

Chapitre 15 : Bilan du traitement des données pluviométriques

Rappelons que nous avons au total 65 années et 6 mois, soit 23.923 jours (hauteurs de précipitations journalières). On considère cette série comme un échantillon prélevé dans une population (hauteur de précipitations journalières), destinée à représenter l'ensemble de la population. Plus l'effectif de l'échantillon est grand, plus l'estimation des caractères descriptifs de la population sera précise.

Dans ce Chapitre, on va présenter tous les résultats des traitements relatifs aux données suivant les méthodes présentées dans la partie II- Chapitre 8 en les interprétant en même temps. Mais avant tout, définissons notre période d'étude et localisons la station.

Section 1 Aperçu général de l'étude

A la figure 39 à la page 98, on a la géolocalisation de la station d'étude. Dans cette section, un aperçu général de l'étude est présenté.

V. Résultats de la vérification initiale

Dans un premier temps, nous avons procédé à une première vérification de la série de précipitation à l'aide de la fonction « recherche » dans Excel, répertoriée dans le Tableau 15.

Tableau 15: Bilan manuelle des données pluviométriques de la station Antananarivo DMH

Source : Auteur

<i>Libellés</i>	<i>Pourcentage (%)</i>
Valeurs manquantes	4.24%
Traces	0.09%
Valeurs complètes	95.67%
Total	100%

En observant le Tableau 15, on a au total 95,67 % de la série qui est complète. Soit 1.015 jours de hauteurs manquantes qui représentent 4,24 % de toute la série. 0,09 % des hauteurs sont des traces (inférieur à 0,1 mm). Dans un premier point de vue, on peut dire qu'on satisfait largement au seuil de 80 % conseillé par l'OMM.



District Avaradrano, hachuré en rouge.
Source : BD500

Figure 39: Géolocalisation de la station Antananarivo DMH dans le district Avaradrano, région Analamanga à partir d'une image satellite haute résolution sur Google Earth pro 2017.
Source : Auteur

Sur la Figure 40, nous avons représenté graphiquement toute la série chronologique des précipitations journalières sans modification, ni traitement. On y observe généralement une importante lacune, qui va du mois d’Avril 2009 à Décembre 2011. La fréquence des valeurs nulles quotidiennes y est aussi très remarquable tout le long de la série. Un maximum journalier est observé le jour du 28 Janvier 1982.

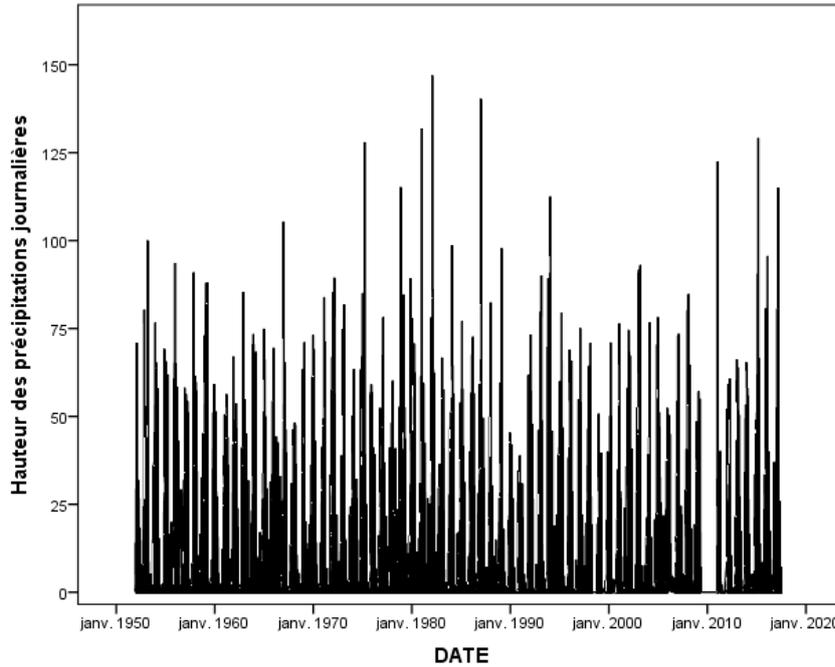


Figure 40 : Hyétogramme des précipitations journalières sur la période de 1952 à juin 2017 à partir des données initiales sans modification ni traitement.

Source : Auteur

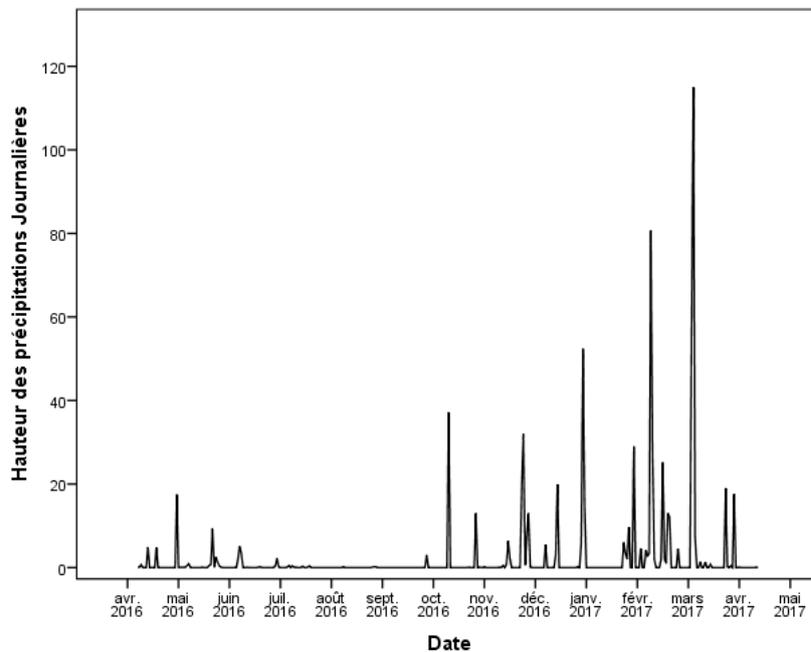


Figure 41: Hyétogramme des précipitations journalières sur la période d'étude à partir des données initiales sans modification ni traitement.

Source : Auteur

Sur la Figure 41, on a représenté la série chronologique journalière de la période d'étude. Pour une bonne nouvelle, la période d'étude ne présente aucune lacune. Le graphique présente un maximum le 08 Mars 2017. On observe une importante alternance des jours sans précipitations qui sont assez long entre les pics, et ceci même pendant la saison des pluies ; très marquée en début de saison persistant sur 4 mois, allant du mois d'Octobre 2016 à Janvier 2017.

Section 2 Rapport du contrôle de qualité des données sous RClindex

Dans ce paragraphe, on va rapporter les résultats du contrôle de qualité sous RClindex en les commentant bien évidemment.

I. Les fichiers générés

A la sortie de la simulation de RClindex sur le Fichier AntananarivoDMH.txt, 8 fichiers sont générés dans un nouveau répertoire « log » créés sous le répertoire mère : 4 fichiers CSV et 4 fichiers PDF. Nous nous intéressons qu'aux 3 fichiers concernant les précipitations, les autres se référant sur les températures maximales et minimales sont exclus de notre rapport.

1. Le fichier « Anatanarivo DMH_nastatistic.csv »

Il donne les détails du nombre de valeurs manquantes dans la série. Sur le Tableau 16, on observe 2 cases colorées en jaunes, rouges et gris contenant le nombre de valeurs manquantes par mois chaque année. Les autres cases représentées par 0 représentent les mois complets sur chaque année.

Les cases en rouges indiquent les précipitations mensuelles contenant les nombres de jours manquants inférieurs ou égaux à 10, ce sont les mois de : Mai 1992 (1jour manquant), Octobre 1994(1jour manquant), Janvier 1997(1jour manquant), Juillet 2007(1jour manquant), Octobre 2008(4Jours manquants), Juillet 2012(6jours manquants), Août 2013(1jour manquant), Octobre 2014(1jour manquant), Février 2015(1jour manquant), Avril 2015(2jours manquants), Mai 2015(2jours manquants), Décembre 2015(2jours manquants), Avril 2016(1jour manquant), Juillet 2016(1jour manquant), Août 2016(1jour manquant), Septembre 2016(2jours manquants), et Mars 2017(1jour manquant). Ces mois constituent les périodes tolérées suivant les normes, donc on n'a pas à exclure ces périodes.

Les cases en jaunes indiquent les précipitations mensuelles contenant les nombres de jours manquants supérieurs à 10. Ce sont les mois de : Août et Septembre 1991, Janvier à Août 2006, Mai à Décembre 2009, janvier à Décembre 2010, Février à Avril 2011 et Octobre à Décembre 2011. Ces mois représentent les périodes dépassant le seuil exigé par l'OMM, de ce fait, on va procéder à leurs éliminations.

Les cases en gris sont les années contenant plus de 3 mois de manque, avec plus de 10 jours de manques par mois. Les années 2006, 2009, 2010, 2011 présentent ces caractéristiques. De même que pour les cases en rouges, on va les éliminées de la série.

Tableau 16 : Répartition des données journalières manquantes par mois chaque année sur la période de 1952 à Juin 2017.

Source : Auteur, RCLimdex

Année	J	F	M	A	M	J	J	A	S	O	N	D
1952	0	0	0	0	0	0	0	0	0	0	0	0
1953	0	0	0	0	0	0	0	0	0	0	0	0
1954	0	0	0	0	0	0	0	0	0	0	0	0
1955	0	0	0	0	0	0	0	0	0	0	0	0
1956	0	0	0	0	0	0	0	0	0	0	0	0
1957	0	0	0	0	0	0	0	0	0	0	0	0
1958	0	0	0	0	0	0	0	0	0	0	0	0
1959	0	0	0	0	0	0	0	0	0	0	0	0
1960	0	0	0	0	0	0	0	0	0	0	0	0
1961	0	0	0	0	0	0	0	0	0	0	0	0
1962	0	0	0	0	0	0	0	0	0	0	0	0
1963	0	0	0	0	0	0	0	0	0	0	0	0
1964	0	0	0	0	0	0	0	0	0	0	0	0
1965	0	0	0	0	0	0	0	0	0	0	0	0
1966	0	0	0	0	0	0	0	0	0	0	0	0
1967	0	0	0	0	0	0	0	0	0	0	0	0
1968	0	0	0	0	0	0	0	0	0	0	0	0
1969	0	0	0	0	0	0	0	0	0	0	0	0
1970	0	0	0	0	0	0	0	0	0	0	0	0
1971	0	0	0	0	0	0	0	0	0	0	0	0
1972	0	0	0	0	0	0	0	0	0	0	0	0
1973	0	0	0	0	0	0	0	0	0	0	0	0

1974	0	0	0	0	0	0	0	0	0	0	0	0
1975	0	0	0	0	0	0	0	0	0	0	0	0
1976	0	0	0	0	0	0	0	0	0	0	0	0
1977	0	0	0	0	0	0	0	0	0	0	0	0
1978	0	0	0	0	0	0	0	0	0	0	0	0
1979	0	0	0	0	0	0	0	0	0	0	0	0
1980	0	0	0	0	0	0	0	0	0	0	0	0
1981	0	0	0	0	0	0	0	0	0	0	0	0
1982	0	0	0	0	0	0	0	0	0	0	0	0
1983	0	0	0	0	0	0	0	0	0	0	0	0
1984	0	0	0	0	0	0	0	0	0	0	0	0
1985	0	0	0	0	0	0	0	0	0	0	0	0
1986	0	0	0	0	0	0	0	0	0	0	0	0
1987	0	0	0	0	0	0	0	0	0	0	0	0
1988	0	0	0	0	0	0	0	0	0	0	0	0
1989	0	0	0	0	0	0	0	0	0	0	0	0
1990	0	0	0	0	0	0	0	0	0	0	0	0
1991	0	0	0	0	0	0	0	20	16	0	0	0
1992	0	0	0	0	1	0	0	0	0	0	0	0
1993	0	0	0	0	0	0	0	0	0	0	0	0
1994	0	0	0	0	0	0	0	0	0	1	0	0
1995	0	0	0	0	0	0	0	0	0	0	0	0
1996	0	0	0	0	0	0	0	0	0	0	0	0

1997	1	0	0	0	0	0	0	0	0	0	0	0
1998	0	0	0	0	0	0	0	0	0	0	0	0
1999	0	0	0	0	0	0	0	0	0	0	0	0
2000	0	0	0	0	0	0	0	0	0	0	0	0
2001	0	0	0	0	0	0	0	0	0	0	0	0
2002	0	0	0	0	0	0	0	0	0	0	0	0
2003	0	0	0	0	0	0	0	0	0	0	0	0
2004	0	0	0	0	0	0	0	0	0	0	0	0
2005	0	0	0	0	0	0	0	0	0	0	0	0
2006	18	14	21	22	26	26	24	28	0	0	0	0
2007	0	0	0	0	0	0	1	0	0	0	0	0
2008	0	0	0	0	0	0	0	0	4	0	0	0
2009	0	0	0	0	31	30	31	31	30	31	30	31
2010	31	28	31	30	31	30	31	31	30	31	30	31
2011	0	28	31	10	0	0	0	0	0	31	30	31
2012	0	0	0	0	0	0	6	0	0	0	0	0
2013	0	0	0	0	0	0	0	1	0	0	0	0
2014	0	0	0	0	0	0	0	0	0	1	0	0
2015	0	1	0	2	2	0	0	0	0	0	0	2
2016	0	0	0	1	0	0	1	1	2	0	0	0
2017	0	0	1	0	0	0						

2. Le fichier Antananarivo DMH_prcpQC.csv

La liste des valeurs négatives de précipitations est répertoriée dans ce fichier. Dans notre cas, RClindex n'a détecté aucune valeur présentant ce genre dans notre série.

3. Le fichier Antananarivo DMH_prcpPLOT.pdf

Ce fichier contient les représentations graphiques de la série en caractérisant les densités de la population et l'occurrence des valeurs manquantes sur toute la série.

Sur la Figure 42, RClindex a construit l'histogramme de la densité des hauteurs de précipitations supérieures ou égales à 1 mm. Sur l'axe des ordonnées, on distingue les valeurs de l'aire représentant la fréquence de chaque classe des hauteurs de précipitations. Sur l'axe des abscisses, les hauteurs de précipitations en mm. Les histogrammes en vert définissent les classes et la courbe en rouge délimite l'aire contenant toutes les classes. On déduit à partir de ce graphique que l'occurrence des valeurs entre 0 et 20 mm occupe une place pondérante dans toute la série.

La Figure 43 représente les répartitions temporelles des valeurs de hauteurs de précipitation quotidiennes manquantes sur la station Antananarivo DMH par série de 9 ans. Les zones marquées soient par des points (représentent entre 1 et 6 jours consécutifs de manques), soient par des traits gras continus (représentent plus de 10 jours consécutifs de manques) en rouge sont les séries manquantes.

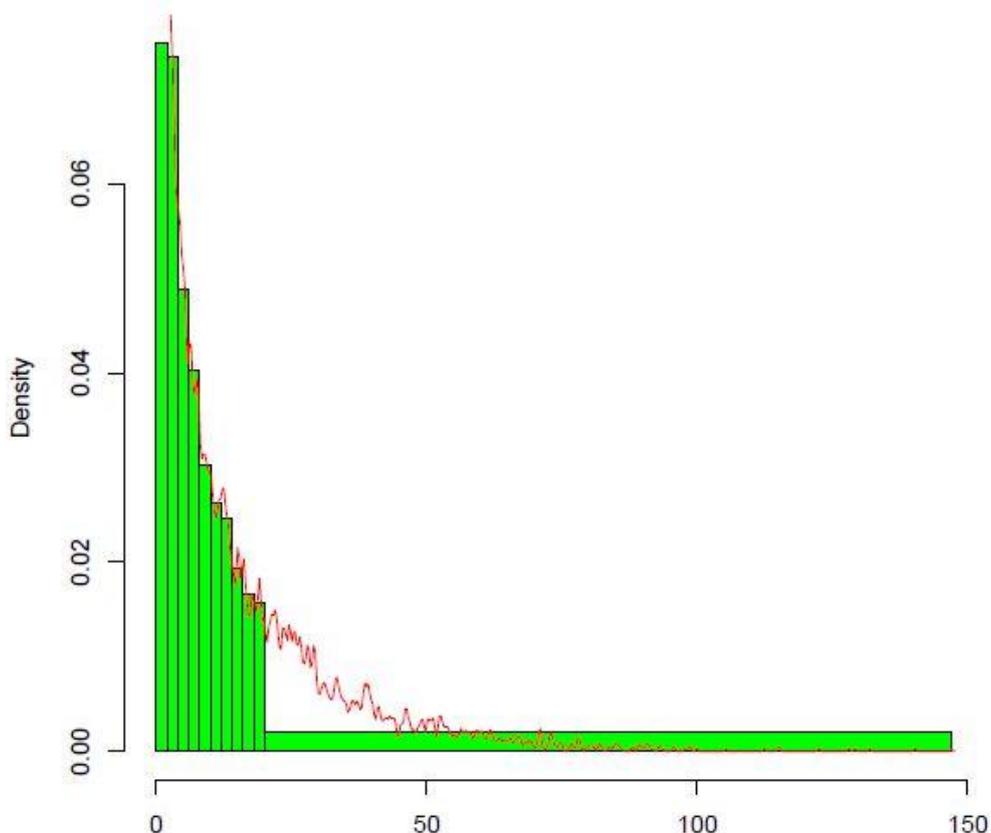


Figure 42: Histogramme de la densité de répartition des hauteurs de précipitations ≥ 1 mm

Source : Auteur, RClindex

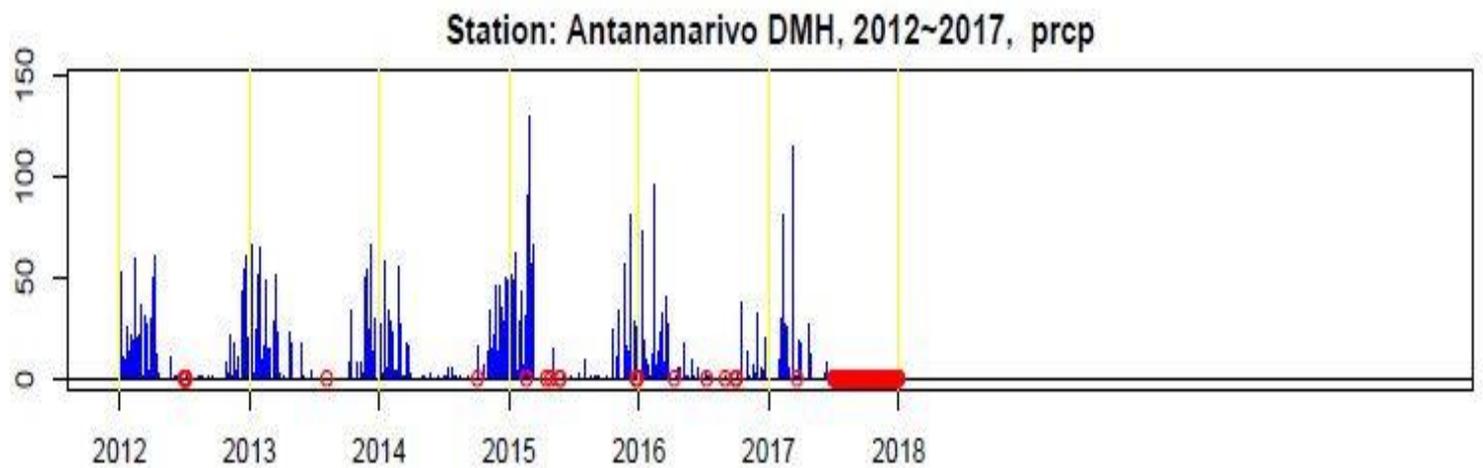


Figure 43 : Répartitions temporelles des valeurs manquantes

Source : Auteur, RCLindex

Contrairement au Tableau 16-page 101 qui montre les effectifs sur chaque mois par année, ce graphique situe les valeurs manquantes temporellement sur toute la série. Il est aussi à noter que RCLindex considère les valeurs sur une année pleine si le cas contraire se produit, le progiciel proclame les mois non-figurant sur l'année comme contenant tous des valeurs manquantes. En effet, ceci est le cas pour l'année 2017 puisque notre série se termine le 30 Juin de l'année. De ce fait, les mois de Juillet à Décembre sont déclarés manquantes.

Sur la figure 43, les cercles en rouges représentent les séries manquantes inférieures à un mois tandis que les traits épais rouges représentent celles supérieures à un mois. Ici, nous avons représenté qu'une figure présentant les répartitions temporelles des valeurs manquantes, le reste est présenté dans l'Annexe IX.

Il est nécessaire aussi de signaler que RCLindex n'a pas détecter de valeurs aberrantes. En tout, les erreurs connues sur la série se réfère seulement aux valeurs manquantes.

Section 3 Traitement des données pluviométriques manquantes

Ici, nous allons restituer les lacunes par les méthodes qu'on a exposé dans la Partie II. Apparemment, l'algorithme EM.

I. Répartition et classification des données manquantes

En premier lieu, RCLindex code les valeurs manquantes en les remplaçant par « -99,9 ». En observant le schéma des données dans le Tableau 16-page 101, on en déduit que le pattern des données manquantes est de type « Multivarié ». Quant à la classe, elle est du type « MCAR (Missing Completely At Random) », les données sont manquantes aléatoirement ; la probabilité d'avoir des valeurs de hauteurs de précipitations manquantes ne dépend pas d'eux, mais uniquement des paramètres extérieurs indépendants.

II. Elimination et restitution des données manquantes

Ici, on a procédé suivant les normes citées dans la partie II. Dans un premier temps, on élimine les mois contenant plus de 10 jours de manque en les codant par le chiffre « -9 » (les cases rouges dans le Tableau 16-page 101), après on restitue les valeurs quotidiennes avec les manques inférieures à 10 jours. Ensuite, on calcule les cumuls mensuels par années en introduisant les « -9 », puis on restitue les « -9 ». A titre de vérification et de précision, on calcule les cumuls annuels et on élimine les années contenant 3 mois au plus de manque en les codant par « -99 » (les cases grises dans le Tableau 16-page 101), puis, on les restitue. Les « -99 » restitués servent de référence par rapport au cumuls annuels obtenus à partir de « -9 ».

Les mois représentés par le code « -9 » : Août et Septembre 1991, Janvier à Août 2006, Mai à Décembre 2009, janvier à Décembre 2010, Février à Avril 2011 et Octobre à Décembre 2011. Les années représentées par le code « -99 » : 2006, 2009, 2010, 2011.

Section 4 Bilan de l'homogénéisation des données avec RHtests_dlyPrpc et RHtestsV4

Ici, on procède en 2 étapes : la validation des séries quotidiennes en introduisant les séries estimées à l'aide de l'algorithme EM. En deuxième temps, la validation des séries mensuelles avec les valeurs imputées.

I. Les séries pluviométriques journalières

Avant l'analyse, il faut définir les paramètres d'entrées dans la fenêtre GUI du package RHtests_dlyPrpc. Dans notre cas, nous n'avons pas de documentation exacte concernant la série, donc, les paramètres par défaut suffiront amplement. Après l'appel de la fonction « FindU », le progiciel produit 5 fichiers dans le répertoire de sortie. Le fichier présenté ci-après récapitule les résultats que nous cherchons.

1. Les fichiers « Antananarivo DMH_1Cs.txt et Antananarivo DMH_mCs.txt »

Ces deux fichiers indiquent le nombre des points de changements identifiés dans notre série testée et répertorient leurs dates et leurs statistiques. Nos résultats sont présentés dans le Tableau 17 ci-dessous :

Tableau 17 : Résultats des fichiers Antananarivo DMH "_1Cs.txt" et "_mCs.txt".

Source : Auteur, RCLindex

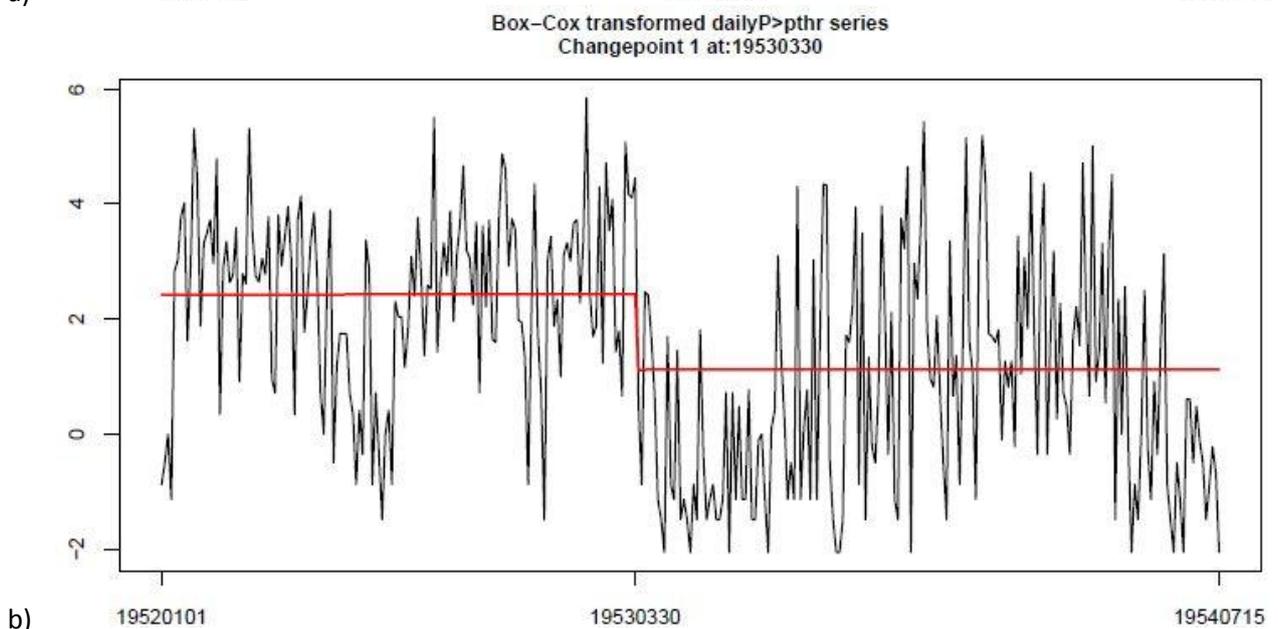
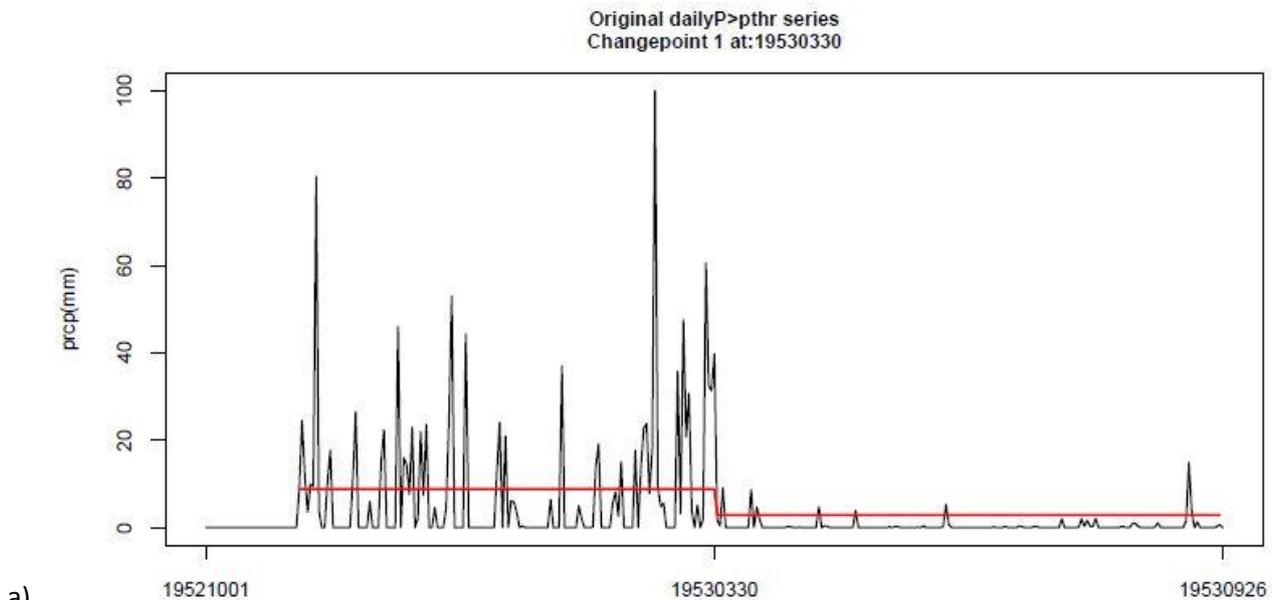
Type	Signification.	Date	P value (0.95)	T.level	PFmax	CV95
1	Yes	19530330	(1.0000 - 1.0000)	0.950	39.5194	(29.2563 - 32.8790)

Dans le Tableau 17, La première colonne (Chiffre 1) est un index indiquant que c'est un point de changement de type 1. La deuxième colonne (Yes) indique que le point de changement est statistiquement significatif pour le type. La troisième colonne énumère la date de changement, ici, le point de rupture se situe au 455ème jour, le 30 Mars 1953. Les nombres dans la quatrième colonne

(1.0000 - 1.0000) sont l'intervalle de confiance de 95% de la valeur p, ce qui est estimé en supposant que le point de changement est documenté (donc la valeur est très élevée pour un point de changement significatif de type 1). La valeur nominale p (niveau de confiance) est donnée dans la cinquième colonne. Les trois dernières colonnes sont les statistiques PFmax = 39.5194 et l'intervalle de confiance à 95% = (29.2563 - 32.8790) des percentiles PFmax qui correspondent respectivement au niveau de confiance nominal.

2. Le fichier « Antananarivo DMH_U.pdf »

Ce fichier présenté 5 graphiques se rapportant à l'ajustement de la série dont les représentations sont ci-dessous (Figure 47-a) à -e)). En se référant aux paramètres d'entrées, RHtests_dlyPrcp a nommé la série « dailyP > pthr series ». pthr est le seuil de précipitation inférieur (les précipitations quotidiennes en dessous de ce seuil seront exclus dans le test), sa valeur est de 0 mm pour le test. La ligne en rouge sur chaque figure représente la tendance et le décalage à la date 30/03/1953.



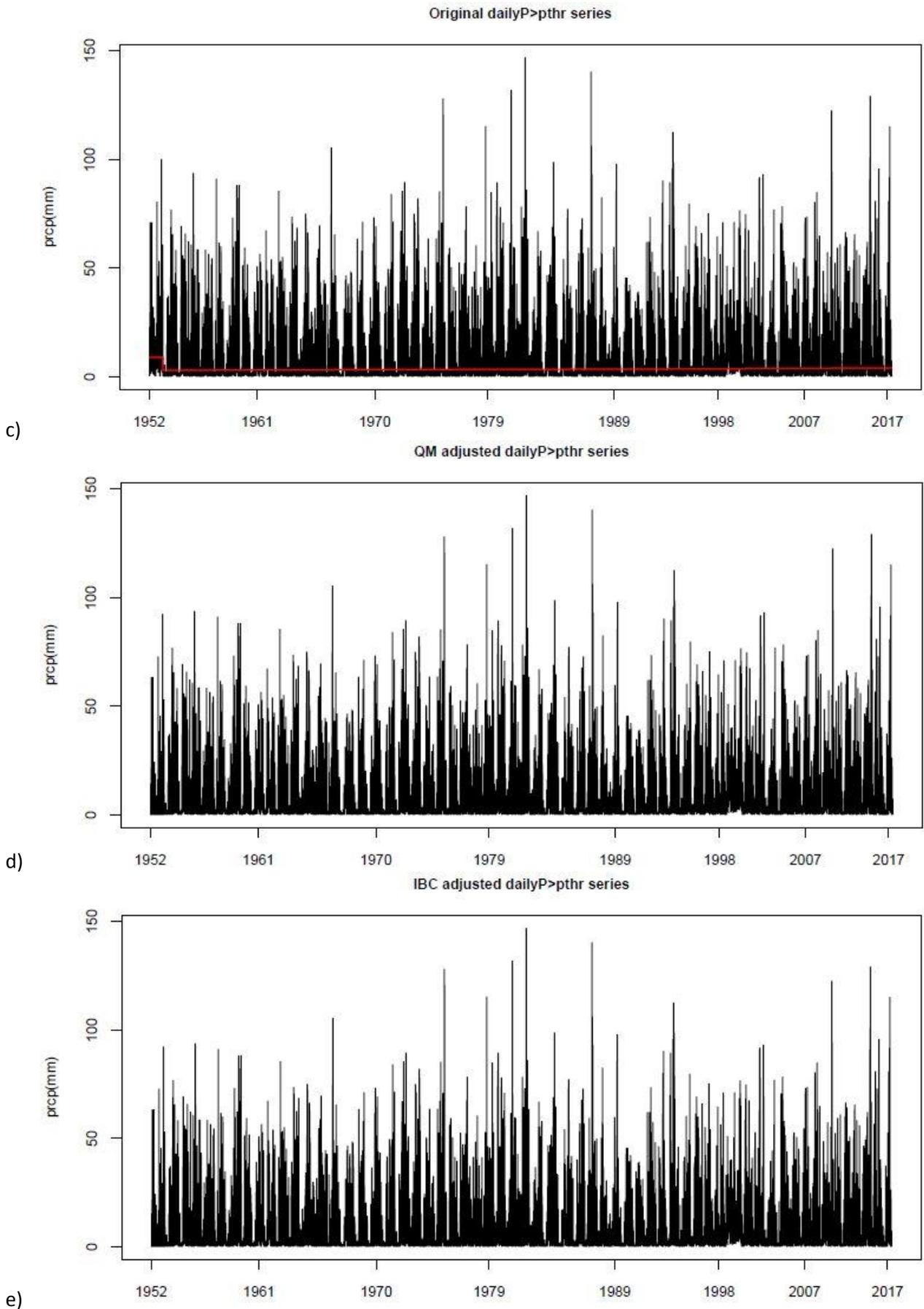


Figure 44 : Représentation de la rupture et des ajustement IBC et QM.

Source : Auteur, RHtests_dlyPrpc

La figure 44-a) présente le segment de la série original « dailyP>pthr » pour les courtes périodes entourant le point de changement.

La figure 44-b) représente les segments de la série « dailyP>pthr » de la transformée de Box-Cox, les variations moyennes estimées et la tendance linéaire pour les courtes périodes entourant le point de changement.

La figure 44-c) représente la série « dailyP> pthr » originale pour toute la période et les moyennes-décalages estimés, avec la tendance linéaire.

La figure 44-d) présente la série « dailyP> pthr » corrigée avec l'ajustement QM pour toute la période.

La figure 44-e) présente la série « dailyP> pthr » corrigée avec l'ajustement IBC pour toute la période.

Les données concernant la Figure 47 sont répertoriées dans le fichier « Antananarivo DMH_U.dat », ainsi que toutes les valeurs d'ajustement et leurs propriétés statistiques.

II. Les séries pluviométriques mensuelles

Après l'homogénéisation des séries quotidiennes, nous avons calculé les cumuls mensuels à partir des données ajustées et imputées. Ensuite, on a procédé à une deuxième homogénéisation avec RHtestsV4. De ce fait, les résultats sont très prometteurs. Après l'appel de la fonction « FindU », le progiciel produit 5 fichiers dans le répertoire de sortie. Le fichier présenté ci-après récapitule les résultats que nous cherchons.

1. Les fichiers « Antananarivo DMH_1Cs.txt et Antananarivo DMH_mCs.txt »

Comme précédemment, les résultats sont répertoriés dans un tableau. Or, celui-ci indique 0 point de changement. On peut tout de suite en conclure que notre série est homogène. Sur ce, nous devons encore vérifier les graphiques pour plus de preuves.

2. Le fichier « Antananarivo DMH_U.pdf »

De même que pour les séries quotidiennes, ce fichier présente 4 graphiques se rapportant à l'ajustement de la série dont les représentations sont ci-dessous. La ligne en rouge sur chaque figure représente la tendance et le décalage.

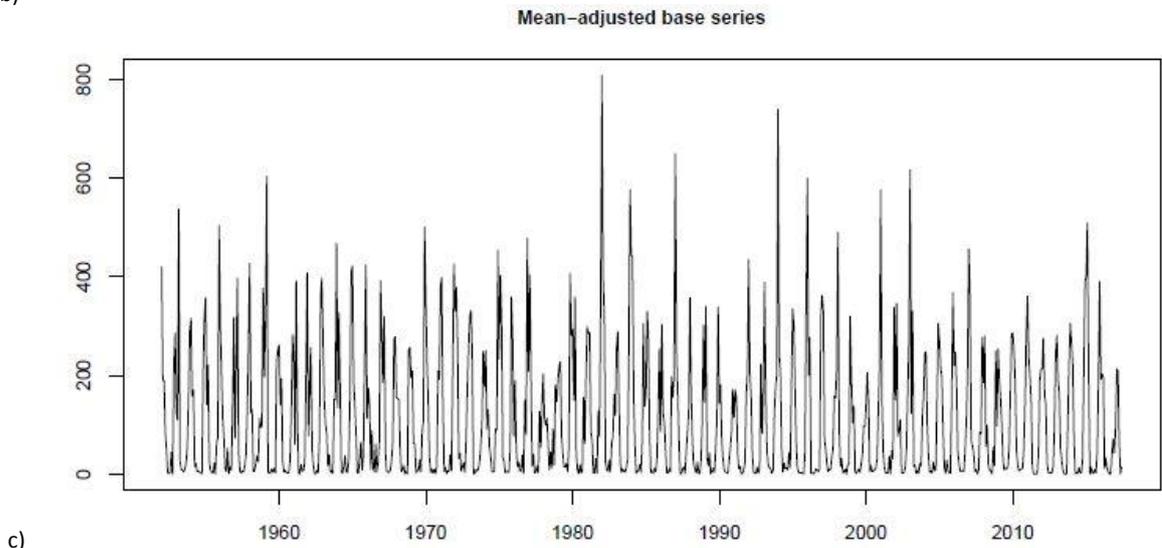
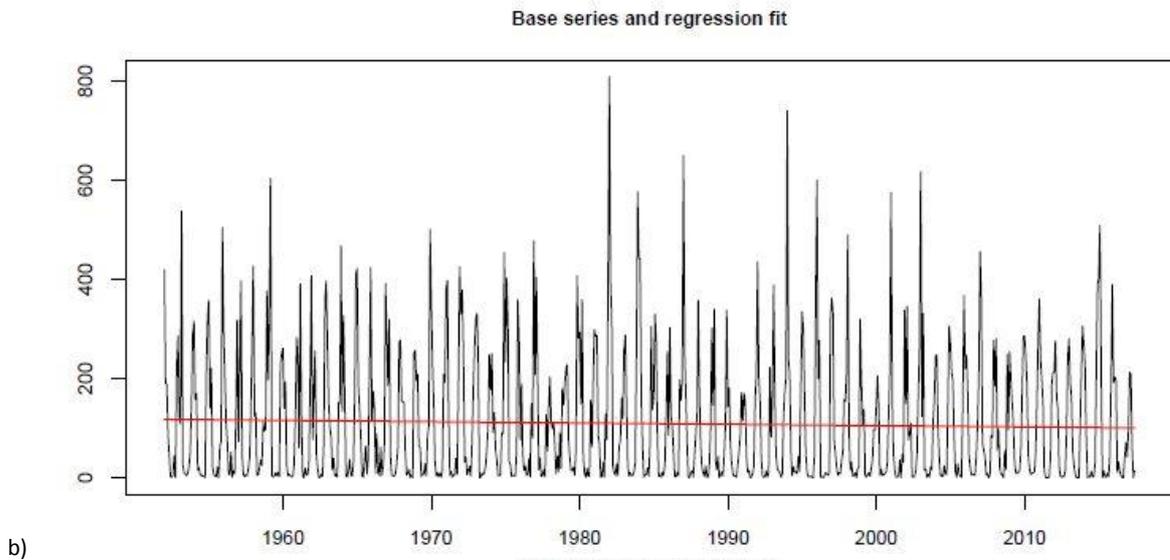
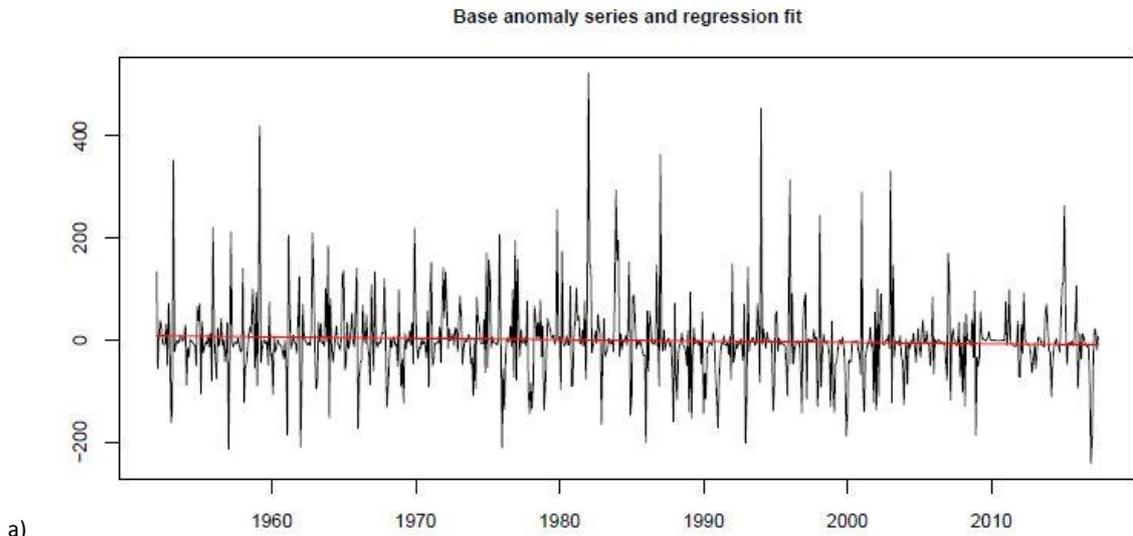
La Figure 45-a) représente la série de base d'anomalies (c.-à-d. anomalies relatives au cycle annuel moyen de la série de base) ainsi que sa correction par modèle de régression à phases multiples.

La Figure 45-b) représente la série de base ainsi que les sauts estimés de la moyenne et la tendance linéaire. Ici, on observe aucun saut.

La Figure 45-c) représente la série de base ajustée selon les moyennes.

La Figure 45-d) représente la série de base ajustée selon QM (les deux dernières ajustées selon le segment de base).

Au final, la série est complète, de bonne qualité et bien homogénéisée. Toutes ces procédures ont été nécessaires puisque nous n'avons pas de série de référence, par conséquent, on doit traiter la série de base de façon à avoir des données fiables et optimales. Donc, dans la suite, nous pouvons calculer les différentes valeurs qu'on souhaite dans les normes.



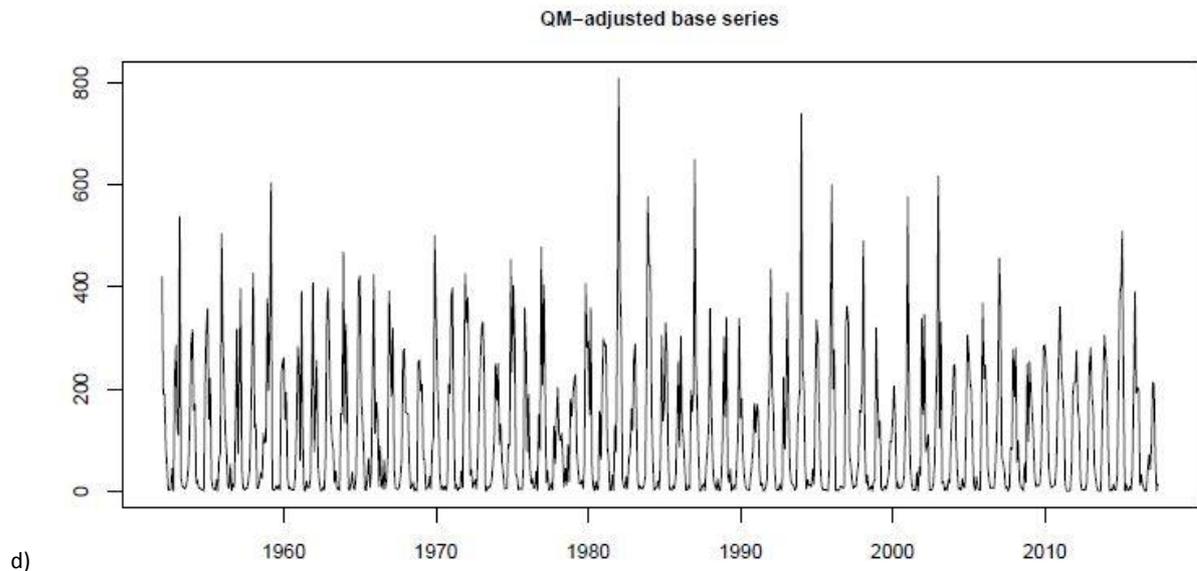


Figure 45 : Représentation des ruptures, des ajustement QM et modèles de régression.

Source : Auteur, RHtestV4

En observant les droites de tendances et de décalage, on peut aussi dire que les précipitations au sein de la station Antananarivo DMH présentent une baisse légère en avançant dans le temps. Bien évidemment, ceci est prouvé par les valeurs de la tendance linéaire, « trend0 » calculées par les deux logiciels ; RHtests_dlyPrcp affiche $\text{trend0} = -0,02554 \text{ mm/mois}$ sur les totaux mensuels originale, RHtestV4 affiche $\text{trend0} = -0,022115 \text{ mm/mois}$. Ceci est dit, les deux valeurs sont négatives et sont à peu près la même à 3 millièmes près. La cause de cette baisse légère est probablement l'augmentation de la température due au changement climatique associée à l'utilisation excessive des réserves d'eau qui assèche les sols influant instantanément sur l'évaporation, l'évapotranspiration et l'humidité de l'air ; facteurs principaux de la formation des précipitations ; conduisant peu à peu aux caractéristiques de la sécheresse.