

Chapitre I : Contrôle de la qualité des données pluviométriques

1. Critique et contrôle des données pluviométriques

Les données pluviométriques proviennent de sources différentes. En Algérie, ces données peuvent être fournies par l'Office National de la Météorologie (ONM), l'Agence Nationale des Ressources Hydrauliques (ANRH) et par l'Agence Nationale des Barrages et des Transfère (ANBT). La disparité des sources pose souvent un problème de la qualité des données.

La critique et le contrôle de la qualité des données hydrologiques sont donc des traitements qui doivent être effectués par l'hydrologue pour s'assurer de l'homogénéité et la fiabilité des séries pluviométriques à utiliser dans une étude hydrologique donnée. Ces traitements peuvent consister en un simple traitement primaire comme ils peuvent consister en un comblement de données manquantes, en une extension de séries courtes, ou en une homogénéisation des séries hétérogènes par le biais de méthodes statistiques, numériques ou graphiques.

1.1. Traitement primaire des données de pluie

Avant toute étude hydrologique ou statistique même très simple, comme le calcul d'une pluie annuelle ou une moyenne inter-annuelle, il est recommandé de faire un traitement primaire des données brutes recueillies par un observateur ou un instrument de mesure. Ce traitement consiste à rendre l'information brute lisible et exploitable comme le passage d'un enregistrement pluviographique à un hyétoگرامme par exemple. Il comprend également un contrôle primaire des données, par exemple déceler d'éventuelles erreurs de saisie, à l'exclusion de tous traitement statistique ou graphique qui consistent à vérifier si la série des pluies annuelles sur laquelle on veut travailler est homogène, c'est-à-dire si l'échantillon fait bien partie de la même population ou de deux populations distinctes, artificiellement groupées à notre insu en une série hétérogène.

Les erreurs les plus souvent rencontrées relèvent de deux catégories :

✚ Les erreurs accidentelles et aléatoires :

Sont les erreurs que peut subir une donnée de pluie relevée au niveau des opérations par lesquelles elle passe avant d'arriver aux différents services qui vont l'utiliser :

- Perte d'eau de pluie au cours de l'observation,
- Absence de l'observateur non signalée,
- Déguisement de la donnée non lue à temps ou décalage de jour,

- Oublis de virgules, mauvaises interprétations des chiffres,
- Calcul des cumuls, moyenne etc.

✚ **Les erreurs systématiques** : elles sont généralement dues à :

- Un déplacement du site d'observation
- Une modification de l'environnement immédiat du poste de mesure
 - Déboisement ou boisement
 - Urbanisation, construction d'un barrage
- Remplacement de l'observateur
- Non-conformité du matériel de mesure ou à des défauts d'appareillage non remarqués par le service gestionnaire : éprouvette ne correspondant pas au diamètre de la bague du pluviomètre, mauvais réglage des augets du pluviographe.

Un examen attentif des bordereaux et fichiers de données peut permettre de détecter des anomalies accidentelles « à l'œil ». C'est toujours nécessaire. Mais des méthodes graphiques et numériques plus élaborées et des tests statistiques seront généralement indispensables pour mettre en évidence l'existence d'erreurs systématiques et contrôler la fiabilité et l'homogénéisation des données.

1.2. Les lacunes dans les séries pluviométriques :

Il est important de signaler que les séries des données de pluie mensuelles sont souvent incomplètes. Le manque de la donnée dans une station peut être dû à une panne de l'appareil, absence de l'observateur, arrêt de la station...

Dans des cas simples, on peut procéder par le comblement d'une lacune de pluie journalière par :

- Remplacer la pluie manquante à celle observée à la station la plus proche. Il faut vérifier la position en altitude des deux stations.
- Estimer la pluie manquante par la moyenne des pluies des stations voisines. Cette méthode est fiable lorsque les précipitations ne sont pas très irrégulières d'un poste à l'autre, une différence de 10% est tolérable.
- Méthode basée sur la tendance annuelle des pluies observées à l'échelle régionale

D'autres méthodes plus élaborées dont la méthode basée sur la régression linéaire entre données de plusieurs stations régionales et ma méthode de l'IDWA (inverse Distance Weighted Averaging) sont recommandées.

Tableau 1 : Précipitations moyennes mensuelles en mm de la station de Pont Bouchet (1975-1995)

Année	Sept	Oct	Nov	Dec	Janv	Fev	Mars	Avril	Mai	Juin	Juil	Aout
1975/76	2,0	10,0	108,0	51,0	38,0	56,0	70,0	M	45,0	16,0	17,0	18,0
1976/77	M	M	151	25	M	8,7	8,4	163,8	48,7	20,2	0	18,2
1977/78	13,4	15,8	92,5	23,8	100,7	73,9	30,5	64,8	35	7,1	0	0
1978/79	6	80,7	93,3	21,1	56,8	72	88,7	132,8	10,1	21,9	0,5	5,3
1979/80	71	29,9	115	21,7	62,3	22,4	70,8	81,3	45,9	0	M	0
1980/81	36	50,3	66,5	141,2	67,1	M	49,2	40,9	8,8	0	0	8,9
1981/82	12,7	26	36,7	146	94,6	131,7	104,5	70,2	33,3	2,5	0	7,2
1982/83	8,4	160,3	150,5	99,6	17,8	4	66,4	3,6	10,7	4,9	M	0
1983/84	125,4	60,3	125,1	62,6	193,5	151,6	50,6	37,6	9,5	5,3	0,9	2,2
1984/85	31,3	234,3	45,1	232,3	79,4	34,6	155,4	43,7	44,1	0	0	M
1985/86	37	19	30,7	50,1	97,8	59,3	81,2	36,1	3,6	16,8	0,4	1,7
1986/87	13,1	66,1	99,8	148,4	97,7	117,6	86,9	37	61,2	0,5	6,6	2,5
1987/88	22,2	34,5	60,9	10,2	67,9	57,7	36,4	15,8	29,5	30,3	0,7	M
1988/89	82,5	2,5	54,8	106,4	28,4	57,1	24,3	M	25,6	10,4	1,7	16,1
1989/90	19,1	68,4	61,3	100,5	146,1	1,9	22,7	27,1	25,9	12,5	M	4
1990/91	3	28	116,3	175,4	75,5	66,8	109,6	26,5	7	40,7	0	4,8
1991/92	44,9	91,8	46,2	20	67,8	101,2	54,1	88,9	79,1	17,1	11,6	0,2
1992/93	5,6	25,6	169,5	155,5	57,3	38,6	58,2	58,2	58,2	3,9	0	1,5
1993/94	27,9	54,2	11,1	113,8	86	98,8	5,8	64,7	20,5	6	M	M
1994/95	26,8	42,8	17,3	114,8	150,2	17,7	80,5	22,2	8,2	29,4	0	21,8

1.3. Le contrôle graphique des séries pluviométriques :

La représentation graphique de la série chronologique de la pluie annuelle et le tracé de la moyenne interannuelle donnent une idée sur la tendance pluviométrique. On peut également ressortir les excédents et les déficits d'apports pluviométriques soit respectivement les années humides et les années sèches enregistrées (*fig.1*).

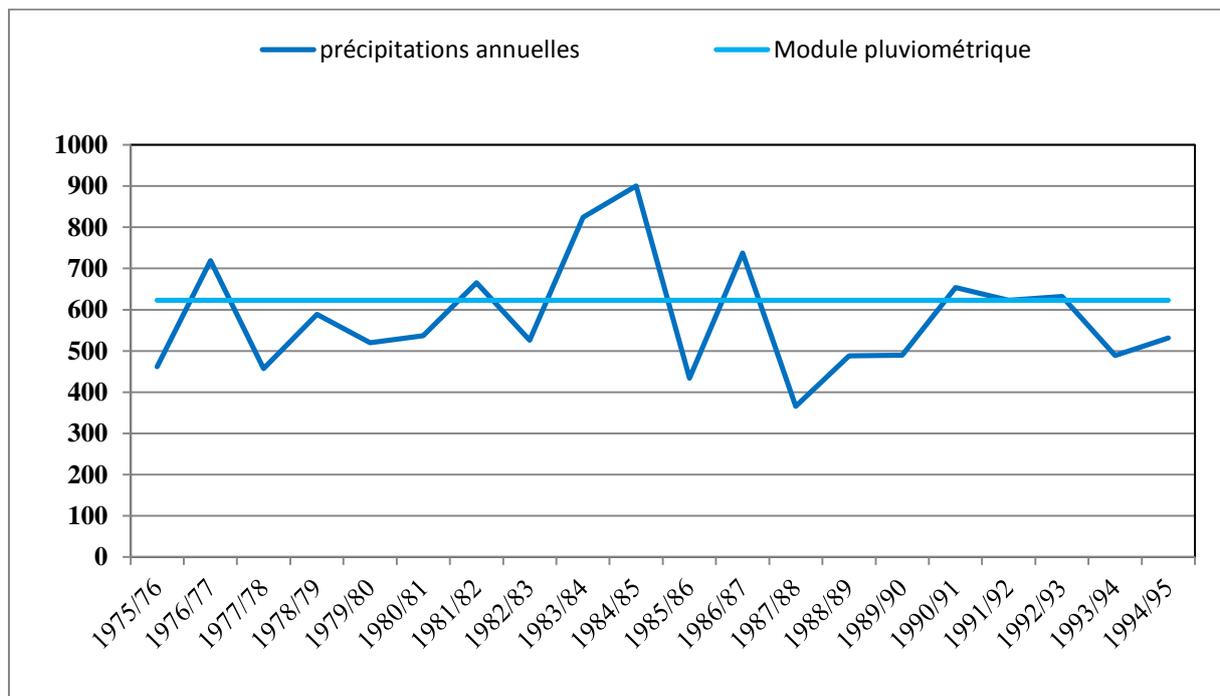


Figure 1 : Précipitations moyennes annuelles de la station de Pont Bouchet complétée.

Cependant des techniques graphiques existent et permettent à l'ingénieur de vérifier et confirmer l'existence d'une hétérogénéité dans une série pluviométrique.

1.3.1. La méthode de simple masse : elle consiste à représenter le cumul des pluies annuelles enregistrées à la station à contrôler en fonction des années. La linéarité du graphique est un indice d'homogénéité de la série se traduit par un changement de pente indiquant l'année de l'hétérogénéité.

1.3.2. La méthode du double cumul : plus efficace et plus répandue, l'homogénéisation par cette technique graphique nécessite la connaissance d'une série de données annuelles homogènes et observées dans une station de référence dite station témoin, ou station de base, voisine et régionale avec la station à corriger. La méthode du double cumul ou double masse a l'avantage de permettre de mettre en évidence la présence d'une anomalie dans la série étudiée et de la corriger.

↳ **Principe de la méthode :**

Il s'agit de comparer la tendance de la station étudiée par rapport à celle de la station témoin, en traçant le graphe des données cumulées à la station étudiée par rapport aux données cumulées de la station de référence ou station témoin.

La méthode est fondée sur le principe qu'en l'absence d'anomalie, deux stations A et B voisines et régionales mesurent chaque année une pluviométrie annuelle dans un rapport sensiblement constant d'une année à l'autre, que l'année soit sèche ou humide :

Soit donc : $PA(i)/PB(i)$ est pratiquement indépendant de l'année i .

En conséquence les points $M(i)$ de coordonnées les pluies cumulées calculées à chaque station A et B jusqu'à l'année i sont pratiquement alignés. En revanche si une erreur systématique à la station étudiée est constatée à partir d'une année i on observe un changement de pente à partir de cette année.

↳ **Interprétation graphique de la méthode**

Soit deux (2) stations X et Y ayant fourni les précipitations annuelles X (x_1, \dots, x_n) et Y (y_1, \dots, y_n).

n : est le nombre d'observations annuelles communes à X et Y, et soit X la station de référence.

On veut donc homogénéiser la station Y à partir des données de la station X.

La corrélation graphique obtenue en représentant le cumul des y_i en fonction du cumul des x_i devrait être linéaire si les deux stations sont situées dans la même région climatique et à une distance relativement faible.

S'il n'existe pas d'anomalie dans la série Y alors la pente de la droite sera constante, c'est-à-dire la série Y est homogène par rapport à la série de référence X (fig.2).

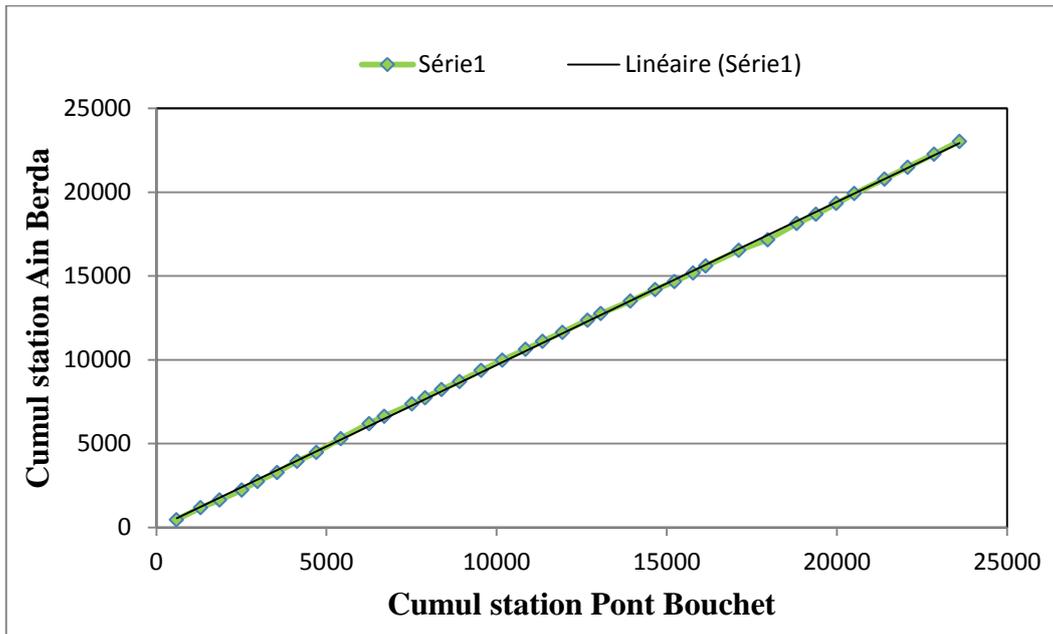


Figure 2 : Graphe des doubles cumuls (station homogène)

Dans le cas où la série étudiée a été perturbée par une modification des conditions de mesures, la droite de double cumul présentera une ou plusieurs points de cassure qui vont être mis en relief par un changement de la pente de la droite (fig.3). donc il faut procéder à l'homogénéisation des données.

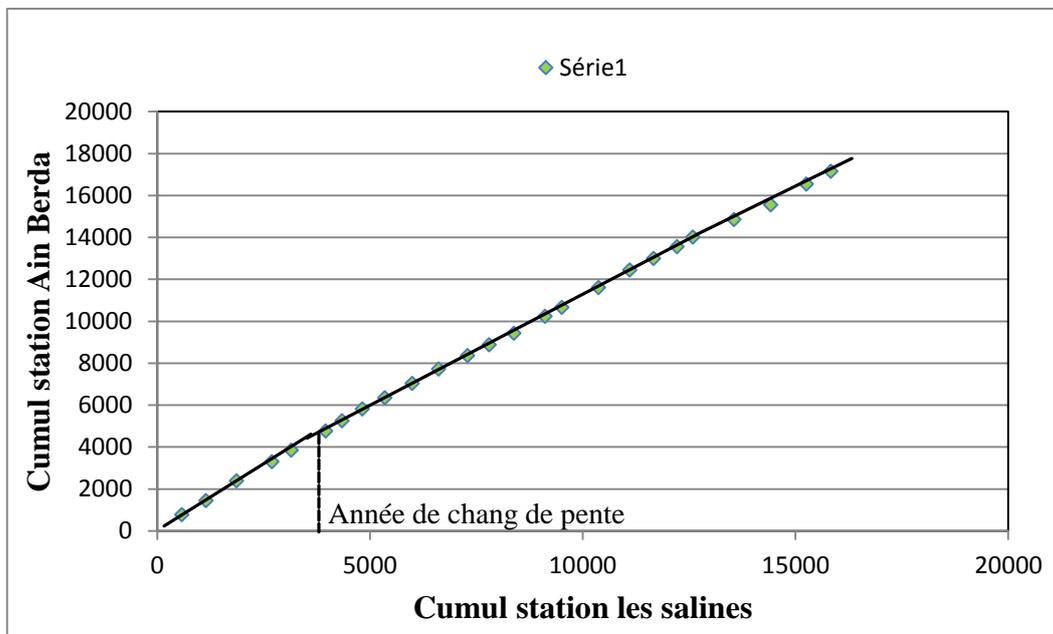


Figure 2 : Graphe des doubles cumuls (station hétérogène)

↳ Procédure de l'homogénéisation

Dans le cas de présence de points de cassure, deux questions se posent :

Quelle portion du graphe faudrait-il corriger, soit donc sur quelle période d'observation il faut corriger les observations de la série Y ?

Et comment procéder à la correction ?

En générale on regarde l'historique de la station : récente, contexte modifié ou non, arrêt du fonctionnement de la station pour une certaine période ou non...etc. La procédure consiste à chercher des éléments indicatifs pouvant faciliter la prise de décision sur la période à partir de laquelle on soupçonne l'introduction d'erreurs systématiques dans les mesures.

Dans ce cas où aucun élément indicatif n'est disponible, on considéra que les données les plus récentes sont les plus fiables.

Afin de minimiser l'influence des erreurs systématiques qui existaient dans l'une ou l'autre des stations jugées de référence (si plusieurs existent), il est préférable d'élaborer une station témoin par la moyenne de ces stations.

La procédure générale d'homogénéisation des données par la courbe des doubles masses consiste à comparer chaque station à la moyenne des autres stations, la corriger et ensuite corriger chaque station successivement. Le processus est répété jusqu'à homogénéisation de toutes les stations.

Ainsi, on peut conclure que le choix d'un groupe de base exige au préalable la comparaison deux à deux de tous les postes susceptibles d'être intégrés dans ce groupe. A défaut, on peut se contenter d'un poste de base par région.

Comment procéder à l'homogénéisation

La procédure de correction des données de la portion du graphe non fiable se fait en prolongeant la pente la plus fiable selon la formule :

$$P_{\text{corrigé}} = \frac{m_{\text{ajusté}}}{m_{\text{observé}}} \times P_{\text{observé}}$$

$m_{\text{ajusté}}$ est la pente de la portion du graphe fiable

$m_{\text{observé}}$ est la pente de la portion du graphe à corriger

Le graphe de la figure 4 indique un changement de pente (point de cassure) entre deux tronçons linéaires. On calcul tout d'abord les pentes graphiquement, on trouve :

Tronçon des données anciennes s'étale de 1950 à 1968 : $m_1=1,25$;

Tronçon des données anciennes s'étale de 1968 à 1979 : $m_2=0,95$;

Ceci justifie que la station a vécu des perturbations des conditions de mesure au cours de période 1968 à 1979.

On doit donc corriger les données observées sur cette période seulement par :

$$P_{\text{corr}} = \frac{m_1}{m_2} \times P_{\text{obs}}$$

Avant ajustement	Après ajustement
29	38,16
32	42,11
39	51,32
25	32,89
30	39,47
23	30,26
37	48,68
34	44,74
30	39,47
28	36,84
27	35,53
34	44,74

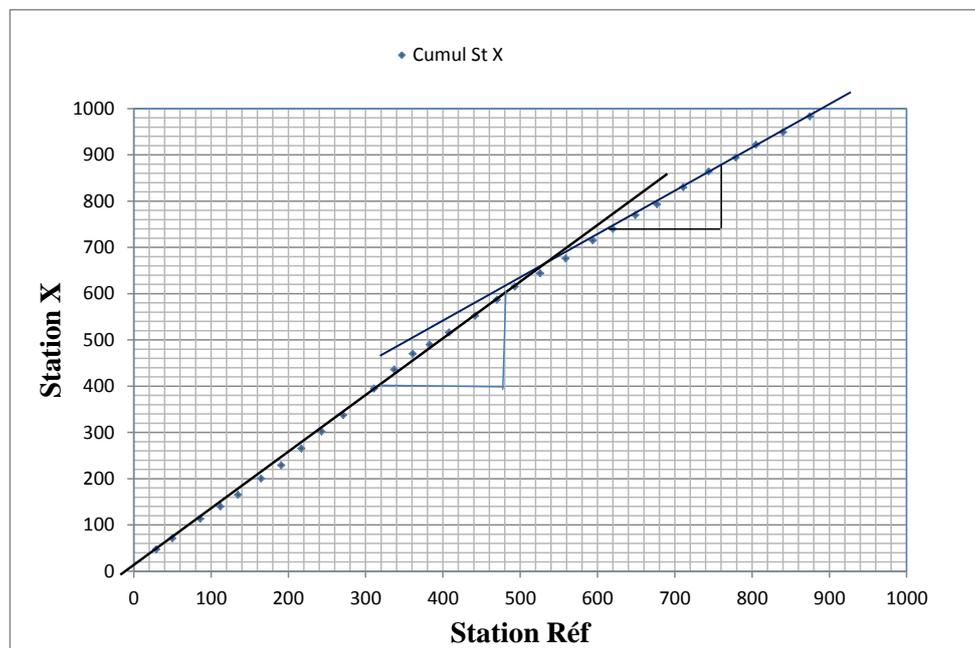


Figure 4 : Homogénéisation des données par la méthode des doubles cumulés

↳ La méthode du vecteur régional

Il est souvent conseillé d'utiliser l'approche du vecteur régional comme méthode de contrôle des erreurs systématiques pouvant exister dans les stations d'une même région. En fait, il s'agit de constituer un vecteur formé par des indices calculés de pluies annuelles des stations d'une zone homogène régionale, de point de vue pluviométrie et climatologie.

L'indice régional à l'année j se calcule, sur la période commune d'observation disponible

entre l'ensemble des stations de la zone par : $V_i = \frac{1}{n} \sum_{i=1}^n \frac{P_{i,j}}{\bar{P}_i}$

$P_{i,j}$: pluie annuelle de l'année j à la station i

\bar{P}_i : module interannuel de la station i

L'approche est basée sur la technique du double cumul et consiste à comparer le cumul de la station à contrôler par rapport au cumul des indices du vecteur régional ainsi calculé

