

UN EXEMPLE DE CONSTITUTION
DE BANQUE DE DONNEES HYDROLOGIQUES

par J. SIRCOULON*

DOCUMENTATION
ANALYSE 2021

960245

ANALYSE

1 - Avant-propos -

Le Service Hydrologique de l'ORSTOM (Office de la Recherche Scientifique et Technique Outre-Mer) effectue des observations systématiques depuis vingt cinq ans sur les cours d'eau de la zone inter-tropicale. Depuis 1947, un millier de stations hydrométriques ont été installées et suivies, particulièrement en Afrique, par ce Service qui contrôle, soit directement, soit en vertu d'accords contractuels, une vingtaine de réseaux hydrométriques nationaux. Parallèlement à cette action, environ cent vingt bassins expérimentaux, constituant un échantillon des différents types de ruissellement de cette région du globe, ont fait l'objet d'études approfondies pendant au moins trois années consécutives.

L'ensemble de ces observations représente une masse considérable d'information qui ne cesse de s'accroître d'année en année. Dans ces conditions, la gestion rationnelle d'une telle accumulation de documents et l'exécution satisfaisante de toutes les études faites soit à la demande d'organismes extérieurs, soit pour les besoins propres de l'ORSTOM, ont fini par rendre peu efficaces les méthodes traditionnelles de calcul.

C'est ainsi que le Service Hydrologique a fait appel en 1967 au calcul automatique pour se constituer une banque de données adaptée aux besoins actuels, effectuer le traitement des données hydrologiques et procéder à des travaux de recherche pure ou appliquée à l'aide d'une vaste programmathèque personnelle.

L'expérience acquise par l'ORSTOM dans le traitement automatique des données sur ordinateur lui a permis de fournir

Fonds Documentaire ORSTOM

Cote: B* 16729

Ex:

* Bureau Central Hydrologique de l'ORSTOM à PARIS.



71429

à bon nombre de pays en voie de développement des schémas opérationnels pour une exploitation rationnelle des données des réseaux hydrométriques, et de gérer à PARIS une vingtaine de fichiers hydrologiques nationaux.

2 - Généralités sur la constitution d'un fichier -

A l'heure actuelle, le stockage informatique des données hydrologiques est en passe de devenir un mode presque classique d'utilisation et son emploi devient de plus en plus courant. Il semble utile toutefois de rappeler certaines règles malgré leur caractère apparent d'évidence.

En effet, la gestion d'une banque de données par l'informatique est très séduisante et ses avantages ne sont plus à démontrer, pourtant pour qu'elle soit vraiment efficace, il ne faut pas perdre de vue que l'exploitation optimale des fichiers dépend non seulement de la puissance de l'équipement du Centre de Calcul, mais également de la façon dont le stockage a été conçu.

Les qualités demandées à un fichier, sont, outre la conformité aux données originales, la simplicité, la souplesse et la fiabilité. Elles doivent se traduire par une grande facilité de manipulation des données :

- extraction ou recherche rapide de la valeur désirée,
- insertion aisée de toute nouvelle information,
- correction de données anciennes déjà stockées,

pour un volume de stockage minimal.

La mise au point d'un classement rationnel et opérationnel nécessite donc une préparation minutieuse des objectifs à atteindre, toute modification non prévue dans le système pouvant avoir des conséquences fâcheuses par les perturbations qu'elle entraîne, aussi bien au niveau de la recherche des données qu'au niveau de leur traitement.

En ce qui concerne la nature matérielle du support constituant la banque de données, les partisans de la carte perforée

s'opposent à ceux de la bande magnétique ou du disque et chaque école loue les avantages indéniables de son support : facilité de manipulation et de vérification de l'information d'un côté, facilités de stockage et rapidité de lecture de l'autre. Il s'agit, en fait, d'un faux problème dans la mesure où, dans un premier temps, ce ne sont pas les moyens informatiques disponibles qui doivent guider le choix du support, mais la qualité même de la donnée observée. Il est plus judicieux ainsi de commencer la constitution d'un fichier par une mise sur carte perforée qui assure le maximum de garantie de fidélité de transcription de la donnée brute pour passer dans un second temps à la transposition du fichier sur bande magnétique ou disque, qui caractérisent un stade d'évolution et donc de maturité du système.

3 - Etablissement de la Banque de données -

Nous prendrons comme exemple de données pour illustrer les procédés utilisés à l'ORSTOM, et dans un souci de simplification les relevés de hauteurs d'eau pour le fichier de données brutes et les débits pour le fichier de données traitées.

3.1 - le fichier de données brutes -

Les données originales qui vont constituer le fichier de données brutes suivent, dans chaque pays dont le réseau hydrométrique est pris en charge par l'ORSTOM, la filière suivante : collecte aux stations du réseau, centralisation de l'information au Service Hydrologique National, transmission au Bureau Central Hydrologique à PARIS dont l'atelier mécanographique assure la mise sur support. Il y a donc deux niveaux très distincts à considérer.

3.1.1 - Acquisition et transmission de l'information -

Entre l'observation assurée par le lecteur d'échelle et la même observation arrivant à l'atelier mécanographique, il existe de nombreuses possibilités de modification de l'information (erreurs de recopies, sélection inadéquate des relevés à garder, etc...), or, il est impératif d'avoir un fichier qui soit l'image

fidèle de l'observation brute (aux erreurs grossières de lecture près), toutes corrections ou changements (par exemple, variation du niveau du zéro de l'échelle) pouvant être faites ultérieurement à l'aide de programme de traitements appropriés. Il est donc nécessaire d'établir un protocole d'acquisition et de transmissions de la donnée afin d'éviter toute altération de celle-ci et de faciliter la transcription sur le support.

Ainsi des consignes très strictes sont fournies aux hydrologues des bassins représentatifs assurant eux-mêmes le dépouillement des appareils enregistreurs et il existe des fiches spéciales pour chaque type d'observation permettant une saisie rapide et efficace de l'information.

De telles consignes ne sont évidemment pas applicables aux données anciennes, qui représentent toujours une masse considérable. Des photocopies ou microfilms des documents originaux sont alors indispensables et préférables aux recopies sur fiches.

3.1.2 - établissement du fichier -

L'atelier mécanographique a finalement à sa disposition trois sortes de documents :

- des données anciennes sous forme de photocopies
- des données récentes sous forme de bordereaux appropriés
- des enregistrements non dépouillés à traiter au lecteur de points.

La mise sur support informatique de l'information doit répondre à des normes rigoureuses et soigneusement établies que nous allons aborder maintenant.

Quelle que soit la nature matérielle du support qui sera utilisé : cartes perforées, bandes magnétiques ou disques, il est nécessaire d'associer sans ambiguïté à la valeur de la donnée à stocker :

- sa nature (type) ; définition du phénomène et de la variable
- son lieu de provenance
- sa position dans le temps
- sa qualité éventuellement.

Le type de donnée n'est pas toujours représenté explicitement par un chiffre ; ainsi à l'ORSTOM où l'on utilise comme

support primaire la carte perforée pour ses facilités de manipulation et de vérification, cette représentation est fournie implicitement par un modèle de cartes approprié à l'information : hauteurs d'eau brutes, jaugeages, renseignements sur l'échelle, etc ... Des contrôles de lecture dans les programmes de traitement rendent inutiles une numérotation préalable de chaque type de donnée.

- L'opération d'identification du lieu de provenance, soit ici la station hydrologique, est beaucoup plus délicate et fait intervenir ce qu'on appelle improprement le codage.

3.1.2.1 - Le codage des stations hydrologiques -

En se plaçant logiquement à l'échelle du continent, l'identification d'une station hydrologique exige l'adoption d'un système tenant compte à la fois de la division politique de ce continent et de sa structure géographique.

Le codage retenu par l'ORSTOM permet de définir une station par une variable de 8 chiffres, pouvant se décomposer en 4 variables de 2 chiffres qui sont : ETAT, BASSIN, RIVIERE et STATION.

Ainsi, en prenant l'exemple du continent africain :

- le code ETAT attribue un numéro de 01 à 49 aux 49 états existants, classés par ordre alphabétique (ALGERIE à ZAMBIE),
- le code BASSIN donne un numéro aux bassins hydrographiques de chaque état, ces bassins étant également classés par ordre alphabétique en laissant des créneaux suffisants pour le traitement futur des bassins non encore couverts par le réseau. On réserve les 29 premiers numéros aux bassins supérieurs à 70 000 km² susceptibles de s'étendre sur plusieurs pays, afin de ne pas leur affecter un numéro différent suivant l'état où ils se trouvent, soit par exemple : BASSIN : 18 pour le Nil et BASSIN : 29 pour le Zambèze.

Dans le cas des îles occupées par un seul état, la précaution des numéros réservés devient inutile, mais il est bon de distinguer entre les divers bassins côtiers suivant leur importance.

. le code RIVIERE est un reflet de la structure du réseau hydrographique. On réserve les numéros 01 à 09 au fleuve principal, celui-ci pouvant changer de nom le long de son cours. Les affluents sont classés par ordre alphabétique et une ventilation est introduite suivant l'ordre de chaque cours d'eau, soit en définitive :

- n° 10 à 29 Affluent de 1er ordre ou petit fleuve côtier,
- n° 30 à 59 Affluent de 2ème ordre,
- n° 60 à 89 Affluent de 3ème ordre ou plus,
- n° 90 à 99 Lacs, affluents, lits majeurs.

. le code STATION affecte un numéro aux stations situées sur un même cours d'eau après classement par ordre alphabétique et aménagement de créneaux pour les stations futures possibles, mais cette disposition n'est pas impérative, un classement par ordre chronologique se basant donc sur la date de création de la station étant tout aussi valable.

Nous donnons (figure 1) un exemple de carte d'identification

3.1.2.2. - la représentation de l'hydrogramme -

Les chroniques d'observation de hauteurs d'eau qui constitue l'information brute sont fournies soit de façon discontinue par les relevés d'échelles limnimétriques, soit de façon continue par les enregistrements des limnigraphes.

La représentation chiffrée du chronogramme est pratiquée à l'ORSTOM suivant deux systèmes de pas de temps qui doivent permettre une reconstitution aussi fidèle que possible de l'hydrogramme, compte tenu de l'information disponible.

3.1.2.2.1 - Système à pas de temps constant -

Dans ce système, le chronogramme est découpé suivant un nombre donné d'intervalles de temps égaux. Il s'applique principalement aux observations manuelles des stations de réseaux. Avec ce procédé, l'adresse dans le temps de chaque hauteur est connue implicitement et conventionnellement par la position de la valeur dans la carte.

.

Le pas de temps constant semble assez rigide, toutefois on peut l'adapter à la vitesse de variation du phénomène mais il doit rester constant au moins pendant une journée.

La hauteur à l'échelle étant généralement donnée au centimètre près, il suffit de quatre chiffres pour exprimer toutes ses valeurs possibles et par convention, les données manquantes sont représentées par 9999 (soit 99,99 m). Dans le cas d'études particulières exigeant la hauteur en millimètre, ce nombre de chiffres reste encore suffisant.

Le repérage dans le temps est fourni sur chaque carte par l'indication de la variable ANNEE et de la variable MOIS et la combinaison de deux autres variables. La variable CAS donne le nombre de lectures journalières et la variable GROUP la position des données de la carte dans le mois.

L'adressage de chaque donnée est ainsi déterminé sans qu'il soit nécessaire de numéroter les cartes.

Pour donner un exemple, la mise sur support des relevés d'une station ayant 2 lectures par jour aboutit à la perforation de 4 cartes par mois.

La variable CAS est égale à 2 et la variable GROUP prend les valeurs :

- GROUP : 1 = 1er au 8ème jour
- 2 = 9ème au 16ème jour
- 3 = 17ème au 24ème jour
- 4 = 25ème jour à la fin du mois.

chaque carte pouvant contenir jusqu'à seize hauteurs d'eau, on pourrait avec ce procédé perforer jusqu'à seize observations par jour. En fait, chaque fois que le nombre d'observations journalières est élevé, il est plus commode, pour avoir une meilleure connaissance des débits, de prendre le système à pas de temps variable. Une combinaison des 2 systèmes permettant de s'adapter à la forme de l'hydrogramme est aussi prévue (il s'agit alors d'un pas de temps "mixte").

3.1.2.2:2 - Système à pas de temps variable -

Ce système est surtout utilisé pour les petits bassins et notamment les bassins représentatifs ou expérimentaux. L'adresse dans le temps de chaque valeur est fournie ici par les variables JOUR, HEU (heures), MIN (minutes) associées à la hauteur HAUT, ce qui porte à 10 le nombre de chiffres nécessaires pour chaque donnée.

En général, les variables sont obtenues par traitement des diagrammes par un lecteur de points connecté à une perforatrice par l'intermédiaire d'un ordinateur de coordonnées, les dérivés de temps ou de hauteur, ainsi que toutes les anomalies possibles d'enregistrement, sont traitées à l'aide de programmes spéciaux.

Les normes employées ne permettent pas de perforer plus de six hauteurs par cartes, mais le pas de temps variable rend possible un choix beaucoup plus judicieux des données significatives en permettant avec un minimum de points, de suivre parfaitement l'évolution de la hauteur d'eau. Signalons également que sur chaque carte figure une variable JUNITE indiquant si les hauteurs sont en cm ou en mm et que les quatre dernières colonnes sont réservées à la numérotation en séquence (NOCART) afin de permettre un contrôle rigoureux du chronogramme. On trouvera sur la figure 2 un exemple de cette représentation.

3.2 - Le fichier des données traitées -

Il s'agit, en se limitant à l'exemple que nous avons retenu des débits obtenus après traduction des hauteurs d'eau du fichier primaire. Suivant les programmes de traitement employés les débits sont donnés, soit sous forme de débits instantanés, soit sous forme de moyennes à une échelle de temps quelconque (débits moyens journaliers, mensuels et annuels). L'amplitude de l'échelle des débits est très grande puisqu'elle peut varier de moins d'un litre par seconde, à plusieurs dizaines de milliers de m^3/s et elle pose donc le problème du volume du fichier. Afin de compacter l'information tout en gardant la précision souhaitable, l'ORSTOM a créé un système de notation exponentielle spéciale couvrant toute la gamme des débits possibles.

Cette notation, qui donne habituellement les débits en m^3/s (JUDEBI = 0) comprend seulement 4 chiffres :

- 3 chiffres significatifs XXX
- 1 chiffre exposant A

Tout débit représenté sous la forme $Q = 0,XXX 10^A$ s'écrira XXXA avec cette notation.

.

A titre d'exemple :

$$Q = 0,001 \text{ m}^3/\text{s} \text{ donne } 0010$$

$$Q = 0,152 \text{ m}^3/\text{s} \text{ donne } 1520$$

$$Q = 14,9 \text{ m}^3/\text{s} \text{ donne } 1492$$

$$Q = 12\,600 \text{ m}^3/\text{s} \text{ donne } 1265$$

Pour certaines études de débits d'étiage, où les dispositifs de mesure permettent de définir de très faibles débits, ceux-ci peuvent être donnés en l/s (JUDEBI = 1) en modifiant simplement la valeur du chiffre exposant A qui devient A + 3.

Quel que soit le système employé, les débits manquants sont toujours représentés par la valeur - 10.

A l'inverse du fichier de données brutes, qui est normalement établi une fois pour toutes et qui représente l'acquis hydrologique, le fichier de données traitées est modifiable à tout instant (si l'on s'aperçoit par exemple qu'un étalonnage est erroné).

ETAT	BASSIN	RIVIERE	STATION	CAMEROUN ETAT	BENOUÉ BASSIN	BENOUÉ RIVIERE	RIAO NOM de la STATION
STATION HYDROMETRIQUE - CARTE D'IDENTIFICATION							
0	0	0	0	0	0	0	0
1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64
65	66	67	68	69	70	71	72
73	74	75	76	77	78	79	80
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9
1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64
65	66	67	68	69	70	71	72
73	74	75	76	77	78	79	80
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9
1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64
65	66	67	68	69	70	71	72
73	74	75	76	77	78	79	80

IBM 6042

O.R.S.T.O.M. - Hydrologie

MODELE COH 302

ETAT	BASSIN	RIVIERE	STATION	ANNEE	MOIS	GROUPE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
HAUTEURS LIMNIMETRIQUES (Lues en m)																						
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79	80	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79	80	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79	80	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

IBM 6044

O.R.S.T.O.M. - Hydrologie

MODELE COH 301

Figure 1

