Homogeneisation des données en hydrologie.

L'homogénéisation des données et une analyse statistique de l'information aidant à une prise de décision conséquente.

Elle consiste en :

- la détection des anomalies dans les séries hydrologiques et d'en chercher la cause ;
- la correction des anomalies par des méthodes appropriés ;
- L'extension des séries hydrologique courtes à partir de séries de base homogènes.

- la correction des erreurs

a/ - la méthode des doubles cumules.

Elle permet de détecter la non-homogénéité d'une série de mesure et de la corriger.

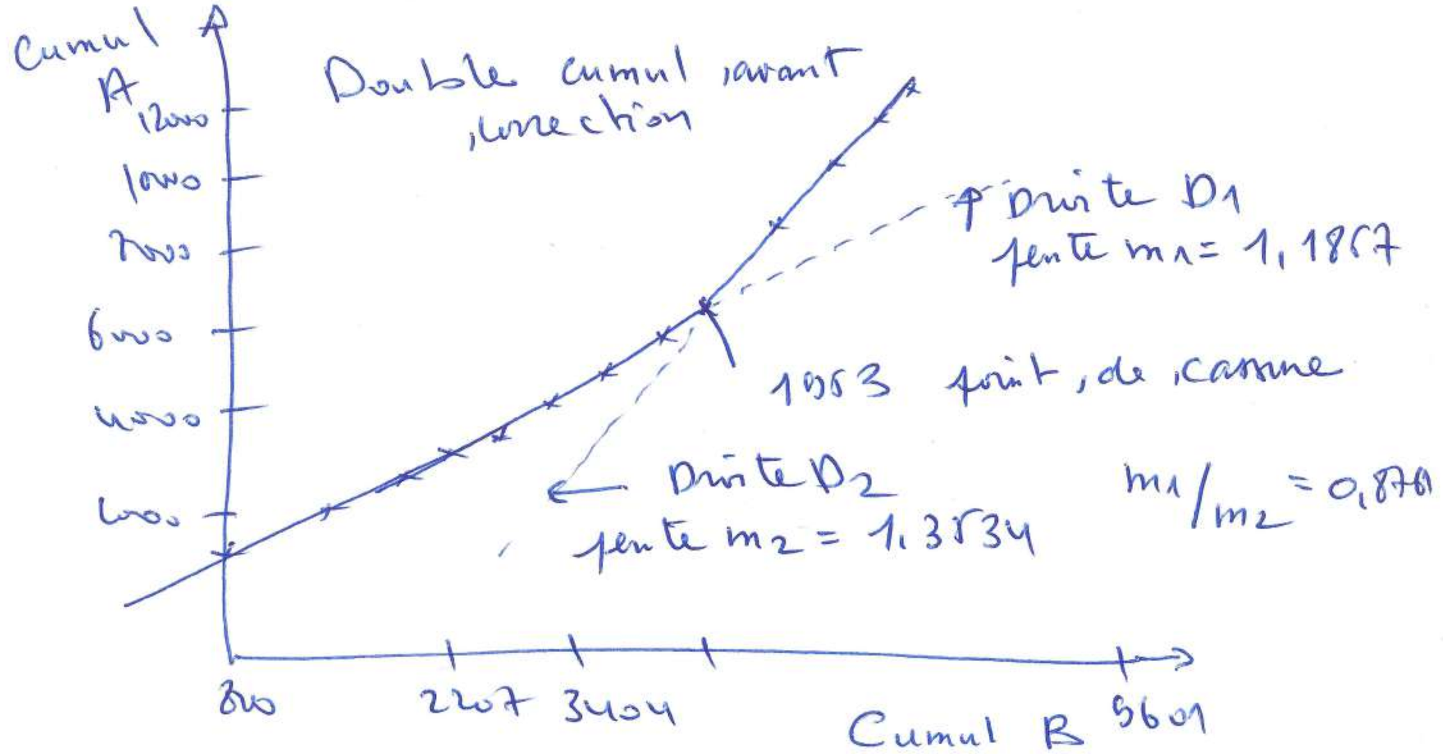
On commence donc par établir le tableau suivant :

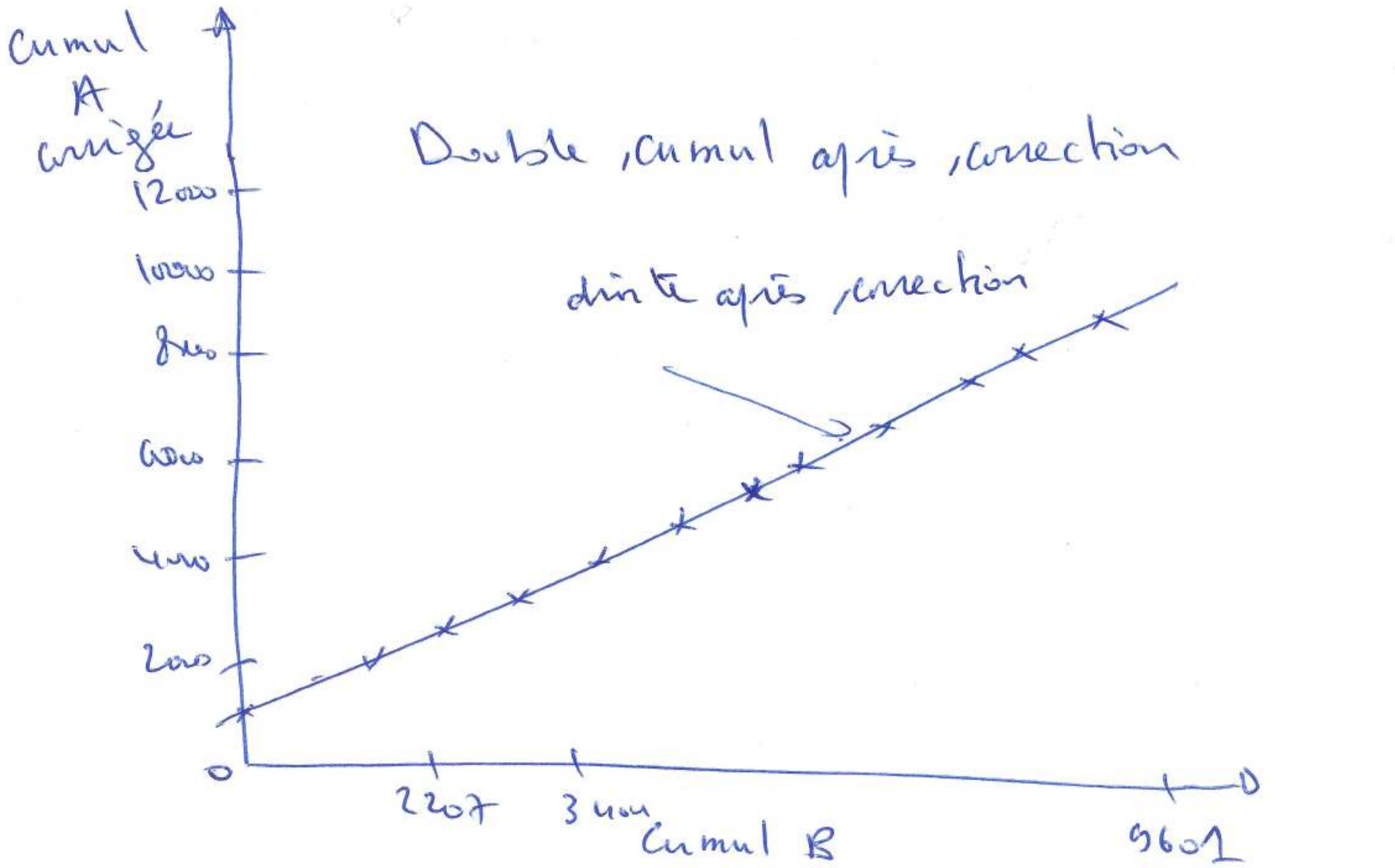
Année	Station A	Station B	Cumul B	Cumul A	A corrigée	A cor. cum
1960	869	800	800	869	869	869
1959	596	549	1349	1465	596	
1958	994	858	2207	2459	994	
1957	643	540	2747	3102	643	
1956	736	657	3404	3838	736	
1955	734	677	4081	4572	734	
1954	699	702	4783	5271	699	
1953	546	393	5176	5817	478,4	
1952	953	820	5996	6770	834,9	
1951	882	841	6837	7652	772,7	
1950	945	732	7569	8597	827,9	
1949	694	459	8028	9291	608	
1948	875	522	8550	10166	766,6	
1947	849	540	9090	11015	743,8	
1946	791	511	9601	11806	693	10996,3

- En suite on porte les valeurs sur du papier millimétré, avec les valeurs de B en abscisses et les valeurs de A en ordonnées.

- On voit sur le graphique que les points s'alignent sur deux segments de droite differents, c-à-d qu'il ya une cassure sur la droite au cours de l'année 1953. On suppose que le deplacement (ou autre cause d'erreur) s'est produit en 1953.

les données mesurées après 1953 sont jugées bonnes, et on ne doit corriger que les données précédentes (1946, 47, 48 49, 50, 51, 52 et 53).





Méthode du double cumul ou double masse.

- la décision de corriger ou non les données de l'année 1953 est prise après une ~~con~~ connaissance détaillée des circonstances de ce l'accident" au cours de cette année.

- On calcule les pentes m_1 du segment de droite qui contient les données de 1960 à 1953, et m_2 de 1952 à 1946.

On calcule le rapport des pentes m_1/m_2 avec lequel on va multiplier les données des années 1953 à 1946 pour les corriger. On porte ces valeurs sur la

- le Test d'homogénéisation de Wilcoxon:

Le test de Wilcoxon se base sur le principe suivant: Si l'échantillon X est issu d'une même population Y, l'échantillon $X \cup Y$ (union de X et de Y) en est également issu.

On procède ainsi:

Soit une série d'observations de longueur N à partir de laquelle on tire deux échantillons X et Y : N_1 et N_2 sont respectivement les tailles de ces échantillons, avec $N = N_1 + N_2$ et $N_1 \leq N_2$.

On classe ensuite les valeurs de notre série par ordre croissant. Par la suite, nous ne nous intéresserons qu'au rang de chacun des éléments des deux échantillons dans cette série.

Si une valeur se répète plusieurs fois, on lui associe le rang moyen correspondant.

On calcule ensuite la somme W_x des rangs des éléments du premier échantillon dans la série commune : $W_x = \sum \text{Rang } x$.

Wilcoxon a montré que, dans le cas où les deux échantillons X et Y constituent une série homogène, la quantité W_x est comprise entre deux bornes W_{max} et W_{min} données par les formules suivante :

$$W_{min} = \frac{(n_1 + n_2 + 1)\, n_1 - 1}{2} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

$$W_{max} = (n_1 + n_2 + 1)\, n_1 - W_{min}$$

$z_{1-\frac{\alpha}{2}}$ représente la valeur de la variable centrée réduite de la loi normale correspondant à $1 - \frac{\alpha}{2}$ (au seuil de confiance de 95%, nous avons $z_{1-\frac{\alpha}{2}} = 1.96$