sub>2</sub> dans l'atmosphère. Nous pouvons également cité l'accident dans la centrale nucléaire de Fukushima de 2011 qui a contribué à cette montée exponentielle de CO<sub>2</sub> dans l'atmosphère. Pour meubler cette sous section, nous traitons la corrélation multiple entre nos trois principaux variables. Nous allons dans la suite représenter une courbe expliquant la distribution du CO<sub>2</sub> et des précipitations selon les températures. Cette courbe a été réalisé dans le language de programmation **R** et le code inclu dans le dossier Prog du projet. Cette courbe a été rendue possible en utilisant la fonction **ggplot2** de ce language et nous permet de prédire les données de précipitations et de CO<sub>2</sub> en fonction des températures.

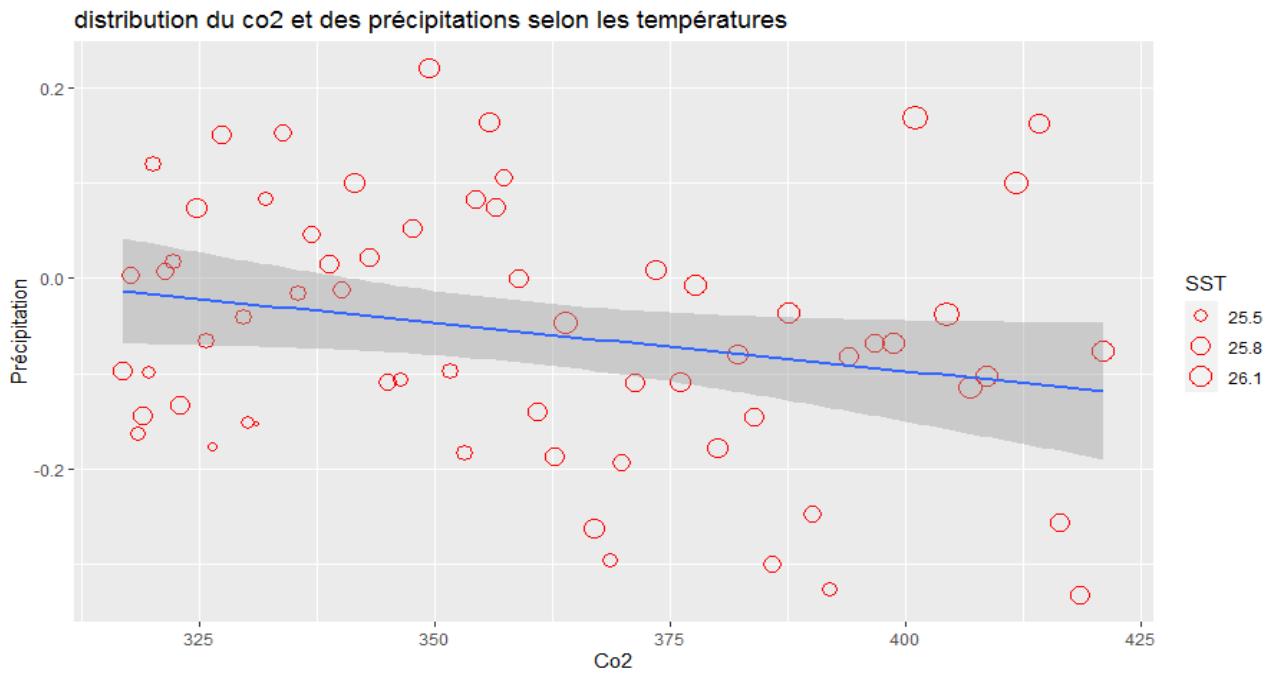


FIGURE 1.13 – Distribution du CO<sub>2</sub> et des précipitations selon les températures.

## 1.5 Analyse en composantes principales

L’analyse en EOF est une méthode de réduction de données qui fournit une description compacte de la variabilité spatiale et temporelle d’une série de données en termes de fonctions orthogonales (ou modes EOF pour Empirical Orthogonal Function) dont les premières expliquent généralement la plus grande partie de la variance [4]. Nous nous intéressons dans ce paragraphe uniquement à l’analyse en composantes principales de la SST dans le Pacifique tropical. Cependant, avant toute chose, le cycle saisonnier de la SST et celui de l’anomalie de précipitation sont illustrés respectivement dans les figures 1.14 et 1.15.

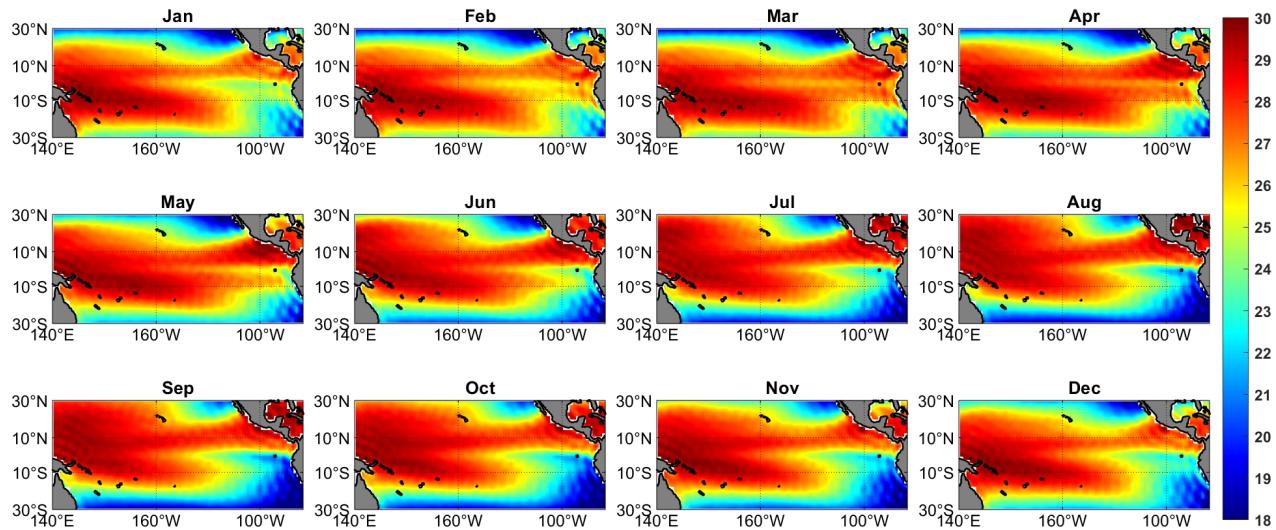


FIGURE 1.14 – Climatologie de la SST.

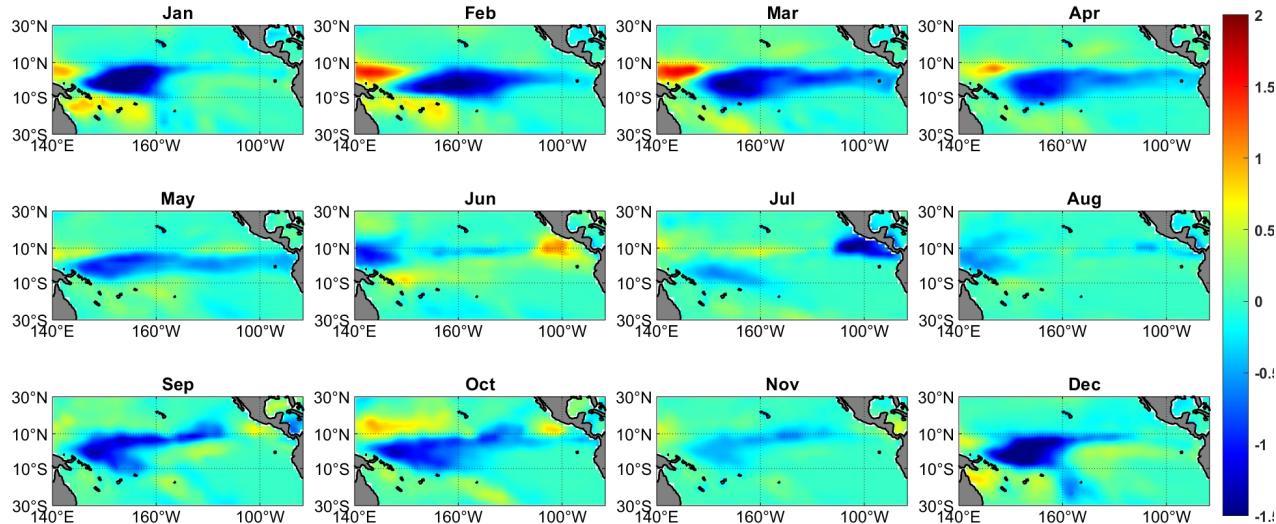


FIGURE 1.15 – Climatologie de l'anomalie de précipitation.

### 1.5.1 Anomalie saisonnière

L'océan est plus chaud entre l'hivers jusqu'au début du printemps(Janvier-Avril), entre janvier et avril nous observons une anomalie positive presque globale et beaucoup plus intense

au sud. En juin-juillet-aout-septembre-octobre et novembre, nous avons un mode chaud au nord qui décroît progressivement jusqu'à devenir froid dans certaines régions au sud. En mai et décembre, nous avons mode légèrement positif sur tous la zone d'étude.

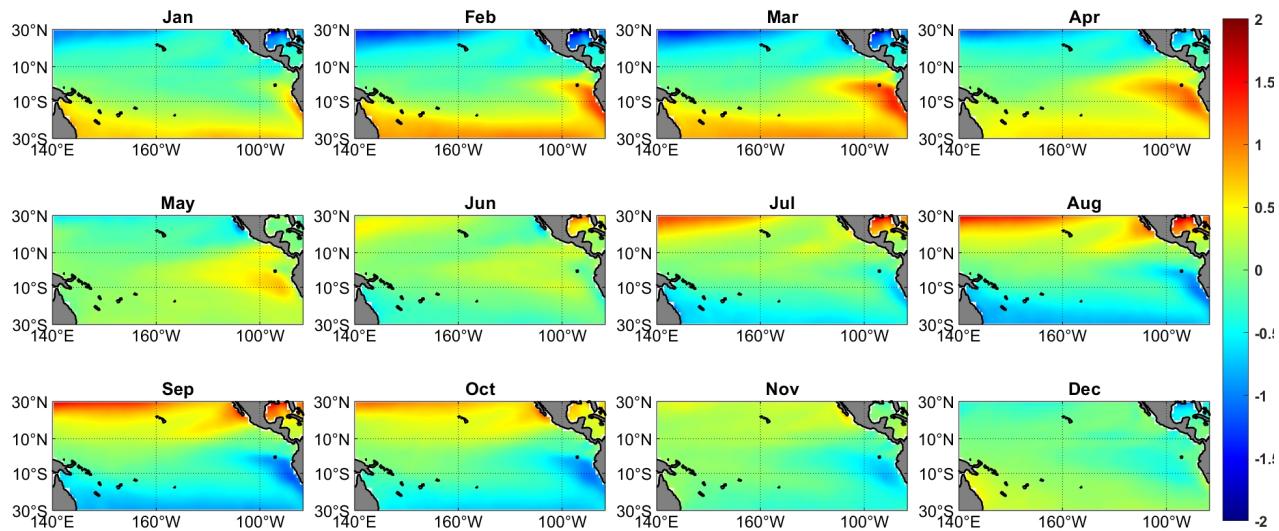


FIGURE 1.16 – Anomalie saisonniere de température.

Nous considérons à présent le champ tridimensionnel d'anomalies de SST par rapport à la moyenne annuelle (anomalie interannuelle), que nous transformons en un champ bidimensionnel de dimension (espace,temps). Nous appliquons ensuite la méthode de décomposition en composantes principales (ou en EOF), qui consiste dans un premier temps à déterminer les valeurs propres  $\lambda$  de la matrice de covariance. La représentation des 20 plus grandes valeurs propres en fonction du numéro de la composante est faite dans la figure 1.18. Le mode est environ 44.37%. Il est à noter que plus la valeur propre est élevée, plus la composante explique une portion significative de la variance totale. L'observation de cette figure 1.18 nous suggère la représentation de l'histogramme des quatre premières valeurs propres pour illustrer mieux la part de l'ensemble de l'information contenue sur chaque composante principale. Cette histogramme est représenté dans la figure 1.19. Ce diagramme permet également de comparer l'importance relative des composantes principales. Nous pouvons lire dans les figures 1.18 et 1.19 que la première composante principale de l'analyse, correspondant à la première valeur propre, concentre à lui seul à peu près la moitié de l'ensemble de l'information (44.37 %). La seconde, la troisième et la

quatrième composante en regroupant chacune respectivement 15.89% ; 10.62% et 5,39%, plus des trois quarts de la variance (76,27%) est expliquée par ces quatre premières composantes principales. On peut donc se contenter de retenir ces quatres axes factoriels pour cette analyse, l'information contenue sur les autres pouvant être considérée comme résiduelle. La décomposition de l'anomalie de la SST sur la période de janvier 1958 à décembre 2023 fait ressortir quatre

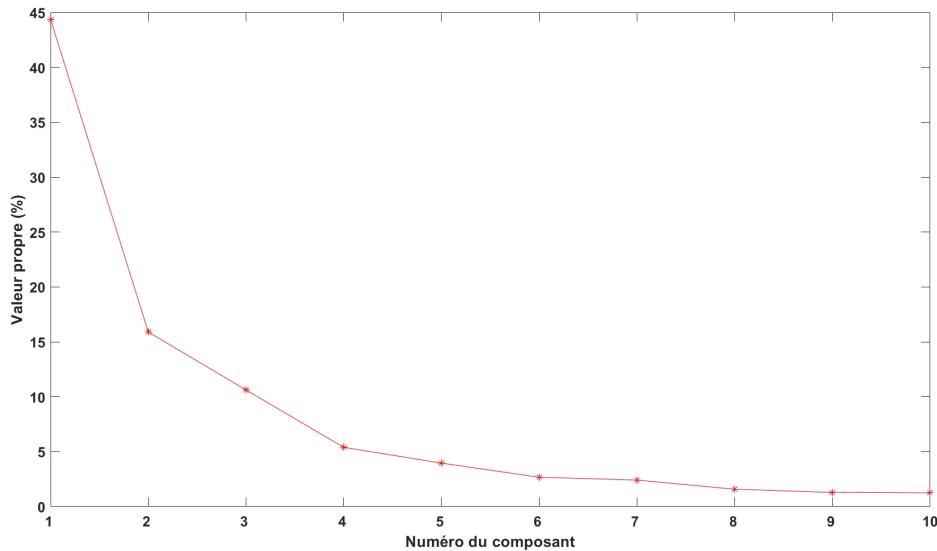


FIGURE 1.17 – Illustration des valeurs propres les plus importantes.

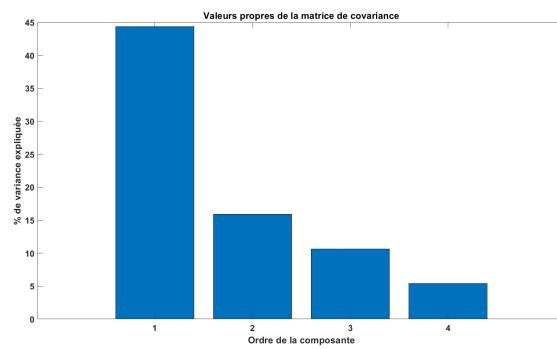


FIGURE 1.18 – Pourcentages relatifs des variances expliquées..

modes dominants illustrés sur les figures 1.20 et 1.21. Le premier mode, comme souligné plus

haut, représente environ 44% de la variance totale et s'apparente au mode équatorial [4]. La figure 1.20 nous donne la composante principale associée à ce mode (valeur propre maximale) ainsi que la carte d'anomalie de SST correspondante. Dans la figure 1.20(b), on peut observer des valeurs de la série chronologique normalisée qui excèdent l'écart-type, c'est-à-dire qui sont supérieures à +1 ou bien inférieures à -1. Ces valeurs indiquent les années où le mode équatorial est bien marqué. Les trois autres composantes principales sont également illustrées

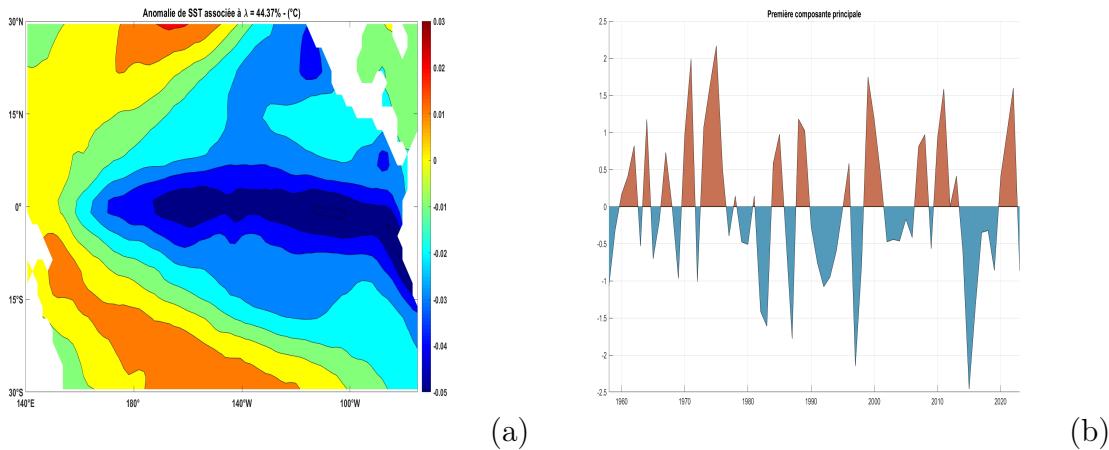


FIGURE 1.19 – (a) Structure spatiale de l'anomalie de SST et (b) série temporelle du premier mode de la décomposition en composantes principales.

sur la figure 1.21. On y observe à gauche les cartes physiques (carte d'anomalie de SST) associées et à droite les composantes principales correspondantes. Précisément, les figures de droites illustrent l'évolution temporelle des différentes EOFs. Les figures 1.21(a) et (b) correspondent au deuxième mode, 1.21(c) et (d) au troisième mode et enfin, les figures 1.21(e) et (f) correspondent au quatrième mode. La figure 1.14(b) montre qu'au delà de l'année 1998, il y a des années pendant lesquels le deuxième mode est considérablement marqué positivement. Tandis qu'avant 1998, il y a des années où ce mode est fortement marqué, cependant négativement. Ce deuxième mode, tout comme le premier s'apparente au mode équatorial.

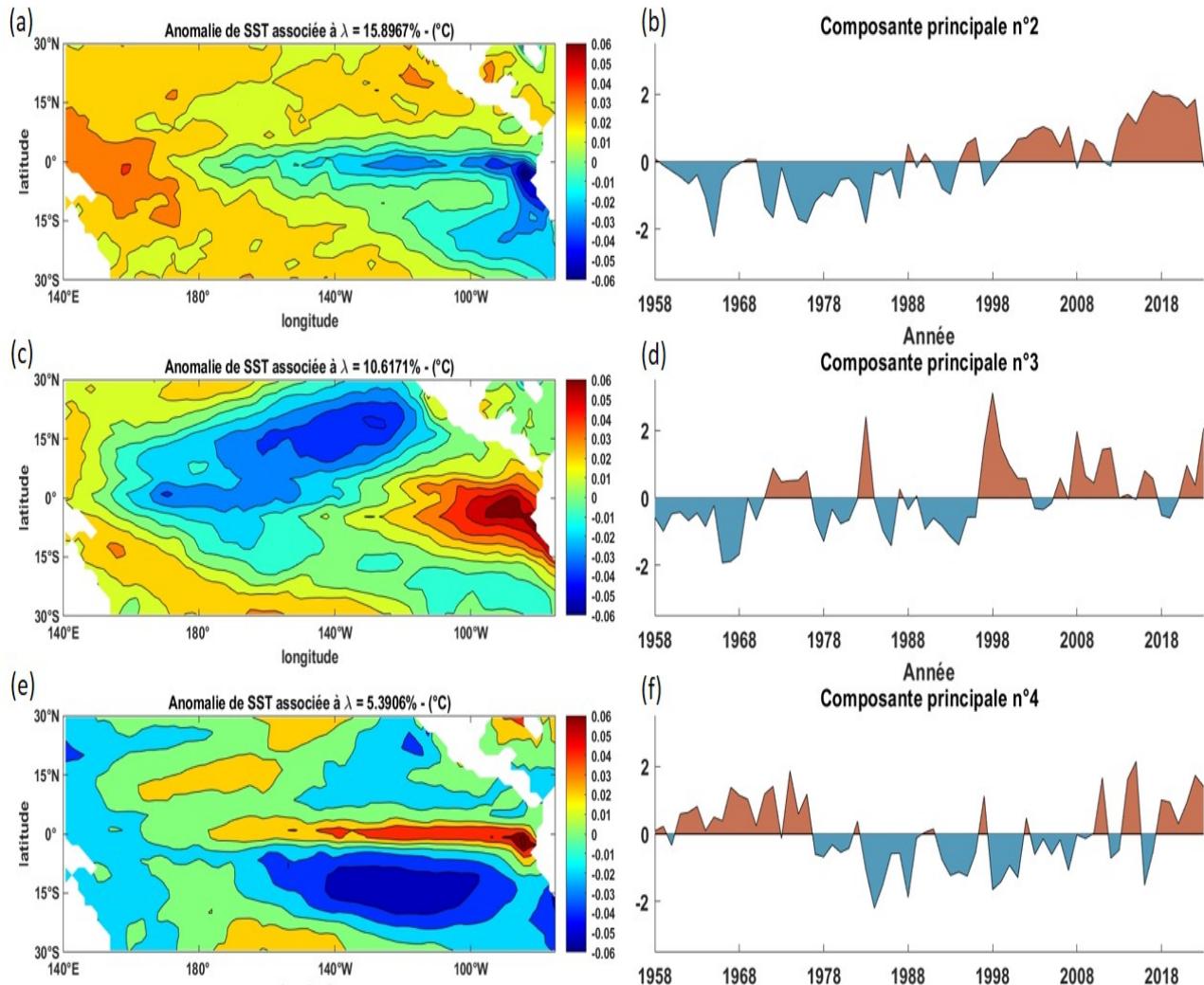


FIGURE 1.20 – Structure spatiale et série temporelle ((a) et (b)), troisième ((c) et (d)) et quatrième ((e) et (f)) mode de l’analyse en EOF.

## 1.6 Conclusion

Dans ce rapport, nous avons mené une étude statistique de la température de surface (SST), de la pluviométrie et de la distribution de CO<sub>2</sub> dans l'océan Pacifique tropical sur les dernières 66 années passées. Il en ressort que l'anomalie de précipitation et la SST ainsi que de concentration de CO<sub>2</sub> sont significativement corrélées dans cette région du globe. En outre, dans l'ensemble, la température dans cette région a augmenté d'environ 0.26 °C ces dernières 27 années tandis que l'anomalie de précipitation a plutôt baissée ces récentes années. Nous avons également trouvé que, pour des décalages temporels appropriés, la SST de cette région reste corrélée à l'anomalie de précipitation du passé (ou du futur). Il en est de même avec l'anomalie de CO<sub>2</sub> et cette corrélation est beaucoup plus importante que celle avec l'anomalie de précipitation et ceci peut s'expliquer par un certains nombre de phénomènes tels que le début de l'industrialisation en 1750 d'après le GIEC en Grande Bretagne, les tests des bombes nucléaires des années 1963 et les accidents nucléaires du 26 avril 1986 (Tchernobyl) qui ont réduit le taux de carbone 14 dans l'atmosphère nécessaire pour faire barrière aux émissions de CO<sub>2</sub>. Nous pouvons également noté l'accident nucléaire dans la centrale de Fukushima au Japon en 2011 qui a contribué d'avantage à cette augmentation de CO<sub>2</sub> dans l'atmosphère. Ce qui n'est pas une bonne nouvelle pour la flore biologique marine car cela fait diminuer la production primaire nécessaire pour son développement. Nous avons également noté une corrélation négative entre l'anomalie de précipitation et de CO<sub>2</sub>. Ceci se justifie par l'apparition de tous ces événements comme cela peut s'observer à travers la courbe de la figure 1.12 où la tendance c'est inversée et cela coïncide au minimètre près aux périodes des dits événements. L'analyse en composantes principales de la SST nous a permis de déterminer quatre modes qui expliquent plus des trois quarts de la variance totale, soit environ 76,27%. Les deux premiers modes s'apparentent bien au mode équatorial observé dans l'Atlantique tropical [4].

---

---

## Bibliographie

---

- [1] [https://downloads.psl.noaa.gov/Datasets/ncep.reanalysis.derived/surface\\_gauss/skt.sfc.mon.mean.nc](https://downloads.psl.noaa.gov/Datasets/ncep.reanalysis.derived/surface_gauss/skt.sfc.mon.mean.nc)
- [2] <https://downloads.psl.noaa.gov/data/gridded/data.gpcp.html>
- [3] <ftp://downloads.ftp.cmdl.noaa.gov/products/trends/co2/co2.mm.mlo.txt>
- [4] A. Founi Mesmin, Modes interannuels de la variabilité climatique de l'Atlantique tropical, dynamiques oscillatoires et signatures en salinité de surface de la mer. Thèse de doctorat (2018).