

Etude sur le Climat et Lien avec les Fièvres à Djibouti

Mohamed ABDI KHAIRE¹, Hawa ADEN FARAH² et Anne-Françoise YAO²

Résumé : Le monde connaît actuellement un réchauffement climatique très important et plusieurs conséquences possibles de ces changements font l'objet d'un consensus scientifique. La République de Djibouti est classée dans la catégorie des pays semi-arides de par ses conditions climatiques sévères. Afin de comprendre le concept de réchauffement climatique, l'évolution de la température et de la pluviométrie de Djibouti-ville seront suivies sur une période de 55 ans allant de 1961 à 2016. L'objectif principal de l'article est la localisation des points de changements sur les tendances estimées pour chacune des séries et ainsi estimer un modèle pour les cas des fièvres en considérant les variables pluviométriques et températures.

Abstract: The world is currently experiencing a very important global warming and several possible consequences of these changes are the subject of scientific consensus. The Republic of Djibouti is classified as a semi-arid country by its severe weather conditions. To understand the concept of global warming, the evolution of the temperature and rainfall of Djibouti-ville will be monitored over a period of 55 years old from 1961 to 2016. The main objective of the article is locating the points of change on the estimated XGs for each series and thus estimate a model for cases of fevers in considering the variable rainfall and temperature.

Mots-clés : Détection de rupture, Modèle prévisionnel, Tendances, ARIMA, Lissage exponentielle, Décomposition de série, Régression locale, Fenêtre de lissage h

Keywords: ChangePoint Detection, Predictive Model, XG, ARIMA, Exponential Smoothing, Decomposition of series, Local regression, Smoothing window h

¹ Laboratoire de Mathématiques, Université Clermont Auvergne et Département de Statistiques, Université de Djibouti.

E-mail : Mohamed.Abdi.Khaire@etu.uca.fr

² Département de Statistiques, Université de Djibouti et

Laboratoire de Mathématiques, Université Clermont Auvergne.

E-mail : anuom_hawa@hotmail.com and E-mail : Anne-francoise.Yao@math.univ-Baibclermont.fr

Introduction

Djibouti [RAYALEH \(2005\)](#) est un petit pays de l'Afrique de l'Est (environ 23000km^2), bordé par la mer rouge d'une part et de l'océan indien d'autre part.

Situé à la limite de l'équateur, son climat est chaud, aride et sec sur quasi les $\frac{3}{4}$ de l'année (saison chaude). Durant cette période de l'année, les températures peuvent atteindre les 48 degrés celsius (maximum), provoquant ainsi des sécheresses et des déplacements des populations nomades vers la capitale (Sédentarisation de la population).

Le manque d'infrastructure dans certain quartier et avec une augmentation de la pluviométrie fait augmenter le risque de prolifération pour tout type de pathologie.

L'objectif principal de l'article est l'étude des ruptures sur les séries températures et pluviométriques. Nous débuterons par une étude temporelle pour les différentes séries afin de modéliser un modèle prévisionnel qui peut être utile pour l'étude du phénomène : réchauffement climatique.

Ensuite, nous détecterons les dates susceptibles d'être considérées comme des années de ruptures. Le nombre des cas des fièvres d'origines inconnues sont souvent influencés par le changement climatique pendant les saisons et la forte tombée de pluie. Donc, pour finir, un modèle sera estimé à partir des variables pluviométriques et températures.

1. Matériels et Méthodes

L'article discutera des estimations de la tendance, de recherche d'un modèle prévisionnel et des détections des différents changements dans des séries temporelles.

Les détections des ruptures d'une série ou les changements des points qui sont étudiés depuis très longtemps sera l'objectif principal de l'article. La détection de rupture est dans la littérature un moyen de détection d'anomalies en général, les tests d'homogénéité, l'ajustement de courbe ou encore le débruitage [Harle \(2016\)](#).

Dans la littérature, les changements des points sont considérés comme des variables à déterminer dans un modèle linéaire, introduit par [BAI et PERRON \(2003\)](#). L'algorithme est implémenté dans **R** avec le *package strucchange*.

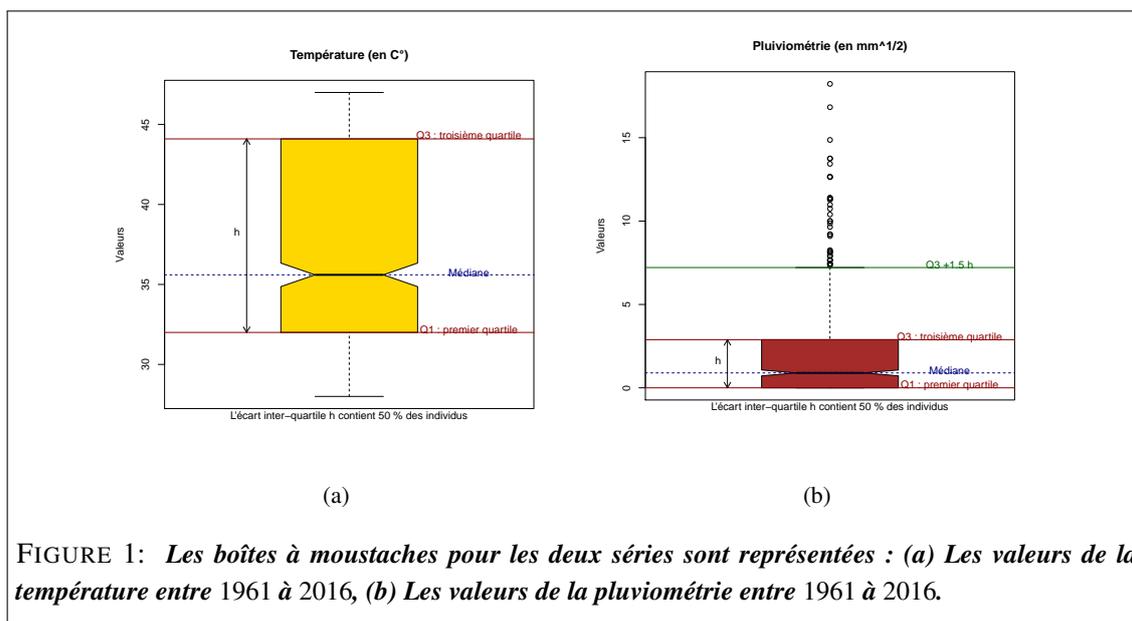
Nous allons aussi utiliser le *package changepoint* pour réaliser des détections en moyennes et voir si les modèles introduit par [BAI et PERRON \(2003\)](#) sont conformes.

L'étude des changements de points dans les séries dépendra de la tendance estimée pour les deux séries. A cet effet, nous commencerons par tester l'existence d'une tendance dans les différentes séries en se basant sur le test de **Mann Kendall**. Puis une régression locale sera effectuée pour l'estimation des tendances avec différents lissages. Et nous étudierons un modèle prévisionnel pour les deux séries.

1.1. Présentation des données :

Les données évoluent entre 1961 et 2016. Nous nous sommes intéressés à deux variables quantitatives : la température et la pluviométrie.

Nous avons représenté les boîtes à moustaches des deux séries dans la **Figure [1]**. Les valeurs de la températures varient de 25 à 45 celsius, la plupart des valeurs de la température présentés dans **Figure[1](a)** dépassent les 36 celsius en moyennes. Puis le boîte à moustache des valeurs de la



pluviométrie sont représentés dans la **Figure [1](b)**.

La plupart des valeurs sont égales à 0 et la première quartile est aussi égale à 0. Même la médiane est proche de 0 et le troisième quartile est égal à 3. Ce qui démontre la rareté de la pluie à Djibouti-ville.

1.2. Estimation d la tendance

L'estimation de la tendance dépendra de la stationnarité des séries. Pour tester cela nous utiliserons un test de **Ducky Fuller Augmenté** (ADF) testant une hypothèse nulle H_0 : *La série admet une racine unitaire* contre l'hypothèse alternative H_1 : *La série n'admet pas de racine unitaire* pour plus de détail à consulter [ARAGON \(2011\)](#).

Nous combinons ce test à un test non paramétrique afin de détecter la présence d'une tendance dans les séries. Nous parlons du test de **Mann Kendall**, très souvent utilisé dans la détection des tendances monotones dans la série d'étude.

Le test est basé sur deux hypothèses qui sont : H_0 : « pas de tendance dans la série » contre H_1 : « la série présente une tendance » d'après [Xiong et Guo \(2004\)](#). La statistique de test est calculé comme suit :

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sign}(X_j - X_k) \quad (1)$$

Avec,

$$\text{sign}(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -1 & \text{si } x < 0 \end{cases} \quad (2)$$

Soumis au Journal de la Société Française de Statistique

File: Soumission_text_long_ABDI_KHAIRE_Mohamed_Environnement_Climat_SFDS_2019.tex,

compiled with jsfds, version : 2018/06/13

date: 15 avril 2019

Par la suite, nous estimerons la tendance existante dans les séries grâce à une régression locale. La régression locale, ou LOESS, est une méthode de régression non paramétrique fortement connexe qui combine plusieurs modèles de régression multiple.

Nous allons varier la valeur de la fenêtre de lissage h – plus le h est grand. Plus le lissage de la série s’approche d’une droite ou d’une courbe. Dans [Kitagawa et Gersch \(1984\)](#), la fenêtre de lissage est déterminée grâce à une validation croisée pour estimer la tendance et la saisonnalité.

1.3. Modèle de détection de rupture

Nous souhaitons estimer une rupture inconnue dans les séries. Un premier test, appelée F statistique consiste à détecter une potentielle rupture dans la série. Le calcul F-test est lié aux estimateurs OLS et il est défini par :

$$F = (RSS - ESS) / (ESS / (n - 2k)), \quad (3)$$

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2.$$

[Zeileis et al. \(2003\)](#) n : taille de l’échantillon, k : le nombre des régresseurs dans le modèle \hat{y} : correspond à l’objet contenant les valeurs prédites. \bar{y} : correspond à la moyenne de la variable réponse. F est la statistique de test calculée pour une détection, aussi appelé le test **sup-Wald** par [Bai \(1997\)](#). Ce test se base sur les paramètres RSS (en anglais **R**estricted **S**um of **S**quare) et le ESS (en anglais **E**rror **S**um of **S**quares). Le F statistique est utilisé pour estimer une seule rupture dans la série et elle est détectée par l’existence d’une valeur maximale dans les valeurs de F (un pic) [Andrews \(1993\)](#).

$$S_t(k) = \operatorname{argmax}(F) \quad (4)$$

où t est le temps et k la rupture potentielle détectée par le maximum du test. Nous avons combiné ce test avec celui de **Pettit’s test** expliqué dans [J.E.PATUREL et al. \(1996\)](#). Il teste H_0 : « pas de changement » contre H_1 : « l’existence de changement ». La statistique non-paramétrique est définie par :

$$K_t = \max |U_{t,T}| \quad \text{avec} \quad U_t = \sum_{i=1}^t \sum_{j=t+1}^T \operatorname{sign}(X_i - X_j)$$

$1, \dots, T$ correspond aux instants de la série d’étude X_t et K_t fournit la statistique de test avec l’instant de la rupture t .

Par la suite, le test F sera étendu par [BAI et PERRON \(2003\)](#) qui teste 0 contre l points de ruptures et l contre $l + 1$ points de ruptures, cette procédure d’hypothèse est discutée dans [Bai \(1997\)](#). [Bai en 1994](#), a introduit le test ou l’évaluation des écarts par rapport à la stabilité dans le modèle de régression linéaire classique.

$$Y = \beta X_i + \varepsilon_i. \quad (5)$$

Pour étendre à un cas multiple, [Bai et Perron en 1997](#) [BAI et PERRON \(2003\)](#) et [Zeileis et al. \(2003\)](#) supposent l’existence de m points ruptures où les coefficients passent d’une relation de régression stable à une autre. Ainsi, il y a $m + 1$ segments dans lesquels les coefficients de régression sont constants, et le modèle peut être réécrit comme :

$$Y_{i_j} = \beta X_{i_j} + \varepsilon_{i_j} \quad (6)$$

où j désigne l'indice de segment. En pratique, les points de rupture i_j sont rarement donnés de manière exogène, mais doivent être estimés. Pour plus de détails sur ces régressions à voir [HORVATH \(1995\)](#), [Christiano \(1992\)](#), [BAI et PERRON \(2003\)](#) et [Bai \(1998\)](#).

2. Résultats

2.1. étude de la température de Djibouti de 1961 à 2016

Un réchauffement climatique est enregistré dans la planète ces dernières décennies. La température de Djibouti est toujours élevée et varie en moyenne de 30 à 45°C.

Par rapport au monde, Djibouti n'a que deux saisons : une saison fraîche et une chaude. La durée de la saison chaude est plus longue que celle de la saison fraîche.

La population Djiboutienne est familiarisée avec ce climat sec et aride. Et cette population est facilement susceptible de tomber malade lors d'une forte baisse de la température. Pendant les saisons fraîches les nombres d'épidémies de pathologie inconnue (comme les fièvres d'origines inconnues) émergent.

A cet effet, nous étudierons dans cette section l'évolution de la température de Djibouti pendant la période 1961 à 2016 afin de détecter une quelconque variation de la température au cours de cette période.

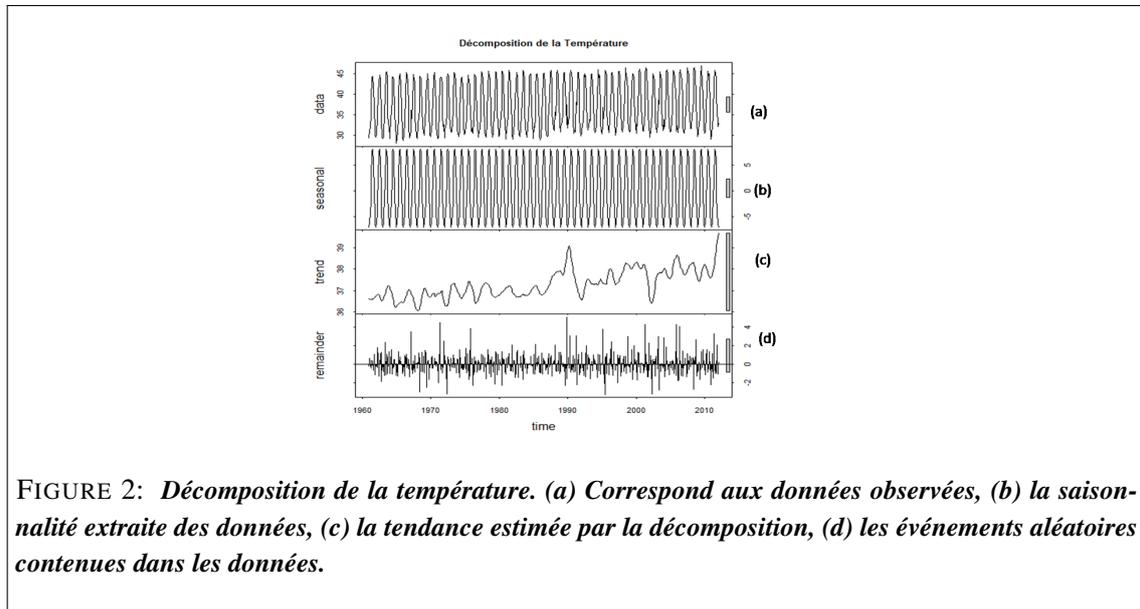
Cette étude temporelle sera suivie par une prévision sur les années. Nous finirons l'étude par une analyse de détection de rupture – détecter la période où une hausse ou une diminution de température est enregistrée, comme dit dans [Harle \(2016\)](#) « détecter les changements dans les caractéristiques du signal et les localiser ».

2.1.1. études temporelles

L'évolution de la série est présentée dans la **Figure[3]**, l'unité de mesure est en degré celsius. Nous pouvons aussi voir de la saisonnalité dans les données. Djibouti est un pays situé à l'équateur, le pays a simplement deux saisons – une saison chaude et une saison fraîche.

Dans la saison chaude les valeurs en degré celsius atteignent plus de 45 et dans la saison fraîche le degré ne baisse pas plus de 25 en moyenne. La **Figure[3]** ne nous montre pas l'évolution de la tendance au cours de ces 50 ans. Un test de stationnarité a été réalisé sur la série avec un test de **Duckey Fuller Augmenté**. Ce dernier confirme la stationnarité de la série avec un risque de 5%. En appliquant le test sur la série d'étude, avec un risque de 5% nous avons l'existence d'une tendance dans la série avec une statistique de test égale à 15056. Nous confirmons la non stationnarité de la série de température, la série admet une tendance et peut aussi admettre un effet saisonnier.

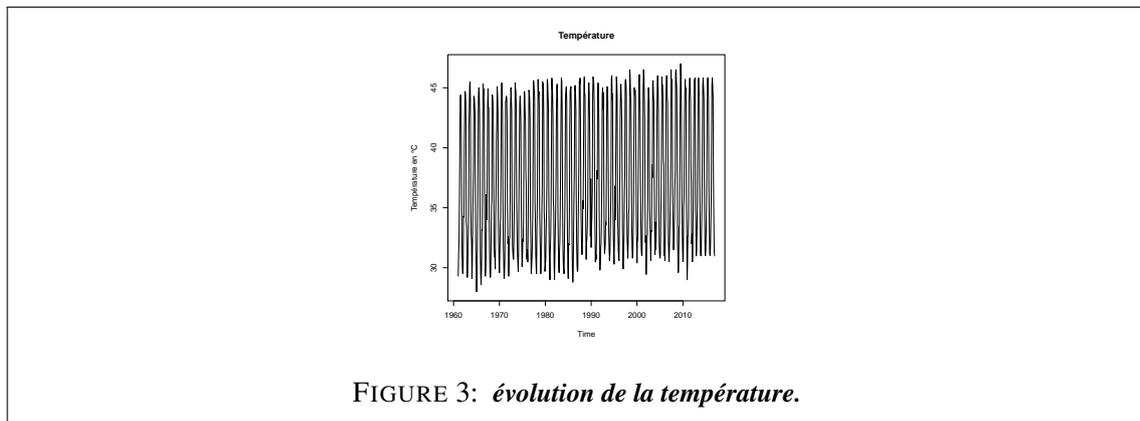
Une décomposition de la série est réalisée pour estimer la saisonnalité et la tendance qui est illustrée sur la **Figure [2]**.



La décomposition de la série est représentée dans la **Figure[2]**. La série est décomposée en trois parties : la saisonnalité, la tendance et les résidus (ou les variations accidentelles).

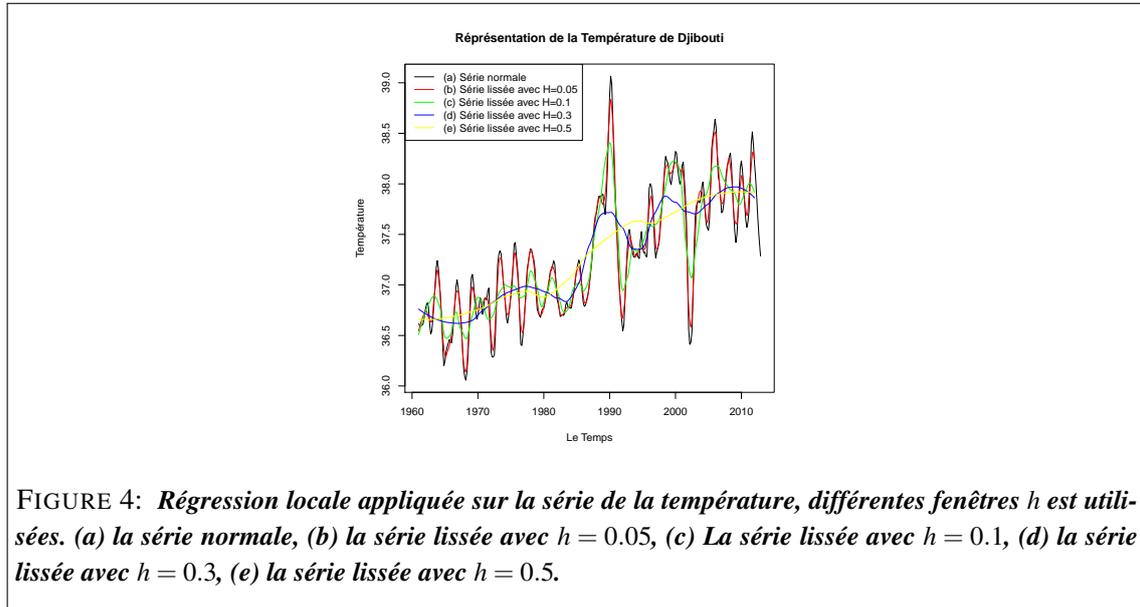
La tendance démontre une croissance de la température à Djibouti ville. Les plus fortes températures sont enregistrées en 1990. Dans l'estimation de la tendance, nous observons la persistance de la saisonnalité.

Pour éliminer cette saisonnalité de notre estimation et lisse de plus en plus la série, nous utiliserons la régression locale et allons varier le filtre h .



Nous utiliserons la tendance estimée par la décomposition est utilisée pour le lissage de la régression locale.

Nous utilisons une régression locale et différents lissages sont réalisés en changeant à chaque fois la fenêtre h , ces estimations sont visualisées dans la **Figure[4]**.



La tendance démontre une croissance de la température à Djibouti ville. Les plus fortes températures sont enregistrés en 1990. Dans l'estimation de la tendance, nous observons la persistance de la saisonnalité.

Pour éliminer cette saisonnalité de notre estimation et lissée de plus en plus la série, nous augmentons la filtre h . En augmentant de plus en plus la fenêtre, l'estimation se rapproche d'un polynôme avec un $h = 0.5$.

La température de Djibouti n'a cessé d'augmenter au fil du temps – signe du réchauffement climatique. Nous allons réaliser une prévision sur les années à venir de la température après 2016. Plusieurs modèles ont été comparés sans faire une différenciation de la série. Des opérateurs de différenciations sont appliqués sur la série dans le but de la rendre stationnaire (différenciations) et aussi pour éliminer la tendance. La différenciation est définie par :

$$\Delta^r X_t = (X_t - X_{t-1})^r \quad (7)$$

le degré de la différenciation r dépend du degré du polynôme de la tendance. Une tendance de degré r est éliminée par $r + 1$ différenciation. Le modèle optimal qui s'ajuste le plus à notre série est un $SARMA(1,0,5)(0,0,2)$ avec une périodicité de 12. Un processus y_t est un $SARIMA(p,d,q)(P,D,Q)_s$ s'il obéit à :

$$(1-B)^d(1-B^s)^D y_t = c + \frac{(1 + \delta_1 B + \dots + \delta_q B^q)(1 + \omega_1 B^s + \dots + \omega_Q B^{sQ})}{(1 - \phi_1 B - \dots - \phi_p B^p)(1 - \psi_1 B^s - \dots - \psi_P B^{sP})}$$

ARAGON (2011), d'où $(1 - B)$ correspond à l'opérateur de différenciation et $\delta, \omega, \phi, \psi$ sont des coefficients à estimer par le modèle.

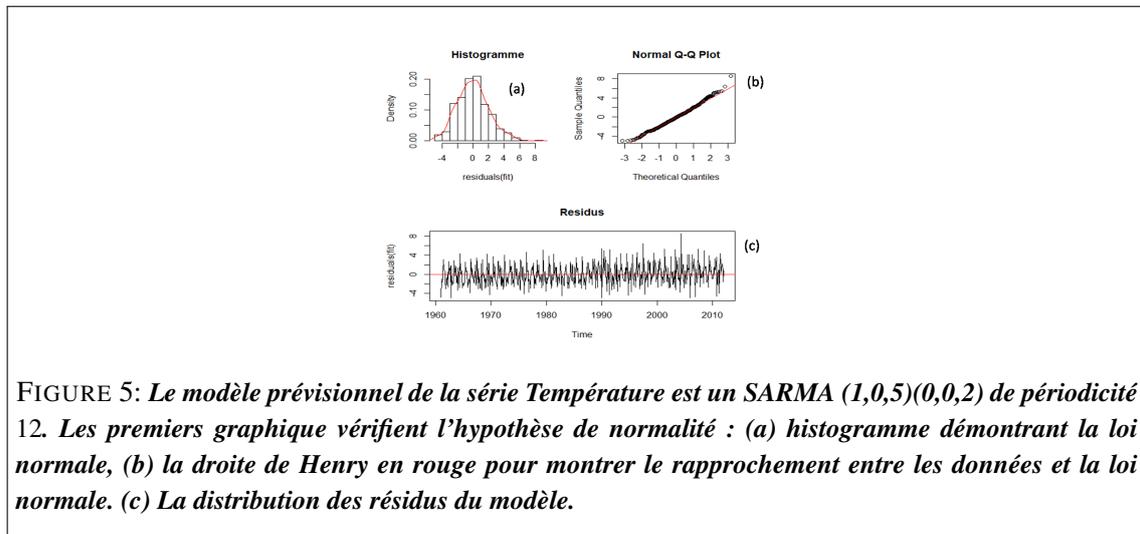
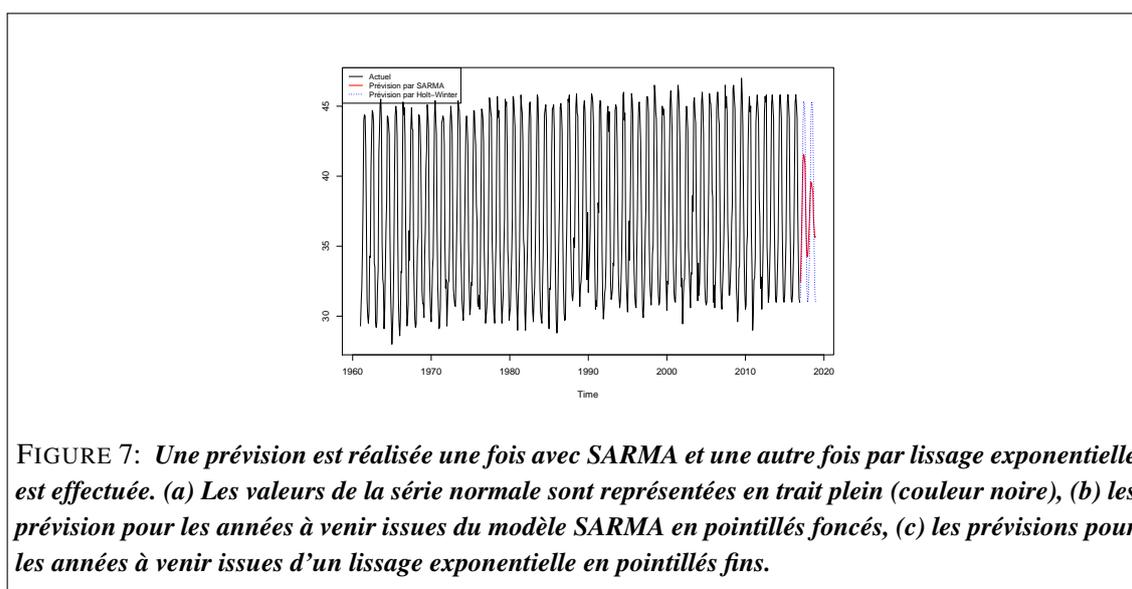
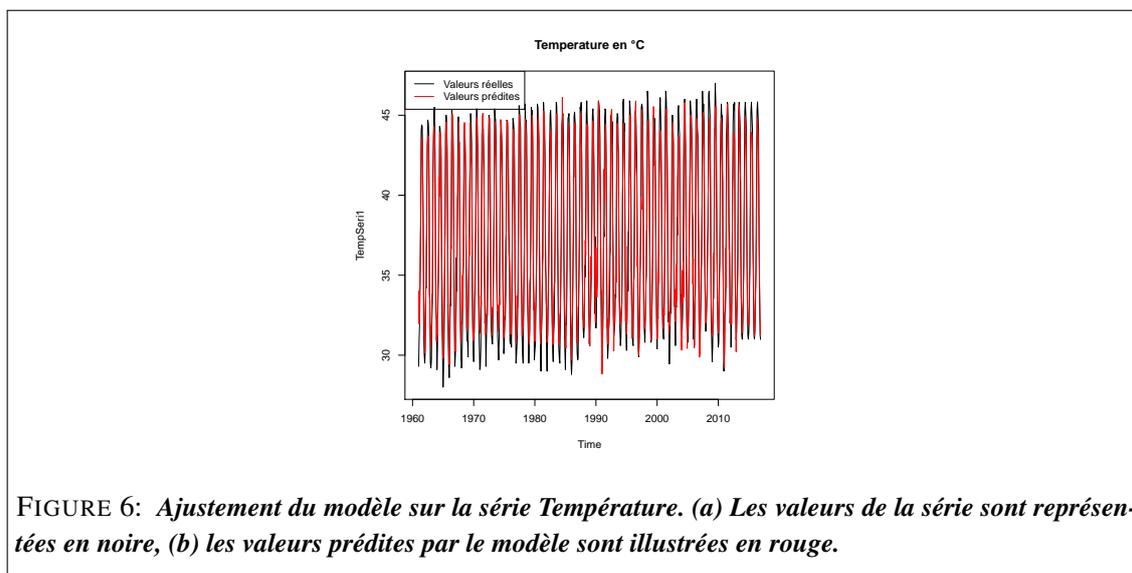


FIGURE 5: *Le modèle prévisionnel de la série Température est un SARMA (1,0,5)(0,0,2) de périodicité 12. Les premiers graphique vérifient l'hypothèse de normalité : (a) histogramme démontrant la loi normale, (b) la droite de Henry en rouge pour montrer le rapprochement entre les données et la loi normale. (c) La distribution des résidus du modèle.*

Par la suite, un diagnostic des résidus a été réalisé. Cette analyse consiste à vérifier la significativité des hypothèses de la normalité, de l'indépendance des erreurs et de l'homoscédasticité. La **Figure[5]** regroupe trois différents graphiques permettant de vérifier les hypothèses citées visuellement.

La **Figure[5]** montre que l'hypothèse de normalité est vérifiée d'après les histogrammes et la droite de Henry. Puis nous observons une distribution presque homogène des résidus, l'hypothèse d'homoscédasticité est vérifiée et l'hypothèse d'indépendance des résidus est vérifiée avec un test de portemanteau. Pour aller plus loin, nous avons regardé si le modèle s'ajuste parfaitement à la série température, pour cela la **Figure[6]** a été illustrée. Elle est une superposition de la série normale et des valeurs prédites par le modèle, l'ajustement est bon donc le modèle explique bien notre série.

Après vérification des différentes hypothèses, nous passons à la prévision sur les années à venir. Nous utilisons le modèle optimal estimé et un lissage exponentielle (Holt-winter) pour faire une prévision. La **Figure[7]** regroupe la prévision des deux modèles sur la série. Nous observons que le lissage exponentielle est influencé par les phénomènes de la tendance et de la saisonnalité de notre série. Tandis que le modèle SARMA prévoit des valeurs compris entre 35 et 40.



La prévision a été réalisée pour les deux prochaines années de la série, c'est à dire pour les années 2017 et 2018. Par la suite, nous choisissons de travailler avec le modèle SARMA pour sa fiabilité à court terme et moyen terme.

2.1.2. étude de la pluviométrie de Djibouti de 1961 à 2016

La pluviométrie est utilisée comme mesure de la sécheresse dans un pays. A cet effet, un suivi sur les 55 dernières années a été réalisée à Djibouti-ville.

Ce phénomène naturel dévient de plus en plus rare pour le pays. Le graphique illustrant l'évolution est visualisé sur la **Figure[9]** Le climat de Djibouti est de type désertique, c'est-à-dire aride et chaud. La pluviométrie y est faible avec moins de 200 mm de pluies par an.

Nous observons des pics très élevés sur certaines années. En 1990, la ville de Djibouti a recensé des valeurs dépassant plus de 200 mm. Mais la plupart des valeurs sont égales à 0. Aucune augmentation ou diminution ou saisonnalité n'est observée dans la série car il y a des fortes écarts entre les valeurs.

Pour étudier la série et chercher un modèle, nous appliquons la racine carrée sur la série afin d'éviter la confusion entre les valeurs très élevées et les valeurs atypiques.

La stationnarité de la série n'est pas vérifiée avec le test de **Ducky Fuller Augmenté**. Nous avons combiné ce test avec celui de **Mann Kendall** pour plus de certitude.

Le test de Mann Kendall rejette l'hypothèse nulle avec un risque de 10%, il affirme l'existence d'une tendance dans notre série avec un $pvalue = 0.08$.

La tendance et la saisonnalité contenue dans la série a été estimée. Ces estimations sont représentées sur la **Figure[8]**.

Nous avons commencé par rechercher la tendance de la série, afin de voir l'évolution de la précipitation de la sécheresse à Djibouti-ville. Pour cette raison, nous utiliserons une régression locale et déterminerons la fenêtre de lissage h .

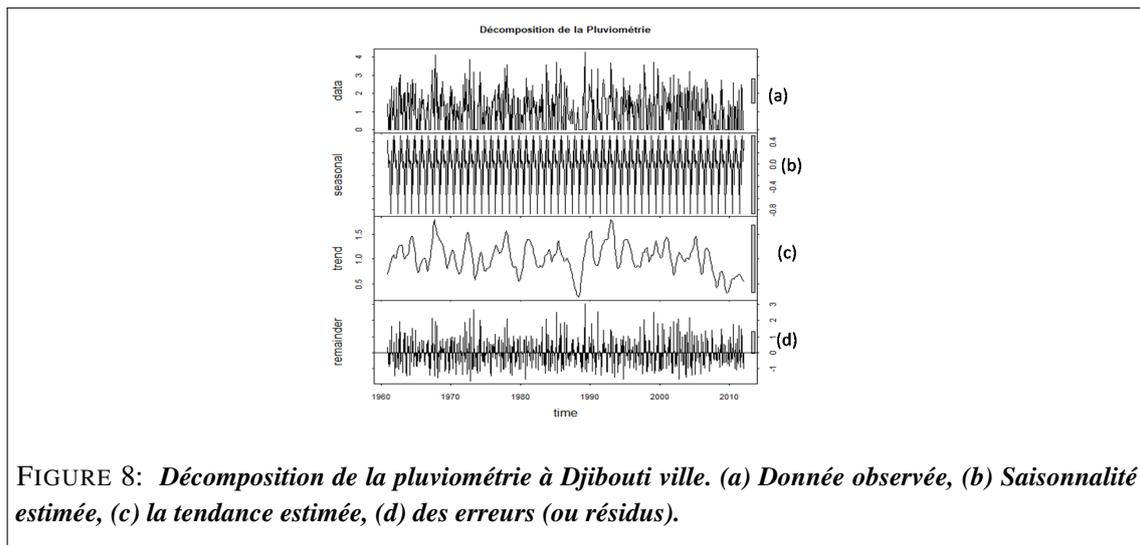


FIGURE 8: *Décomposition de la pluviométrie à Djibouti ville. (a) Donnée observée, (b) Saisonnalité estimée, (c) la tendance estimée, (d) des erreurs (ou résidus).*

Une décomposition est réalisée sur la série en question. Une tendance est estimée et aussi une saisonnalité de la série (voir la **Figure[8]**).

Pour estimer cette tendance, nous utilisons une régression locale et nous varions les fenêtres de lissage, dans le but de voir le lissage le plus approprié. L'estimation de la tendance avec plusieurs fenêtres est visualisée sur la **Figure[10]**. La tendance estimée montre des variations homogènes des valeurs jusqu'à l'année 1989 (où une chute de la pluviométrie est enregistrée). Après on note une reprise de la série avec une diminution progressive qui commence à 2004.

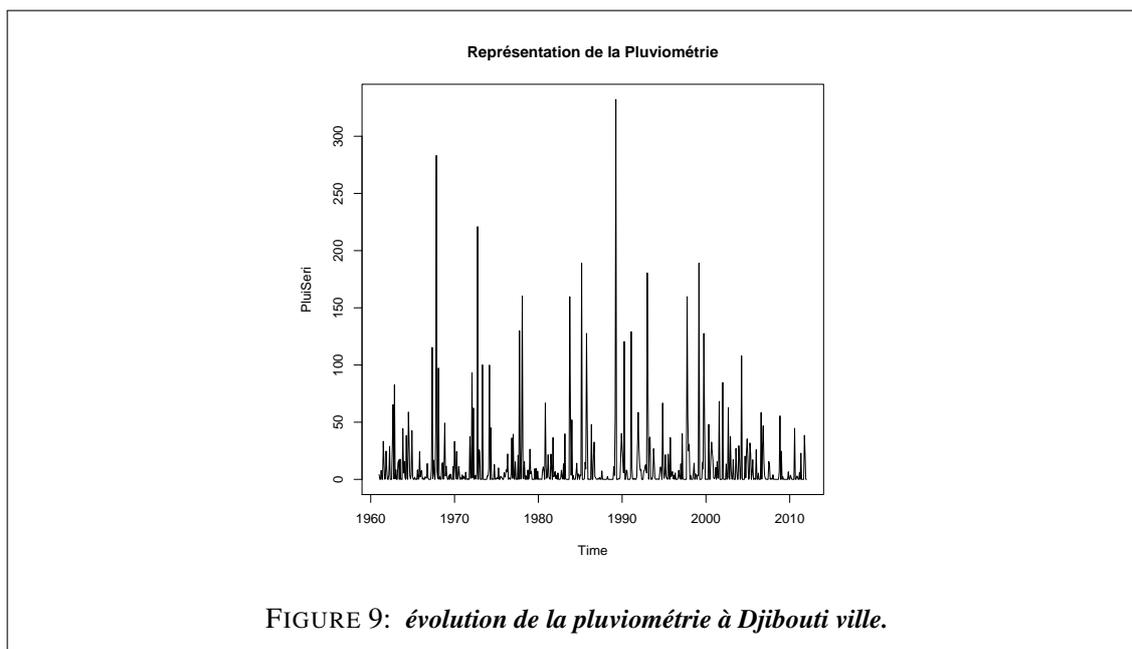


FIGURE 9: *évolution de la pluviométrie à Djibouti ville.*

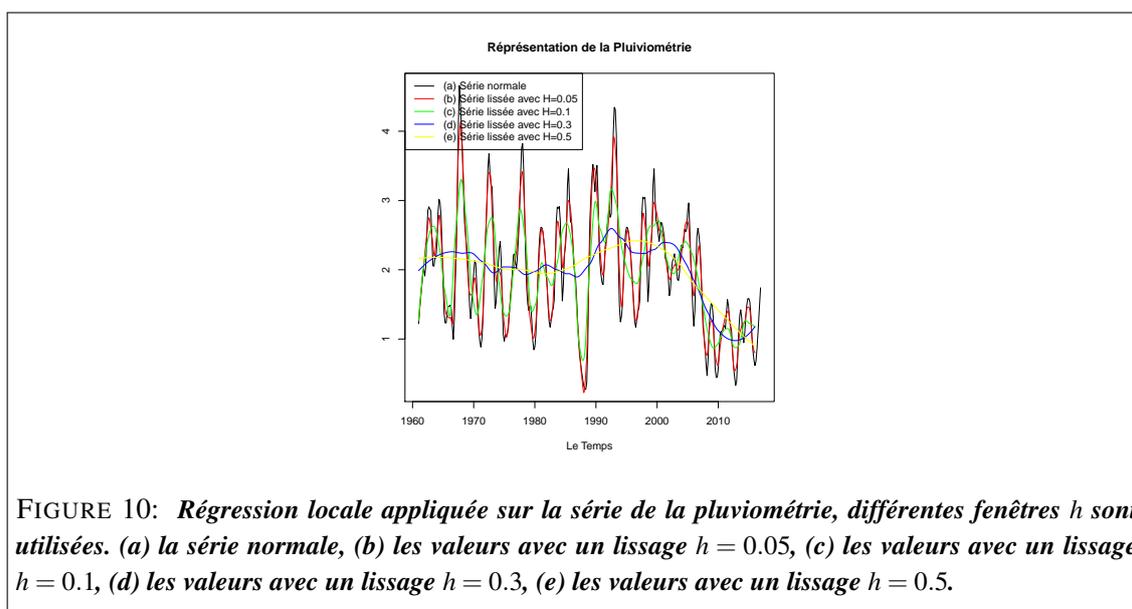


FIGURE 10: *Régression locale appliquée sur la série de la pluviométrie, différentes fenêtres h sont utilisées. (a) la série normale, (b) les valeurs avec un lissage $h = 0.05$, (c) les valeurs avec un lissage $h = 0.1$, (d) les valeurs avec un lissage $h = 0.3$, (e) les valeurs avec un lissage $h = 0.5$.*

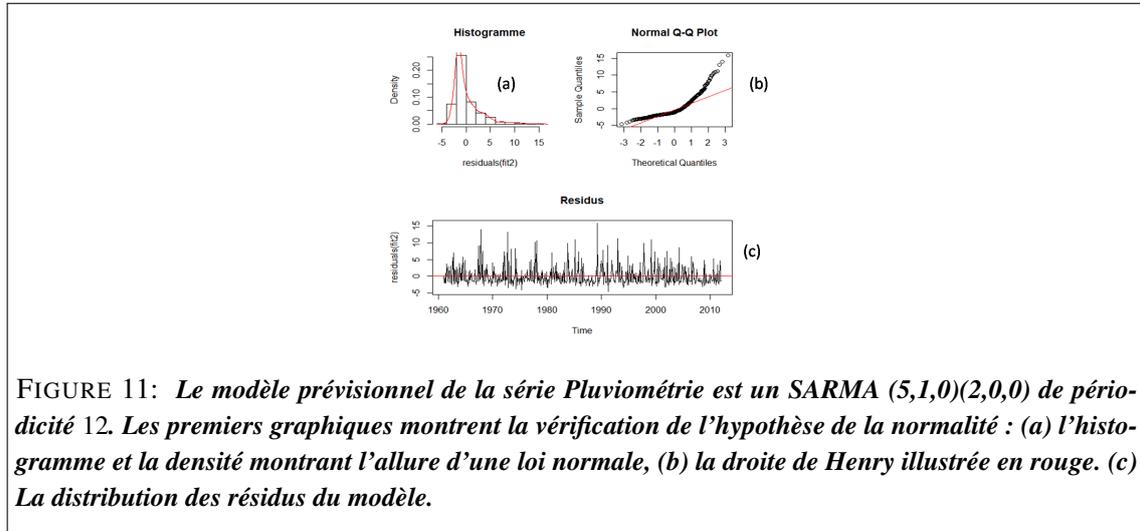


FIGURE 11: *Le modèle prévisionnel de la série Pluviométrie est un SARMA (5,1,0)(2,0,0) de périodicité 12. Les premiers graphiques montrent la vérification de l'hypothèse de la normalité : (a) l'histogramme et la densité montrant l'allure d'une loi normale, (b) la droite de Henry illustrée en rouge. (c) La distribution des résidus du modèle.*

Nous considérons l'estimation de la tendance pour un $h = 0.05$ puisque le lissage n'est pas très grossier. Un $h = 0.5$ donne un polynôme qui diminue au fil des années sauf entre les années 1990 à 2000. Mais après 2000 la diminution de la pluviométrie reprend son cours et cette dernière s'accroît de plus en plus à partir de 2004. Donc avec un $h = 0.5$, nous déduisons la diminution de la pluviométrie au cours du temps à Djibouti-ville.

En utilisant la série appliquée avec la racine carrée nous recherchons un modèle optimal pour la prévision. Le modèle optimal pour l'étude de la pluviométrie est un modèle SARMA(5,1,0)(2,0,0) avec une périodicité de 12. Pour valider ce modèle un diagnostic sur les résidus du modèle est effectué et les résultats sont représentés sur la Figure[11].

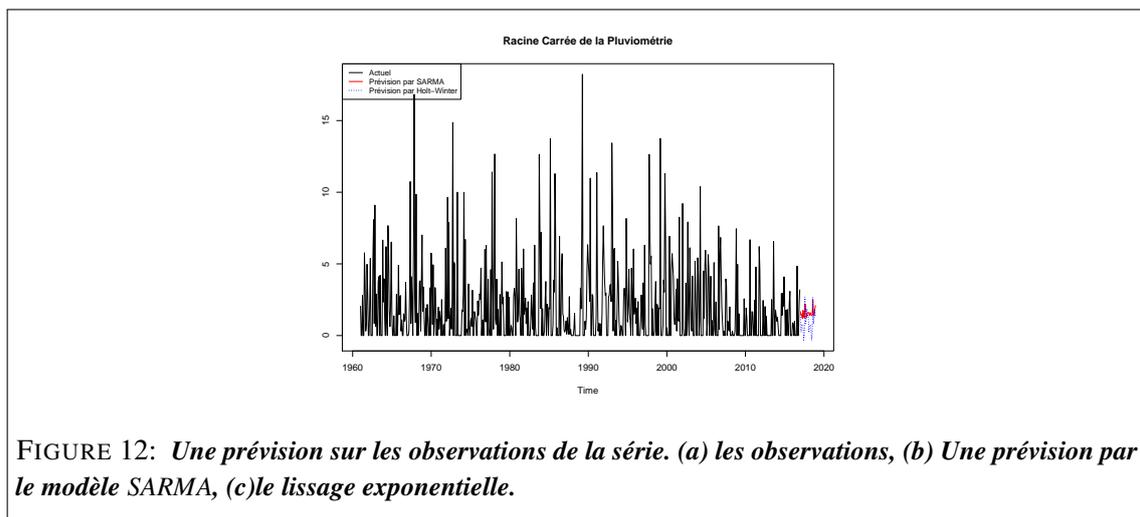


FIGURE 12: *Une prévision sur les observations de la série. (a) les observations, (b) Une prévision par le modèle SARMA, (c) le lissage exponentielle.*

Les résidus du modèle suivent la loi normale et l'homoscédasticité des résidus est vérifiée (les résidus oscillent aux alentours de 0). Les résidus du modèle vérifient les hypothèses. Nous étudions la temporalité de cette série dans le but de faire les prévisions des années à venir

du phénomène étudié. Cependant, il existe plusieurs modèles prévisionnels très intéressants : le lissage exponentiel, ARIMA ou SARIMA et la régression de la tendance.

Pour l'instant, nous nous intéressons au modèle *SARMA* qui donne des valeurs prévisionnelles fiables avec un pourcentage d'erreur très minime pour des instants proches.

Afin de montrer une quelconque comparaison entre un modèle *SARMA* avec les autres modèles, nous avons choisi parmi tous les modèles le lissage exponentiel. Nous comparons les prévisions de ces deux modèles. Ces comparaisons sont visualisées sur la **Figure[12]**.

Les valeurs prévisionnelles du modèle *SARMA* semblent devenir constantes au cours du temps. Et celles du lissage exponentiel ne cessent de diminuer comme les tendances estimées.

2.2. Détection de rupture

2.2.1. Détection de rupture pour la température

La détection de rupture a été très souvent utilisée dans la littérature pour les signaux afin de détecter des réseaux anormaux. La rupture est aussi définie par un changement dans la loi de probabilité des variables aléatoires dont les réalisations successives définissent les séries chronologiques étudiées [J.E.PATUREL et al. \(1996\)](#).

Cette sous-section a pour objet d'identifier des ruptures éventuelles dans les séries chronologiques températures annuelles. Les valeurs de la série normale (température) ne sont pas tous bien lisibles et la représentation de l'évolution de la série est quasi constante à l'œil nu.

Dans ce cas, la tendance estimée par la régression locale avec une fenêtre de lissage $h = 0.05$, sera utilisée dans le but de détecter des quelconques ruptures (comme dans [Xiong et Guo \(2004\)](#)). Nous avons calculé la statistique de test pour les différents instants, les valeurs du test sont visualisées dans la **Figure[13]**.

Le plus haut pic présenté dans la **Figure[13]** est le signalement d'une rupture significative dans les données. Cette dernière est localisée en 1987.

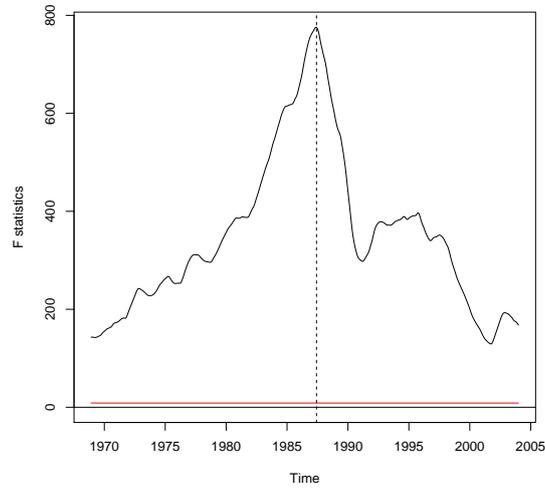


FIGURE 13: *F* statistique pour les différents instants de la série.

Nous avons cumulé cette méthode au découpage homogène moyen sous l'hypothèse que la variance est constante. Les résultats sont illustrés sur la **Figure[14]**.

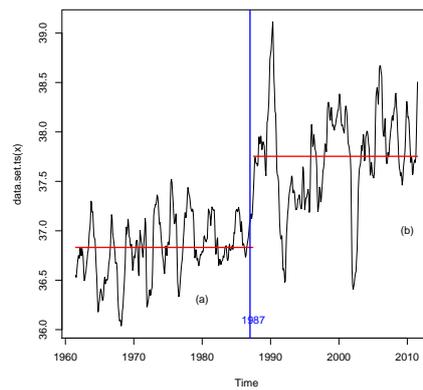


FIGURE 14: *Détections des changements de la série en moyennes.*(a) Correspond à la moyennisation des valeurs avant la rupture, (b) correspond à la moyennisation des valeurs après la rupture.

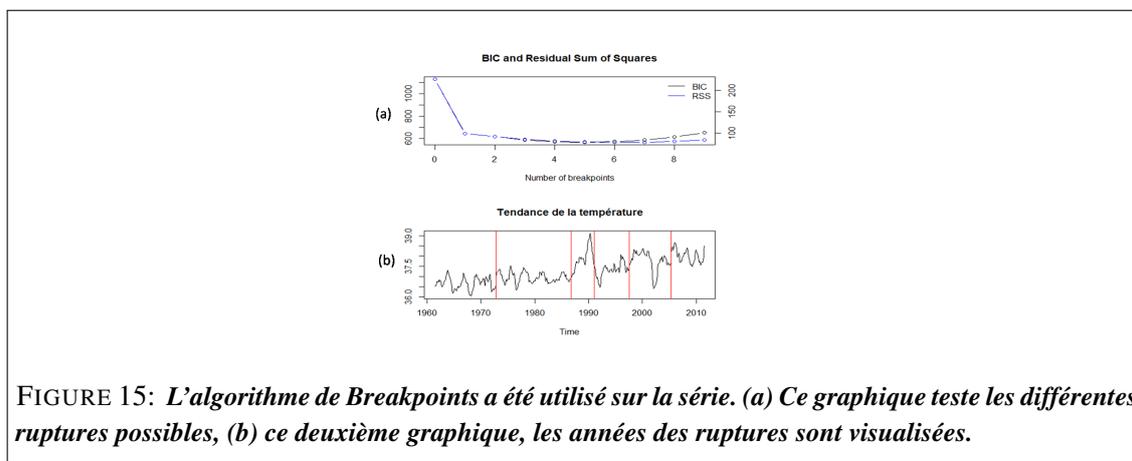


FIGURE 15: *L'algorithm de Breakpoints a été utilisé sur la série. (a) Ce graphique teste les différentes ruptures possibles, (b) ce deuxième graphique, les années des ruptures sont visualisées.*

La série est divisée en deux parties, le découpage se fait au niveau de l'année 1987 comme l'avait estimé le premier test. Ces moyennes nous montrent que le réchauffement climatique débute en 1987 puisque les valeurs de la série prennent un déclin à cette année. Les valeurs moyennes avant la rupture étaient de 36.82 et après la rupture ils atteignent la valeur 37.75.

Cette méthode est très efficace et nous permet de lisser notre série. Un second test est réalisé, pour donner plus de robustesse à l'estimation. Il s'agit du **Pettitt's test**(1979) très souvent utilisé pour détecter une seule rupture sur les séries hydrologiques ou climatiques.

Le test rejette l'hypothèse nulle et retient l'hypothèse alternative avec un risque de 5%. Il confirme la rupture estimée en $t = 1987$ avec un $K_t = 78891$.

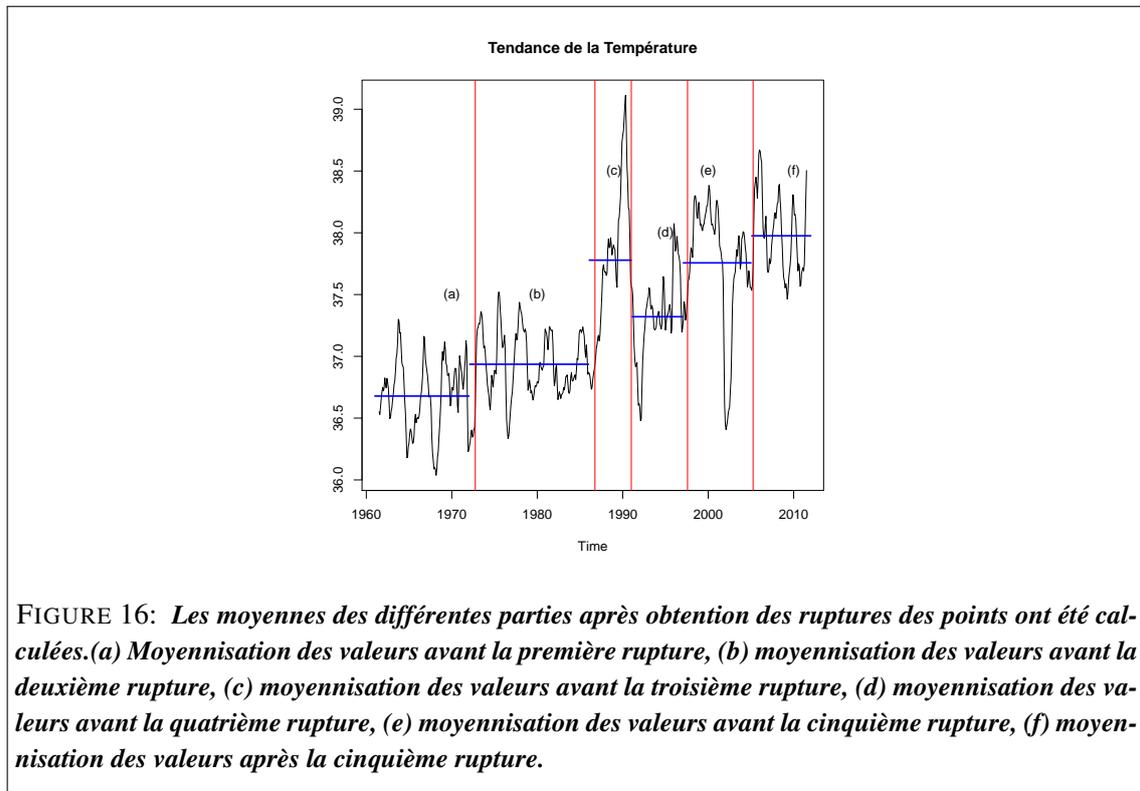
Cependant, l'information est réduite au simple fait qu'une rupture significative est présente dans la série et ne renseigne pas sur toutes les ruptures possibles. Donc pour stimuler ces détections, nous passons à une détection plus grossière et plus informative.

Les points de ruptures sont estimés en minimisant la somme résiduelle des carrés (RSS) et aussi le critère BIC. En utilisant l'algorithm de **Bai** sur la tendance, une régression linéaire avec une constante a été réalisée pour expliquer la tendance de la série.

Pour une première approche, nous avons choisit de travailler avec le plus simple modèle et voir ces résultats. Ces derniers sont affichés sur la **Figure**[15].

L'algorithm détecte 5 points de ruptures. Ces ruptures sont localisés en 1972, 1986, 1991, 1997 et en 2005. Une moyennisation des valeurs pour chaque partie a été visualisée sur la **Figure** [16]. Les moyennes ont fortement augmenté entre 1986 et 1991, atteignent 37.8 degré celsius. Cette croissance brusque est suivie d'une diminution, nous pouvons la considérer comme une reprise normale de la croissance de la température.

Le thermomètre a enregistré des fortes chaleurs à Djibouti pour les années 1986 et 1991.

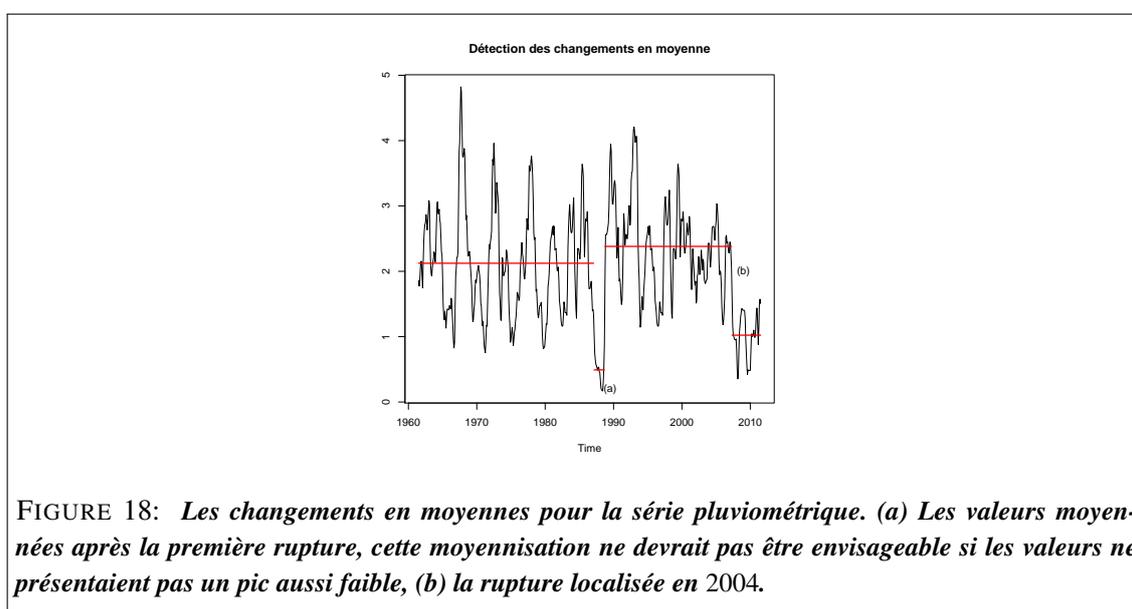
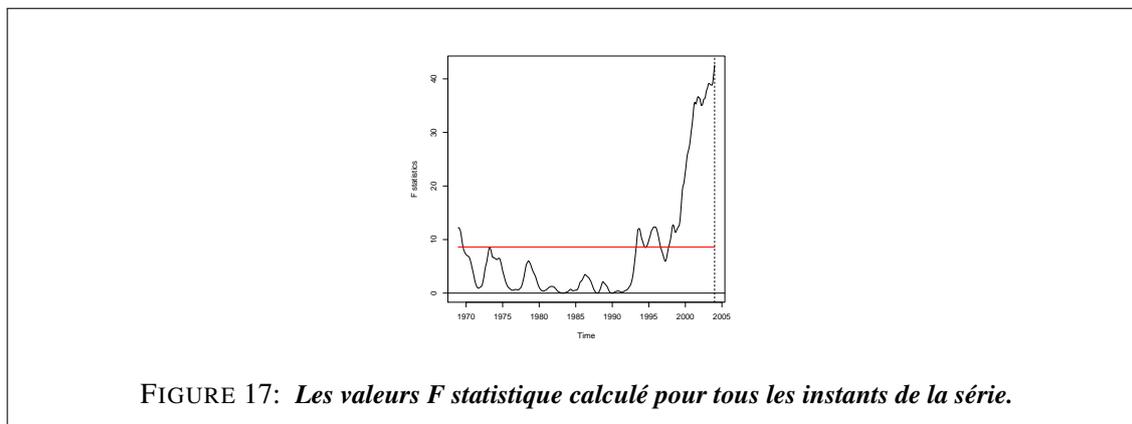


2.2.2. Détection des ruptures pour la pluviométrie

La pluie dévient de plus en plus rare à Djibouti-ville. Mêmes dans les saisons fraîches les quantités de pluies se font rares et dans certaines périodes, la quantité d'eau de pluie est substantielle et provoque des inondations dans la ville.

Nous voudrions connaître les périodes de sécheresse ou les périodes à risques d'inondations. L'étude de la détection de rupture pour cette série sera réalisée comme dans la sous-section de la température.

Pour cette raison, nous commençons par étudier le test F statistique. Ce test détecte une simple rupture dans la tendance de la série. Les valeurs de ce test sont visualisées sur la [Figure\[17\]](#).



Le pic est situé à l'instant 2004, cela signifie que nous avons une rupture à cette date. Nous pouvons même confirmer cette rupture à l'œil nu dans l'estimation de la tendance. Nous apercevons que les valeurs de la pluviométrie commencent à diminuer à partir de l'année 2004. Le maximum de ce test correspond à la valeur minimale des estimateurs RSS et ESS .

Pour plus de certitude, le **pettitt's test** a été réalisé sur la tendance de la série. Le test a confirmé avec un risque de 5%, l'existence d'une rupture en 2004 avec un $K_T = 23216$.

Cependant, la détection de changement en moyenne n'est pas du même avis. La **Figure[18]** représente les détections des ruptures en moyennes de la série. La **Figure[18]** représente en (a) les valeurs moyennées après la première rupture et en (b) la rupture localisée en 2004.

Le changement moyen a détecté 3 différentes ruptures. La première en 1986, la seconde est en 1988 et la troisième en 2004. Nous supposons, que cette détection est influencée par les valeurs très faibles de la période comprise entre 1986 et 1988, les valeurs de la pluviométrie sont les plus basses.

Mais, les moyennes calculées pour ces différentes périodes sont égales. L'algorithme de chan-

gement des moyennes aurait détecté une seule rupture en conformité avec les tests précédents sans les faibles valeurs enregistrées entre 1986 et 1988 (a). Et la seule rupture devrait être à (b) l'année 2004.

Par la suite, nous avons supposé l'existence de m ruptures dans la série. En s'aidant de l'algorithme de Bai et de Perron (1997) et de la régression de l'équation (6), nous avons réalisé une détection de rupture en expliquant la tendance de la série par une constante.

L'estimation est affichée dans la Figure [19]. L'algorithme estime 14 points des ruptures avec un $RSS = 199.6$ et $BIC = 1235.1$. Ces points sont situés à 1964, 1966, 1969, 1971, 1974, 1976, 1979, 1983, 1986, 1988, 1993, 1997, 2001 et 2007.

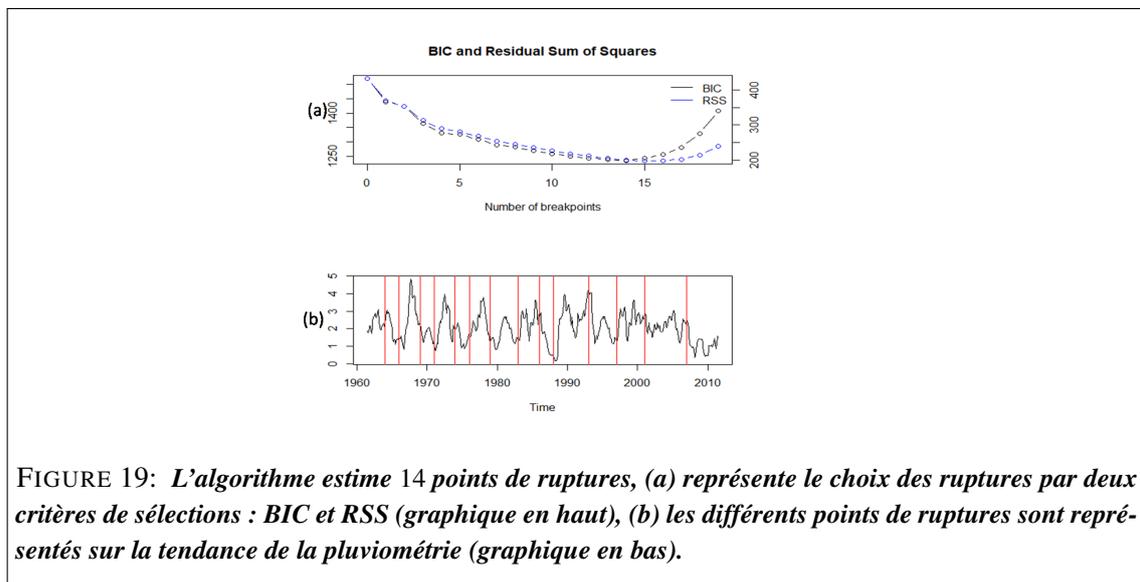


FIGURE 19: *L'algorithme estime 14 points de ruptures, (a) représente le choix des ruptures par deux critères de sélections : BIC et RSS (graphique en haut), (b) les différents points de ruptures sont représentés sur la tendance de la pluviométrie (graphique en bas).*

A chaque variation de la tendance l'algorithme nous détecte une rupture.

En moyenne, la pluviométrie démontrait une certaine monté des quantités d'eau de 1990 à 2000. A partir du XXI^{ème} siècle la quantité de pluie n'a cessé de diminuer et nous avons observé sur les différentes Figures un fort décrochage en 2004.

3. Modélisation et lien avec la fièvre

La modélisation mathématique a pour objectif de définir plusieurs modèles pour expliquer un phénomène et dans notre cas ; il est question de déterminer un moyen de visualiser et de comprendre l'évolution selon les années des fièvres par l'étude des facteurs climatologiques.

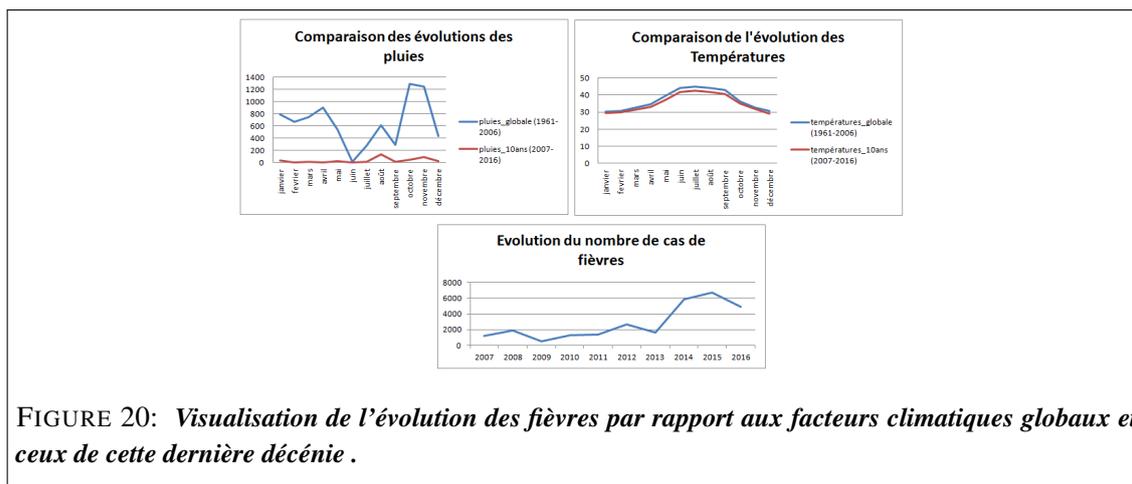


FIGURE 20: *Visualisation de l'évolution des fièvres par rapport aux facteurs climatologiques globaux et ceux de cette dernière décennie .*

mois	% pluviométrique (2007-2016) Vs (1961-2006)	Ecart Température (2007-2016) Vs (1961-2006)
janvier	4,90038855	-0,628204685
février	1,009546148	-0,910872625
mars	1,822860819	-1,201595622
avril	0,789151779	-1,705542186
mai	5,978689457	-2,127920507
juin	0	-2,622440476
juillet	7,116794888	-2,168168203
août	22,31839561	-2,209179954
septembre	5,598462109	-2,238940448
octobre	4,254134245	-1,219607719
novembre	7,88537979	-0,777947024
décembre	7,25702794	-1,412321429

FIGURE 21: *Comparaison des facteurs climatologiques globaux et ceux de cette dernière décennie.*

Nous avons créé deux variables indicatrices, correspondant aux variables températures et pluviométrie et pour la création de ces indicatrices, nous nous avons utilisé les différentes ruptures détectées (Figure [22]).

$$I_{Temp} = \begin{cases} 1 & \text{si Existence d'une rupture pour la température} \\ 0 & \text{sinon} \end{cases} \quad (8)$$

$$I_{Pluvio} = \begin{cases} 1 & \text{si Existence d'une rupture pour la pluviométrie} \\ 0 & \text{sinon} \end{cases} \quad (9)$$

FIGURE 22: *Description des Indicatrices associées à la températures et à la pluviométrie.*

Modèle de régression	R^2 -Ajusté
$Fievre = Temperature \times 1_{Temp} + Pluviometrie \times 1_{Pluvio} + 1_{Temp} \times 1_{Pluvio}$	0.9497
$Fievre = Temperature \times 1_{Temp} + Pluviometrie \times 1_{Pluvio}$	0.796
$Fievre = Temperature + Pluviometrie$	0.2789
$Fievre = 1_{Temp} \times 1_{Pluvio}$	0.409
$Fievre = Temperature + Pluviometrie + 1_{Temp} \times 1_{Pluvio}$	0.5376

TABLE 1. Regroupement des possibilités de modèles.

Variabes	Coefficients estimés	Pr(> t)	Significativités
Intercept	-632195.2	0.0301	*
Température	17035.5	0.0295	*
1_{Temp}	598688.6	0.0445	*
Pluviométrie	-2567.6	0.0498	*
1_{Pluvio}	1763.1	0.4609	
$Temperature \times 1_{Temp}$	-16026.3	0.0441	*
$Pluviometrie \times 1_{Pluvio}$	555.3	0.7419	
$1_{Temp} \times 1_{Pluvio}$	-2993.2	0.0859	.

TABLE 2. La significativité des variabilités.

On observe d'après la **Figures [21]** et la **Figures [20]** que l'on assiste depuis une dizaine d'années à un changement climatique assez évident et c'est d'autant plus intéressant car on voit une évolution des fièvres qui lui est associée **Sanitaire (2015)**.

Par la suite, nous utiliserons la régression linéaire pour expliquer les cas des fièvres d'origines inconnus par la variation des phénomènes temporelles : Température et pluviométrie.

Pour des raisons de précisions nous considérons les 10 dernières années. Les données des fièvres sont recensées pendant la période 2007 à 2016.

Donc pour équilibrer les données, nous avons effectué une moyennisation annuelle sur les phénomènes temporelles. Avant de moyenniser les données températures et pluviométriques, nous avons fait un récapitulatif sur les ruptures.

Les données de la température ont montré 5 ruptures qui sont localisés aux instants : 2008, 2009, 2010, 2011 et 2013. Et ceux de la pluviométrie sont au nombre de 7, localisés en 2008, 2009, 2010, 2012, 2013, 2014, 2015.

Les indicatrices sont ensuite utilisées dans les différents modèles présentés dans le **Tableau [1]**.

Nous avons testé toutes les combinaisons possibles entre les différentes variables citées dans le **Tableau [1]**. Les qualités des différents modèles sont la plupart satisfaisantes avec une qualité plus que la moyenne. Le premier modèle admet une qualité presque parfaite et la plupart de ces variables sont significatives avec un risque de 5%.

le modèle est globalement bon d'après le test de **Fisher** avec un risque de 5%. La fièvre à Djibouti ville est expliquée par la température aux moments correspondants les ruptures et la pluviométrie.

Mais les ruptures pluviométriques n'ont aucun lien avec la fièvre. Cela est expliqué par la non significativité de cette variable. Ce modèle est suivi par le second, qui admet une qualité approximativement égale à 80%.

Il est significativement bon d'après le test de **Fisher** avec un risque de 5%. Nous jugeons bon de

travailler avec le premier modèle qui démontre une certaine logique.

Le nombre des cas des fièvres d'origines inconnus augmente avec les saisons. La saison fraîche enregistre plus des cas par rapport à l'autre saison. Par exemple, le nombre des cas de KF augmente à cause des fortes retombées de pluies et de la faible température [KHAIREH et FARAH \(2012\)](#). La plupart des maladies vectorielles sont influencées par le climat (l'humidité, température,...) [Graux \(2011\)](#).

4. Discussions et Perspectives

La tendance de la température a démontré une croissance du degré de la température à partir de l'année 1987 qui était une année de rupture.

La rupture est montrée par un lissage plus raffiné de la température visualisé sur la **Figure [4]** avec un $h = 0.5$, nous apercevons une augmentation de la température à partir de l'année 1987. Et montre aussi une variation de 5.55% de la température entre 1987 à 2012.

Nous pouvons aussi voir plusieurs détections sur la **Figure [15]** pour avoir un aperçu sur l'évolution de la température pendant la période d'étude. Les moyennes ont fortement augmenté entre 1986 et 1991, atteignant 37.8 degré celsius.

Cette croissance brusque est suivie d'une diminution ou nous pouvons la considérer comme une reprise normale de la croissance de la température. Ces détections de ruptures peuvent nous ramener à un lissage de la tendance plus affiné que celle de la régression locale comme réalisée dans le rapport de [Dion et al. \(2015\)](#).

Le modèle prévisionnel est aussi un atout pour suivre l'évolution de la température au fil des années et permettra de faciliter la tâche de la récolte et de la prévision auprès des services météorologiques. Le modèle pluviométrique vérifiait aussi toutes les hypothèses de validité.

La pluie se fait de plus en plus rare ces dernières décennies. La tendance obtenue par la régression locale nous démontre que la diminution de la pluviométrie est devenue progressive à partir des années 2000, voir la **Figure[10]** avec un $h = 0.5$ nous observons une décroissance de plus en plus accrue.

En réalisant quelque test d'unicité de rupture nous avons remarqué que la rupture la plus significative se trouve en 2004. Cependant, un modèle multiplicatif de rupture détecte plus de 10 ruptures. Le modèle est influencé par les changements brusques des valeurs enregistrées.

Mais, un fort lissage de la série peut réduire le nombre des ruptures détectées. Nous nous sommes intéressés au même algorithme pour l'estimation des ruptures après diminution des données en 10 an. La tendance est devenue plus lisse pour la pluviométrie et pour la température et l'algorithme a pu détecter plusieurs ruptures en ces 10 ans.

Nous nous focalisons sur ces ruptures afin de créer des variables indicatrices et ayant utilisé pour la modélisation. Le modèle consistait à expliquer le nombre total des cas des fièvres d'origines inconnues par la variable température et pluviométrie pénalisée par les indicatrices issues de l'existence des ruptures.

Nous avons testé plusieurs modèles et le meilleur modèle expliquait à 90% le nombre des cas. A cause du manque d'infrastructure, une forte tombée de pluie laisse certains quartiers inondés ou crée des nids de propagation de certaines maladies et cela a des répercussions sur l'hygiène des habitants et la santé.

Soumis au Journal de la Société Française de Statistique

File: Soumission_text_long_ABDI_KHAIRE_Mohamed_Environnement_Climat_SFDS_2019.tex,

compiled with jsfds, version : 2018/06/13

date: 15 avril 2019

Conclusion

Un réchauffement climatique est observé à Djibouti-ville, le climat a fortement augmenté depuis 1987. La tendance de la température est croissante au fil du temps.

Nous observons aussi une sécheresse de plus en plus accentuée à Djibouti-ville. Les quantités des pluies sont faibles. La tendance estimée pour la pluviométrie démontre aussi une décroissance de la pluviométrie à Djibouti-ville.

Et cette décroissance est progressive en 2004. Tandis que cette année là, la ville de Djibouti a connu une inondation des plus sévères ayant causé plusieurs décès.

L'étude du climat est intéressante pour un petit pays en voie de développement comme Djibouti-ville. Chacune des séries étudiées dans cet article a un impact sur l'environnement sanitaire.

En commençant par la pluviométrie qui est un effet déclencheur de certaines maladies tropicales liées à l'hygiène. Nous manquons d'infrastructures dans la plupart des quartiers donc une forte pluviométrie est annonciatrice des maladies tropicales ou vectorielles comme les fièvres d'origines inconnues.

Et dans les périodes de sécheresse, le manque d'eau amène un manque d'hygiène et nous pouvons avoir l'émergence des certaines maladies dûes au manque hygiènes.

Puis, les excès de chaleurs provoque des déshydratation, des tensions et des canicules. Et pour les saisons fraîches, nous avons l'émergence des maladies telle que la grippe, le paludisme, le choléra et la diarrhée – ces deux maladies souvent observés après beaucoup de pluie et une faible température.

Les nombres des cas des malades des fièvres d'origines inconnues ou autres peuvent augmenter avec ces deux séries. Ces analyses seront un plus pour le domaine sanitaire ainsi que les domaines sécuritaires, financiers et économiques.

Ces modèles permettront de faciliter la suivie de l'évolution la température et de la pluviométrie, aider même dans les prévisions.

Ces modèles renforceront la sécurité sanitaire à Djibouti ville. Lors des saisons fraîches, Djibouti connaît des fortes évolutions des cas de paludisme dans la plupart des quartiers. C'est un moyen pour aider les responsables des santés à mettre en place des mesures des préventions.

Remerciements

Nous remercions les personnels du ministère de la santé qui nous ont permis de récolter le maximum des données possible ainsi que les personnels du ministère du fond d'habitation et d'aménagement.

Un grand merci aux personnels de l'Agence Météorologique de Djibouti pour la récolte régulière des données.

Egalement, nous remercions particulièrement les responsables de l'Université de Djibouti et ceux de l'Université de Clermont Auvergne d'avoir construit ce partenariat fructueux qui a permis la réalisation de ce travail.

Références

- ANDREWS, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point.
- ARAGON, Y. (2011). *Série temporelle avec R Méthodes et cas*.
- BAI, J. (1997). Estimation of a change point in multiple regression models.
- BAI, J. (1998). *Estimating And testing linear model with multiple structural changes*.
- BAI, J. et PERRON, P. (2003). Critical values for multiple structural change tests.
- CHRISTIANO, L. J. (1992). Searching for a break in gnp.
- DION, K., SLIVITZKY, M., SAID, M. et FAVRE, A.-C. (2015). Analyses des ruptures et des tendances dans les écoulements annuels observés et simulés par le mrcc en climat actuel (1961-1999) pour les 21 bassins versants de la péninsule québec/labrador.
- GRAUX, A. (2011). Evaluation des effets du changement climatique sur la fièvre catarrhale ovine en europe.
- HARLE, F. (2016). Détection de ruptures multiples dans des séries temporelles multivariées : application à l'inférence de réseaux de dépendance.
- HORVATH, L. (1995). Detecting changes in linear regressions.
- J.E.PATUREL, E.SERVAT, KOUAME, B., F.BOYER, J., NIEL, H. L. et MASSON, J. M. (1996). Procédures d'identification de "ruptures" dans des séries chronologiques - modification du régime pluviométrique en afrique de l'ouest non sahélienne.
- KHAIREH, B. A. et FARAH, H. (2012). Towards pre-elimination of malaria in the republic of djibouti.
- KITAGAWA, G. et GERSCH, W. (1984). smoothness priors-state space modeling of time series with trend and seasonality.
- RAYALEH, H.-O. (2005). La gestion d'une pénurie: l'eau à djibouti.
- SANITAIRE (2015). *Annuaire Statistiques de Djibouti 2007-2016*.
- XIONG, L. et GUO, S. (2004). Trend test and change-point detection for the annual discharge series of the yangtze river at the yichang hydrological station / test de tendance et détection de rupture appliquées aux séries de débit annuel du fleuve yangtze à la station hydrologique de yichang.
- ZEILEIS, A., KLEIBER, C., KRAMER, W. et HORNIK, K. (2003). Testing and dating of structural changes in practice.