Zeitschrift: Bulletin de la Société Neuchâteloise des Sciences Naturelles

Herausgeber: Société Neuchâteloise des Sciences Naturelles

Band: 99 (1976)

Artikel: Emploi des structures de listes dans la gestion d'une banque de données sur un ordinateur de 16 K. : application à l'étude des eaux du lac de Neuchâtel

Autor: Humbert, Marc / Dubois, Jean-Pascal

DOI: https://doi.org/10.5169/seals-89097

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 15.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

INSTITUT DE MATHÉMATIQUES, INSTITUT DE GÉOLOGIE,
UNIVERSITÉ DE NEUCHATEL

EMPLOI DES STRUCTURES DE LISTES DANS LA GESTION D'UNE BANQUE DE DONNÉES SUR UN ORDINATEUR DE 16 K. APPLICATION A L'ÉTUDE DES EAUX DU LAC DE NEUCHATEL

par

MARC HUMBERT [1] et JEAN-PASCAL DUBOIS [2]

AVEC 5 FIGURES

1. Introduction

Depuis 1957, le Laboratoire cantonal neuchâtelois poursuit un vaste programme d'analyse des eaux du lac de Neuchâtel. L'interprétation d'un tel volume de résultats (environ 200.000 données) a nécessité la création d'une banque de données rapidement accessibles et permettant en tout temps de traiter n'importe quelle série de valeurs selon diverses techniques.

Ce travail est le fruit d'une collaboration étroite entre l'Institut de géologie et le Centre de calcul électronique de l'Université de Neuchâtel.

L'originalité de la méthode permet de résoudre la gestion d'une banque de données malgré les problèmes inhérents à un ordinateur de petite capacité.

2. L'ordinateur

L'ordinateur du centre de calcul électronique (CCE) de l'Université de Neuchâtel est un IBM 1130 de 16 K (mémoire centrale contenant 16.000 mots). Les données, perforées sur cartes, sont enregistrées sur un disque de 512 K. Quant aux graphiques, ils sont dessinés par un traceur de courbe CIL 6011.

[1] Centre de calcul électronique.
[2] Laboratoire de minéralogie, pétrographie et géochimie.

3. Gestion de la banque de données

L'étude de l'état sanitaire du lac est caractéristique du genre de problèmes posés à tous les chercheurs analysant plusieurs types de paramètres dans l'espace et dans le temps. Lorsque ce volume d'informations est suffisamment important, il devient indispensable de créer une banque de données structurée en fonction du mode d'exploitation ultérieure des résultats.

Dans notre cas, les analyses sont mémorisées sous la forme de vecteurs regroupant tous les paramètres observés pour un lieu de prélèvement, à une date et une profondeur données.

LIEU	ANNÉE	PROFONDEUR	INFORMATIONS + POINTEURS

Lors de leur insertion dans la banque de données, les vecteurs sont enregistrés, par années, dans des fichiers séparés. On divise ainsi le nombre total d'informations en plusieurs « blocs » de dimensions réduites, caractérisés par un paramètre dit « principal ».

A l'intérieur de chaque fichier, on crée une « structure de liste » pour tout paramètre secondaire, c'est-à-dire un *système d'adressage* reliant, entre elles, toutes les analyses prélevées en un même lieu, respectivement à une même profondeur.

Une telle organisation évite à l'ordinateur le tri systématique de chaque enregistrement et, par conséquent, accélère considérablement la recherche des données, quelles que soient les exigences.

Le choix de l'année comme paramètre principal, du lieu et de la profondeur comme paramètres secondaires, est justifié par la répartition spatiale et temporelle des analyses. Dans le cas du lac, par exemple, le nombre d'informations varie peu d'une année à l'autre (la division en fichiers annuels constitue des « blocs » de dimensions similaires), alors que leur distribution par lieu de prélèvement ou par profondeur est très disproportionnée. (Certaines stations sont visitées deux fois par an, alors que d'autres le sont plusieurs fois par mois.)

Le traitement des résultats représente un autre axiome de choix. Dans notre cas, il justifie également la classification établie, puisque les problèmes à résoudre sont fréquemment relatifs à l'évolution annuelle d'un paramètre le long d'un profil (recherche des valeurs pour une année et une station à toutes les profondeurs), ou à une profondeur donnée.

Exemple

On désire étudier l'évolution de la teneur en phosphate de la baie d'Auvernier en 1966.

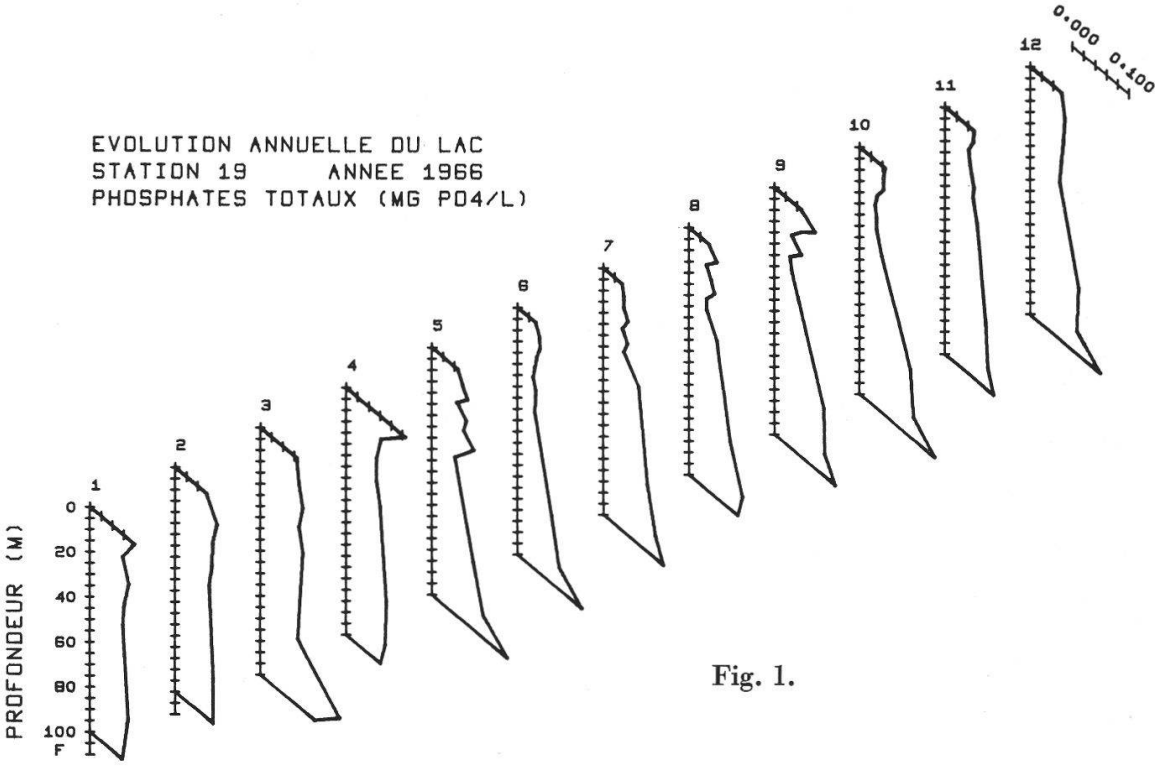

Fig. 1.

Avant de dessiner le graphique, l'ordinateur lit une carte de données sur laquelle est précisée la nature des exigences, soit : 1966, baie d'Auvernier, toutes les profondeurs, phosphates.

Muni de ces renseignements et grâce à la structure de liste décrite en détail au paragraphe 4, il recherche, dans le fichier 1966, toutes les valeurs relatives aux phosphates et à la station d'Auvernier. Il ne lui reste qu'à les trier chronologiquement pour pouvoir dessiner le graphique.

4. Fonctionnement des structures de listes

Pour décrire le fonctionnement des structures de listes, on introduit les notations suivantes : le *fichier annuel*, qui contient les observations et leurs enchaînements, et les *fichiers auxiliaires* facilitant l'entrée dans ce fichier.

4.1. Fichier annuel

On représente un fichier annuel par : fichier (I, J).

Convention : la conception du fichier étant identique pour toutes les années, on simplifie l'écriture en supprimant le paramètre principal ; d'où la notation fichier (I, J), en lettres minuscules. Lorsque le paramètre principal est nécessaire, on lui attribue la valeur Y.

I est l'adresse d'un enregistrement dans le fichier, il prend les valeurs A 1, A 2, A 3, ... AL.

J est l'indice d'une composante du vecteur mémorisé, c'est-à-dire une colonne du fichier. Les valeurs utilisées sont :

PX indice de la colonne contenant le lieu,

PZ indice de la colonne contenant la profondeur,

P 1 (P 2) indice de la colonne contenant les pointeurs de la structure de liste selon les lieux (profondeurs),

P 3 indice de la colonne contenant les pointeurs de la structure de liste liant les cellules libres.

Définition : On appelle « pointeur », l'adresse, dans le fichier du prochain élément de la liste.

4.2. Fichiers auxiliaires

Pour chaque année et pour chaque lieu ou profondeur, il est nécessaire de connaître l'adresse de la première analyse de la liste concernée. De même, pour chaque fichier annuel, il est nécessaire de connaître l'adresse de la première cellule vide. On mémorise ces pointeurs dans trois fichiers auxiliaires, soit :

FIL (Y, X) pointeur du lieu X dans le fichier Y,

FIP (Y, Z) pointeur de la profondeur Z dans le fichier Y,

POINT (Y) adresse de la première cellule vide du fichier Y,

où X, Z et Y représentent tous les lieux et toutes les profondeurs pour toutes les années.

A partir de cette « adresse d'entrée », les listes sont chaînées jusqu'au dernier élément dont le pointeur est nul.

4.3. Description

La figure 2 illustre l'organisation d'un fichier annuel Y et ses entrées à partir des fichiers auxiliaires ; chaque ligne est un enregistrement dont on ne représente que les coordonnées et les structures d'adresses.

a) FIP (Y, Z 3) = 0 pas d'enregistrement à l'année Y et à la profondeur Z 3.

b) FIL (Y, X 3) = A 4 la première analyse du fichier pour la station X 3 se trouve à l'adresse A 4 ;
fichier (A 4, P 1) = 0 pas d'autre analyse pour X 3.

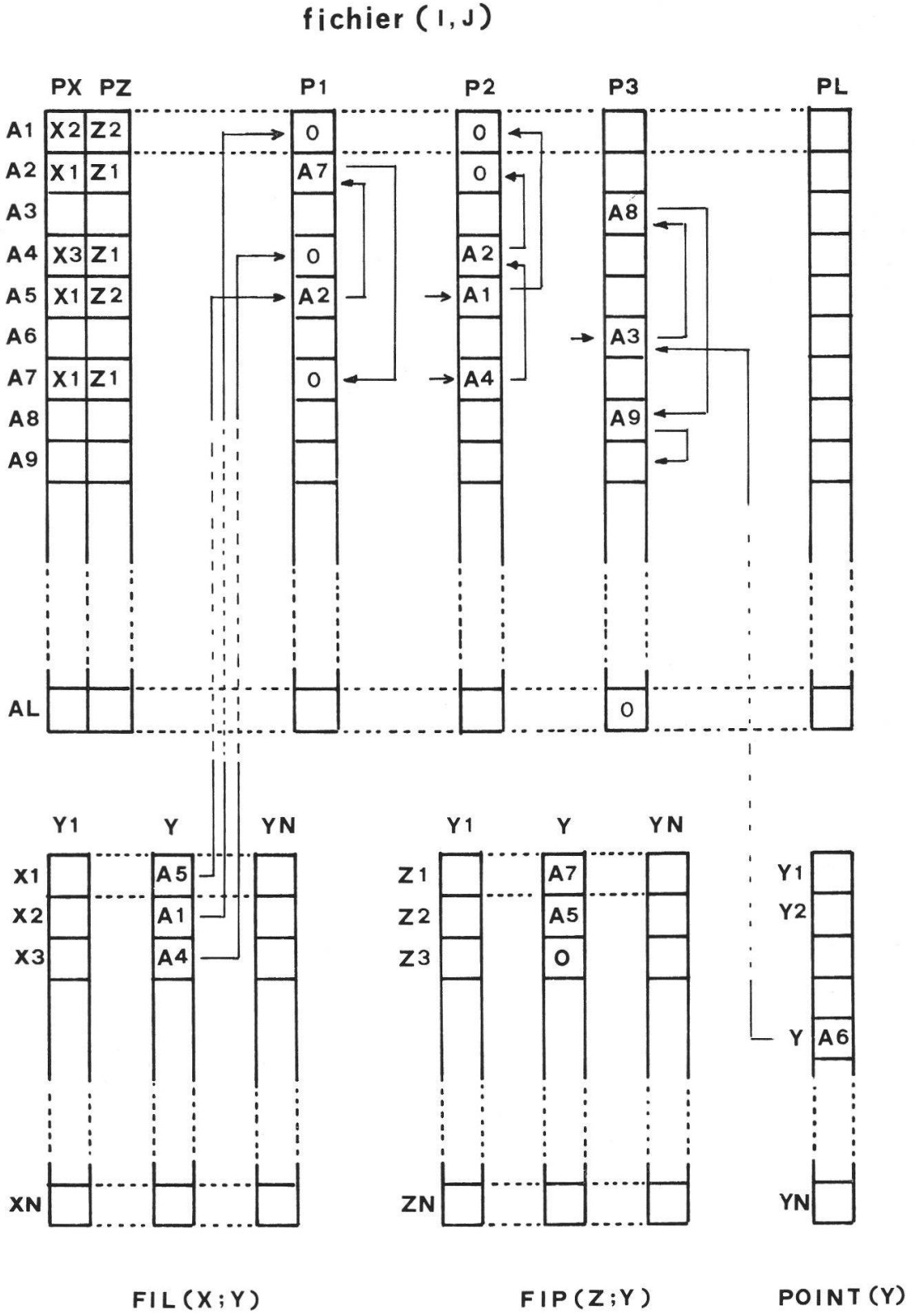

Fig. 2. Fonctionnement des pointeurs dans le fichier annuel Y à partir des fichiers auxiliaires FIL, FIP et POINT.

c) FIL (Y, X 1) = A 5 adresse de la première analyse concernant la station X 1 ;
fichier (A 5, P 1) = A 2 adresse de la deuxième analyse,
fichier (A 2, P 1) = A 7 adresse de la troisième analyse,
fichier (A 7, P 1) = 0 pas d'autre analyse.

d) POINT (Y) = A 6 adresse de la première cellule libre dans le fichier de l'année Y ;
fichier (A 6, P 3) = A 3 adresse de la deuxième,
fichier (A 3, P 3) = A 8 adresse de la troisième,
etc.,
fichier (A L, P 3) = 0 cette cellule est la dernière.

Un tel système est utilisé pour les recherches. De plus, il doit être en mesure de recevoir et (ou) de supprimer des informations. Les paragraphes suivants donnent des indications sur la manière de le faire tout en maintenant les structures.

4.4. Opération d'introduction d'une analyse

La figure 3 montre l'état du fichier après l'introduction de l'analyse de coordonnées (X 1, Y, Z 3). Les valeurs exactes de l'exemple sont explicitées entre parenthèses. Les formules sont valables pour autant qu'il existe une cellule libre. (Pour en avoir la certitude, il suffit de tester la valeur de IA qui doit être différente de 0.) L'algorithme de cette insertion se résume aux formules suivantes :

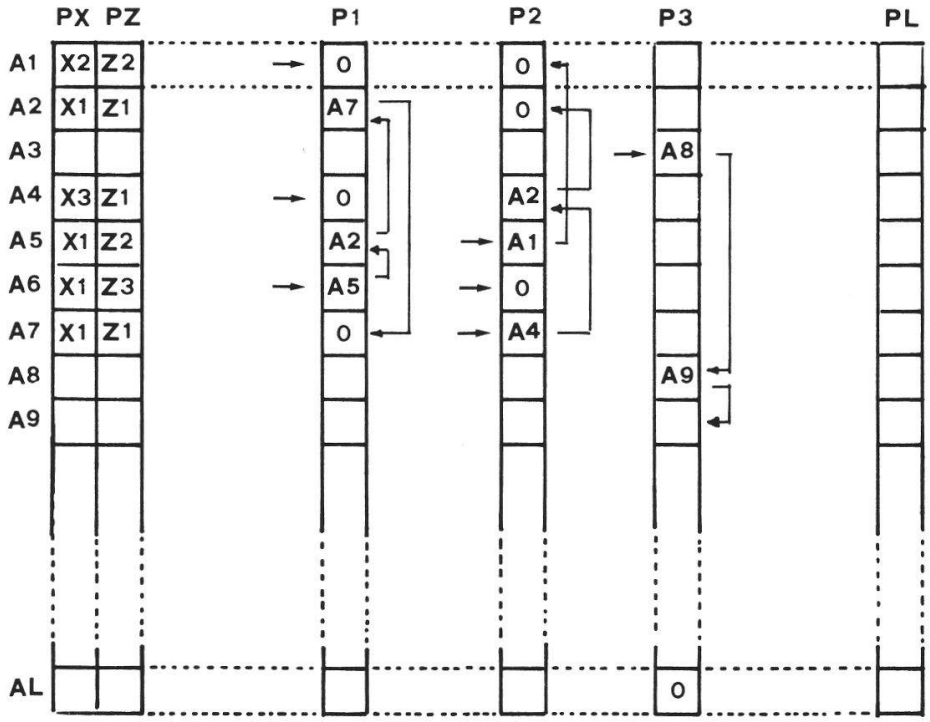

Fig. 3. Etat du fichier annuel Y après introduction de l'analyse de coordonnées (X 1, Y, Z 3).

a) Recherche de l'adresse d'une cellule libre, IA, et correction du pointeur de la liste.
IA = POINT (Y) (IA = A 6)
POINT (Y) = fichier (IA, P 3) (POINT (Y) = A 3)

b) Correction des pointeurs dans la liste selon X 1.
fichier (IA, P 1) = FIL (Y, X 1) (fichier (A 6, P 1) = A 5)
FIL (Y, X 1) = IA (FIL (Y, X 1) = A 6)

c) Correction des pointeurs dans la liste selon Z 3.
fichier (IA, P 2) = FIP (Y, Z 3) (fichier (A 6, P 2) = 0)
FIP (Y, Z 3) = IA (FIP (Y, Z 3) = A 6)

4.5. Opération de suppression d'une analyse

L'opération de suppression d'un élément est la plus délicate du maintien des fichiers. En effet, pour supprimer l'élément de coordonnées (X 1, Y, Z 2), il s'agit :

a) de déterminer son adresse dans le fichier ;

b) de retrouver les analyses dont les pointeurs contiennent son adresse, afin de corriger les structures de listes (la correction ne portant que sur FIL ou FIP, si l'élément est le premier) ;

c) de récupérer son adresse dans la liste des cellules vides.

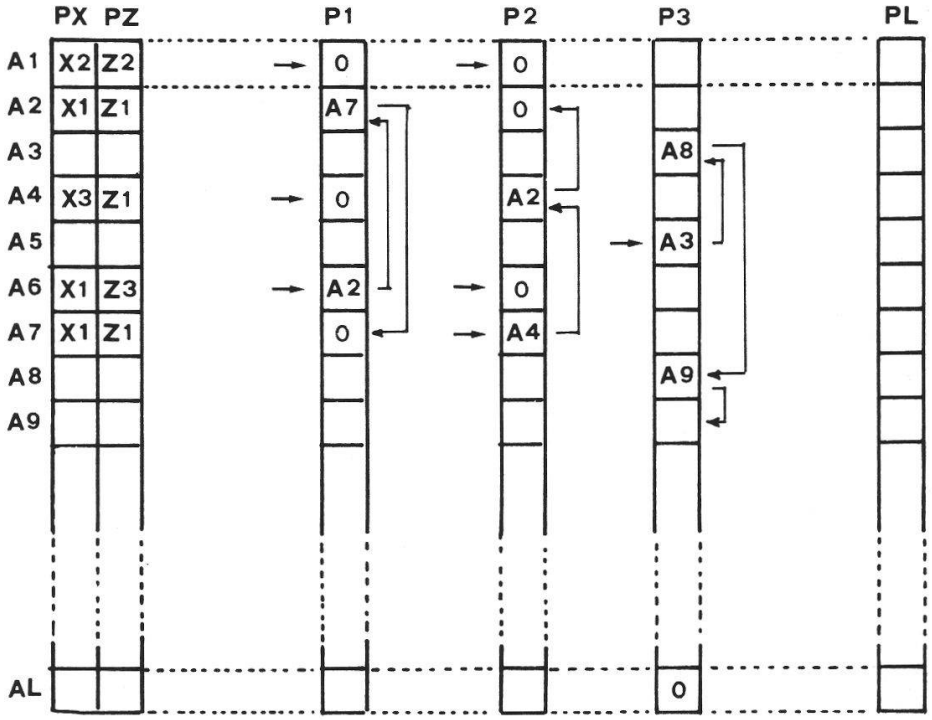

Fig. 4. Etat du fichier annuel Y après suppression de l'analyse de coordonnées (X 1, Y, Z 2).

La figure 4 montre le résultat final de l'opération. Le lecteur trouvera au paragraphe 5 l'organigramme et les commentaires relatifs à la suppression d'un élément.

5. Complément technique

On a remarqué que la suppression est une opération délicate. Le lecteur intéressé trouvera, dans ce paragraphe, le déroulement explicite du travail de recherche et de correction fourni par l'ordinateur.

L'organigramme de la figure 5 utilise les variables suivantes :

IR adresse de l'enregistrement à supprimer ;

IL adresse de la précédente analyse concernant le même lieu X 1 ;

IP adresse de la précédente analyse concernant la même profondeur Z 2.

Les autres définitions correspondent à la description des fichiers.

Cases 1 à 2

IR prend initialement l'adresse de la première analyse concernant le lieu X 1, puis parcourt la liste pour chercher l'enregistrement dont la deuxième coordonnée sera Z 2.

Les sorties sur α indiquent qu'il n'existe pas d'élément correspondant aux coordonnées (X 1, Z 2).

La séquence IL = IR
IR = fichier (IR, P 1)
résume le transfert des pointeurs ; IL mémorise l'adresse du dernier élément testé, IR prend l'adresse du prochain élément de la liste.

Cases 2 à 3

IP prend initialement l'adresse de la première analyse concernant la profondeur Z 2, puis parcourt la liste pour chercher l'enregistrement dont le pointeur du lieu sera IR.

Les sorties sur β (et les tests s'y référant) ne sont en principe pas nécessaires, puisque l'élément existe. Elles préviennent toute autre source d'erreur.

Les pointeurs sont modifiés par les opérations :

FIP (Y, P 2) = fichier (IR, P 2)

si IP = IR, c'est-à-dire si l'analyse est la première de la liste et que le pointeur à modifier se trouve dans le fichier auxiliaire FIP ;

fichier (IP, P 2) = fichier (IR, P 2)

si IP $\neq$ IR, c'est-à-dire si l'analyse se situe dans la liste et que le pointeur à modifier se trouve dans le fichier annuel.

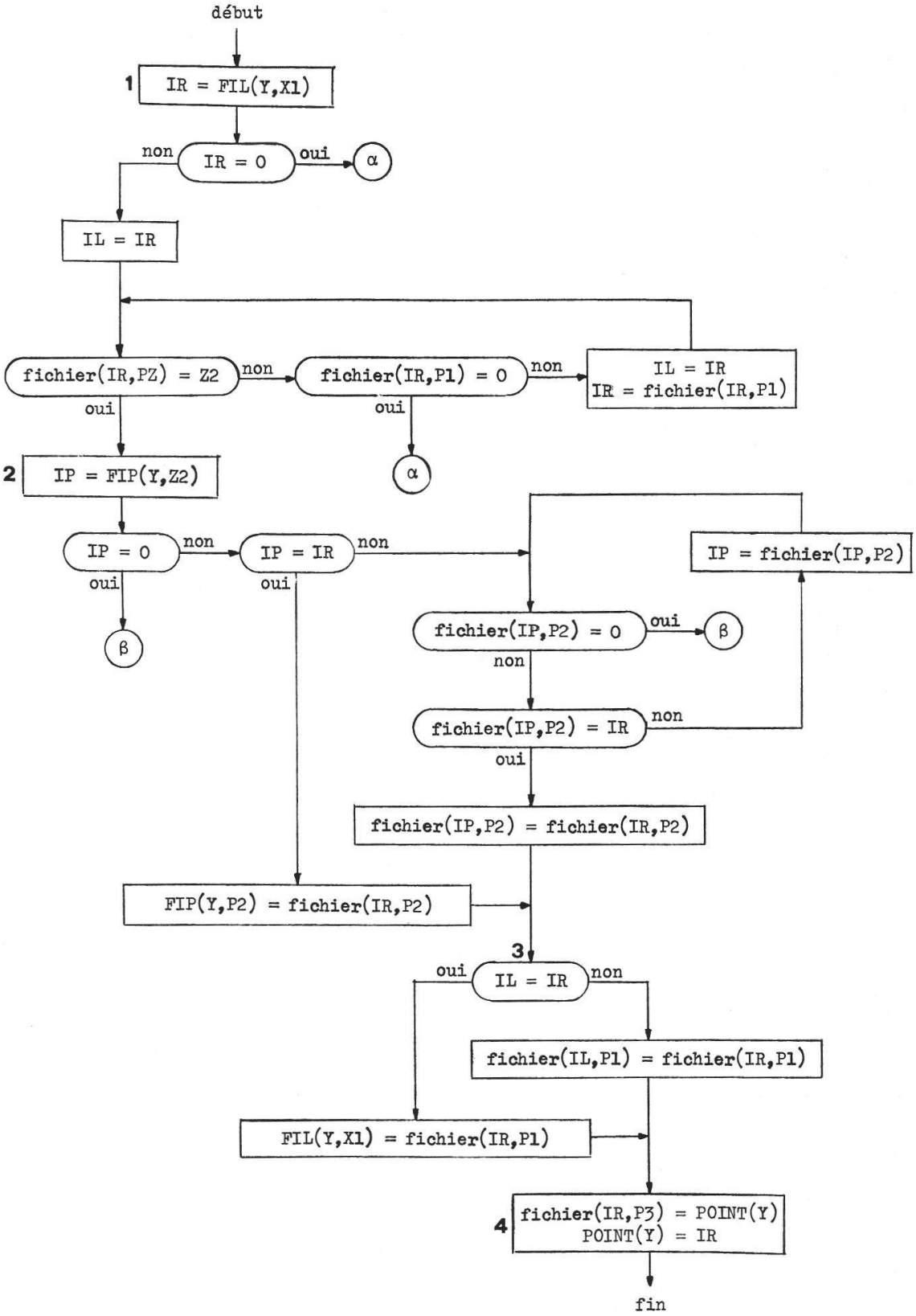

Fig. 5. Organigramme représentant la recherche puis la suppression de l'enregistrement de coordonnées (X 1, Y, Z 2).

Cases 3 à 4

Modification des pointeurs de la liste selon les lieux avec les mêmes considérations que pour les profondeurs. Ces opérations n'ont pas été effectuées avant la case 2, afin de ne pas devoir corriger en cas de sortie sur β.

Case 4

Enchaînement de l'enregistrement IR à la structure d'adresse des cellules vides et modification du pointeur d'entrée dans le fichier annuel Y.

6. Conclusions

Le système développé ici, bien qu'adapté à un problème précis, peut être étendu sans difficultés à tout problème de gestion de données caractérisées par plusieurs coordonnées (temps, lieu et profondeur, par exemple). Dans de tels cas, on étudiera préalablement les différents types de traitements prévisibles et, si possible, la répartition des analyses, afin de choisir valablement le paramètre principal caractérisant chaque fichier et les paramètres secondaires pour lesquels une structure de liste est utile.

Afin de ne perdre aucune place, lors de l'effacement d'un enregistrement, il conviendra de créer, par fichier, une liste d'adresses supplémentaires reliant, entre elles, les cellules vides ; cette opération n'est pas faite automatiquement par l'ordinateur.

Résumé

La gestion d'une banque de données dépend du mode d'exploitation prévisible des résultats et de la priorité accordée aux différents paramètres.

L'utilisation de structures de listes (adressage automatique des cellules de la mémoire) accélère la recherche des analyses et permet l'autogestion des fichiers (introduction, correction ou suppression d'un enregistrement). La méthode exposée a été appliquée avec succès à l'étude des eaux du lac de Neuchâtel.

Zusammenfassung

Die Führung einer Datenbank hängt stark von der Art ab, wie die Resultate ausgewertet werden, sowie von der Wichtigkeit die den Parametern beigemessen wird. Der systematische Gebrauch von File-Strukturen (mit automatischer Adressierung der Speicherzellen) hat die

Beschleunigung der Abfragung der Bank zur Folge, dazu eine weitgehende Automatisierung der File-Operationen (Einschieben, Ändern bzw. Löschen einer Eintragung). Die hier beschriebene Methode wurde für die Untersuchung des Wassers im Neuenburgersee mit Erfolg angewendet.

Summary

Databank management works according to the manner in which the processing of results is planned and depends on the priority given to various parameters. By using file-structures (automatic addressing of the memory cells) particular analyses are retrieved more rapidly and auto-management of files can be achieved (introduction, corrections or suppression of records). The water of the lake of Neuchâtel has been studied successfully by applying the above method of processing.

BIBLIOGRAPHIE

FORSYTHE, A., KEENAN, T., ORGANICK, E. et STENBERG, W. — (1969). Computer Science, A First Course (John Wiley & Sons, Inc.).

KNUTH, D. E. — (1969). The Art of Computer Programming, vol. 1 (Addison-Wesley Publishing Company).

WIRTH, N. — (1975). Algorithmen und Datenstrukturen (B. G. Teubner, Stuttgart).