

GEOMER : BASE DE DONNEES DE GEOPHYSIQUE MARINE

François Missègue
(Géophysique)
(UR 106)

Centre ORSTOM de Nouméa
BP A5
Nouméa Cédex (Nouvelle Calédonie)

RESUME - GEOMER est une base de données de géophysique marine. La mise en oeuvre a été motivée par la masse de données de plus en plus importante devenue très difficile à manipuler. Le seul impératif était d'avoir un accès direct aux données par croisière, par profil ou par fenêtre géographique.

L'ossature choisie est une grille géographique sur laquelle un module de base se positionne. La donnée de base est un point de mesure. Il est lié à un module par sa position dans l'espace et à une croisière par son temps. Des réseaux de pointeurs interactifs permettent d'identifier le module récepteur et de positionner les points de mesure sur les fichiers. Quel que soit le mode d'extraction choisi, l'opération est une recherche spatiale à sélectivité variable.

La base est implantée sur HP1000 et dispose de 30 MO de mémoire de masse. L'utilisation en E/S du format international d'échange MGD77 rend GEOMER compatible avec toutes les grandes bases actuelles. Le codage binaire a permis de réduire au minimum le volume de mémoire de masse.

En développement, tous les logiciels seront traduits en Fortran 77 MicroSoft 3.31 afin d'assurer une bonne portabilité et de permettre une implantation sur compatible IBM PC-AT avec possibilité de mise en réseau local type Omninet et de branchement futur au réseau TOMPAC.

Un logiciel de création est à l'étude. Il permettra une création de base de données à la carte. La grille géographique, le type de données et les structures internes des fichiers devront être définis par l'utilisateur.

1. OBJECTIFS

La base de données GEOMER a été mise en oeuvre afin d'archiver, de conserver et de faciliter l'accès à l'importante masse de données récoltées pendant plus de dix ans par l'équipe de géologie-géophysique ORSTOM de Nouméa. Une première analyse a permis d'établir un cahier des charges en fonction, d'une part, des besoins formulés, et d'autre part des impératifs techniques.

Les impératifs techniques étaient :

- implantation sur le HP 1000 du centre ;
- allocation de 30 MO de mémoire de masse ;
- masse des données à stocker : environ 1.000.000 de points de mesure ;
- une compatibilité des formats E/S permettant des échanges de données avec les bases extérieures existantes.

Les besoins formulés par les futurs utilisateurs se résument à un accès direct aux données selon trois options :

- par fenêtre géographique ;
- par croisière ;
- par profil.

La structuration et l'organisation interne étant laissées au choix du concepteur.

2. DONNEES

Les données incorporées à GEOMER sont des données de géophysique marine. Elles sont situées dans le temps et dans l'espace. La donnée de base est le *point de mesure* qui comporte 11 champs. GEOMER met actuellement à la disposition des utilisateurs environ 350.000 mesures de bathymétrie, 200.000 mesures de gravimétrie et 350.000 mesures de magnétisme.

La cadence actuelle d'incorporation de croisières anciennes qui doivent être entièrement retraitées et des croisières nouvelles, incorporées au fur et à mesure de leur réalisation, permet d'évaluer la masse de données qui sera stockée fin 1988 à environ 1.200.000 mesures de bathymétrie, 1.200.000 mesures de magnétisme et environ 500.000 mesures de gravimétrie.

3. CONCEPTION ET FONCTIONNEMENT

La situation des points de mesure dans le temps et l'espace étant les deux seuls critères d'identification possible, le mode d'incorporation retenu devait respecter la chronologie et la position spatiale des données.

La structure de base de GEOMER est une grille géographique, qui a été volontairement limitée à la zone de travail des utilisateurs actuels. Ses limites sont 10°N et 40°S en latitude, et 90°E et 120°W en longitude (Figure 1).

Cette grille est découpée en modules appelés stations selon un pas de 1/2 degré. Une station correspond à une couverture géographique de 1/4 de degré carré (Figure 2). Le tronçon de croisière entrant dans une station est appelé traversée de station.

Des réseaux de pointeurs interactifs permettent, lors de l'incorporation, de situer tout point de mesure dans l'espace grille et de les retrouver lors des recherches pour extraction.

3.1. Les fichiers

L'organisation de GEOMER s'articule autour de six fichiers principaux : INFOG, DGNDX, DGLST, PROFIL, STATN, CROISR contenant les mises à jour des pointeurs, les réseaux de pointeurs, les informations scientifiques, techniques et administratives des croisières incorporées et les descriptifs des traversées de station et des profils. Les données équipées de leurs différents pointeurs sont stockées dans des fichiers MESxxx. Chaque croisière a son propre fichier MESxxx.

INFOG est un compteur contenant la mise à jour des pointeurs.

DGNDX est la grille géographique. Chaque enregistrement, formé d'un champ, correspond à un des 30.301 1/4 de degré carré. L'adresse de l'enregistrement est calculée à partir de la position géographique du sommet Sud-Ouest du 1/4 de degré carré ou station.

L'état du champ indique si la station a été ou non traversée par une croisière quelconque. Si la station a été traversée, il donne l'adresse à laquelle la première traversée réalisée est répertoriée dans le fichier DGLST.

DGLST contient par ordre chronologique de réalisation toutes les traversées de station. Un enregistrement correspond à une traversée ; il contient toutes les informations permettant :

- d'identifier la croisière au cours de laquelle la traversée de la station a été exécutée ;
- de situer la station sur la grille géographique ;
- d'accéder au fichier descriptif de station STATN ;

GEOMER INCORPORATION SEQUENCE 1

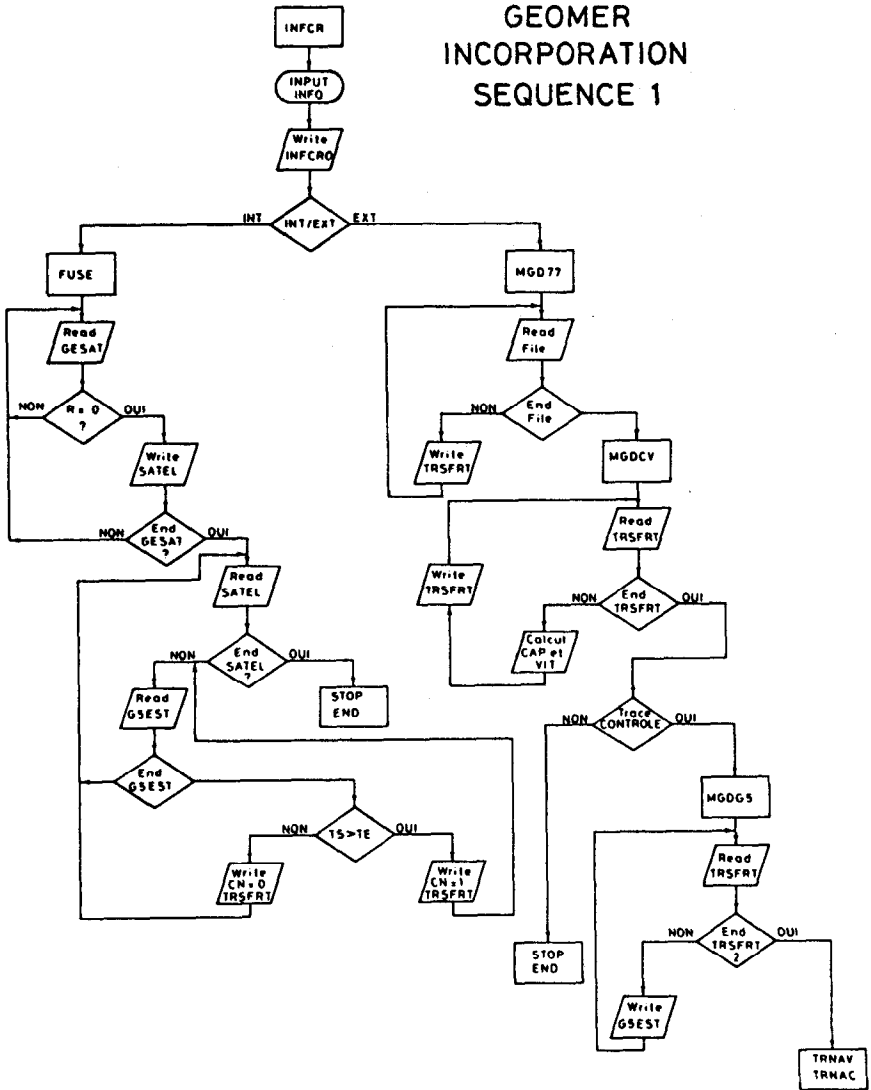


Fig.2

- de savoir grâce à l'état d'un pointeur s'il y a eu d'autres traversées de cette station soit par la même croisière, soit par une autre croisière.

Ce dernier pointeur permet d'accéder directement à l'adresse où est répertoriée la traversée suivante.

PROFIL contient la liste de tous les profils répertoriés. Chaque enregistrement correspond à un profil et contient les informations permettant d'identifier la croisière et le profil et de situer ce dernier dans le temps et l'espace grille.

STATN contient le descriptif de toutes les traversées de station répertoriées. Un enregistrement correspond à une traversée de station. Il contient les informations permettant d'identifier la croisière et la traversée ainsi que la situation dans le temps et l'espace grille de cette dernière.

CROISR contient tous les renseignements scientifiques, techniques et administratifs de chaque croisière incorporée. Ceux-ci indiquent les noms de la croisière, du navire utilisé, du chef de mission et de la zone géographique couverte, le pointeur de croisière, la situation dans le temps et dans l'espace grille de la croisière, la liste descriptive des moyens mis en oeuvre et la liste des bases gravimétriques sur lesquelles la croisière est calée.

3.2. Organisation et principe de fonctionnement

Incorporation L'incorporation d'une croisière se réalise selon les phases suivantes :

- allocations des pointeurs de croisière, de station, de profil et de traversée de station ;
- création du fichier MESxxx ;
- activation des réseaux de pointeurs ;
- identification du premier point de mesure ;
- positionnement du point sur l'espace grille ;
- recherche de la station support ;
- identification de la traversée de station ;
- positionnement et notation de la traversée ;
- allocation des pointeurs de croisière et de station au point traité ;
- mise à jour des pointeurs dans INFOG ;
- mise à jour des fichiers STATN, PROFIL, DGNDX et DGLST ;
- incorporation du point au fichier MESxxx ;
- traitement du point de mesure suivant.

Quand tous les points de mesure sont traités et incorporés :

- mise à jour du fichier CROISR.

Extraction Lorsqu'une opération d'extraction est lancée, le mode d'extraction est laissé au choix de l'utilisateur. Quel que soit le mode choisi, fenêtre géographique, croisière ou profil, l'opération finale est la même et seuls la sélectivité et le temps de recherche varient.

Une extraction se déroule selon les phases suivantes :

- détermination des stations à consulter ;
- identification de toutes les traversées de ces stations ;
- positionnement des traversées dans l'espace grille ;
- élaboration d'une liste des traversées sélectionnées, tri et mise en ordre chronologique de croisière et de traversée ;
- recherche de tous les pointeurs permettant un accès direct aux données à extraire des différents fichiers MESxxx ;
- positionnement des points de mesure dans l'espace grille ;
- extraction des données ;
- transfert sur bande magnétique.

3.3. Positionnement des données dans l'espace-grille

De toutes les opérations et phases d'opération citées, seules celles concernant le positionnement dans l'espace grille, soit des traversées de station, soit des points de mesure, sont d'un intérêt particulier et mettent en valeur l'ossature de la base sous la forme d'une grille géographique et les relations interactives de deux réseaux de pointeurs.

Incorporation (Figure 3) La position d'un point de mesure dans l'espace géographique permet de situer celui-ci sur la grille GEOMER et sur une station déterminée.

Quand la station support est déterminée, on connaît le numéro de l'enregistrement de DGNDX où elle est répertoriée, autrement dit, on la situe sur l'espace grille et les séquences suivantes sont exécutées :

- lecture de DGNDX ;
- si le pointeur $INDX = 0$, la station est vierge et n'a jamais été traversée par une croisière ;
- la nouvelle traversée est identifiée à l'aide du pointeur de croisière et de numéro chronologique de traversées de station au cours du déroulement de la croisière. Ce dernier est automatiquement incrémenté à chaque changement de station ;
- la traversée est située dans l'espace grille en attribuant à $INDX$ la valeur de $PLST$ le pointeur de liste des traversées ;
- la traversée est répertoriée sur $DGLST$ à l'adresse $PLST$ après assignation du pointeur de station suivante $PSTS$ à la valeur zéro, la traversée étant pour l'instant unique ;

GEOMER INCORPORATION SEQUENCE 2

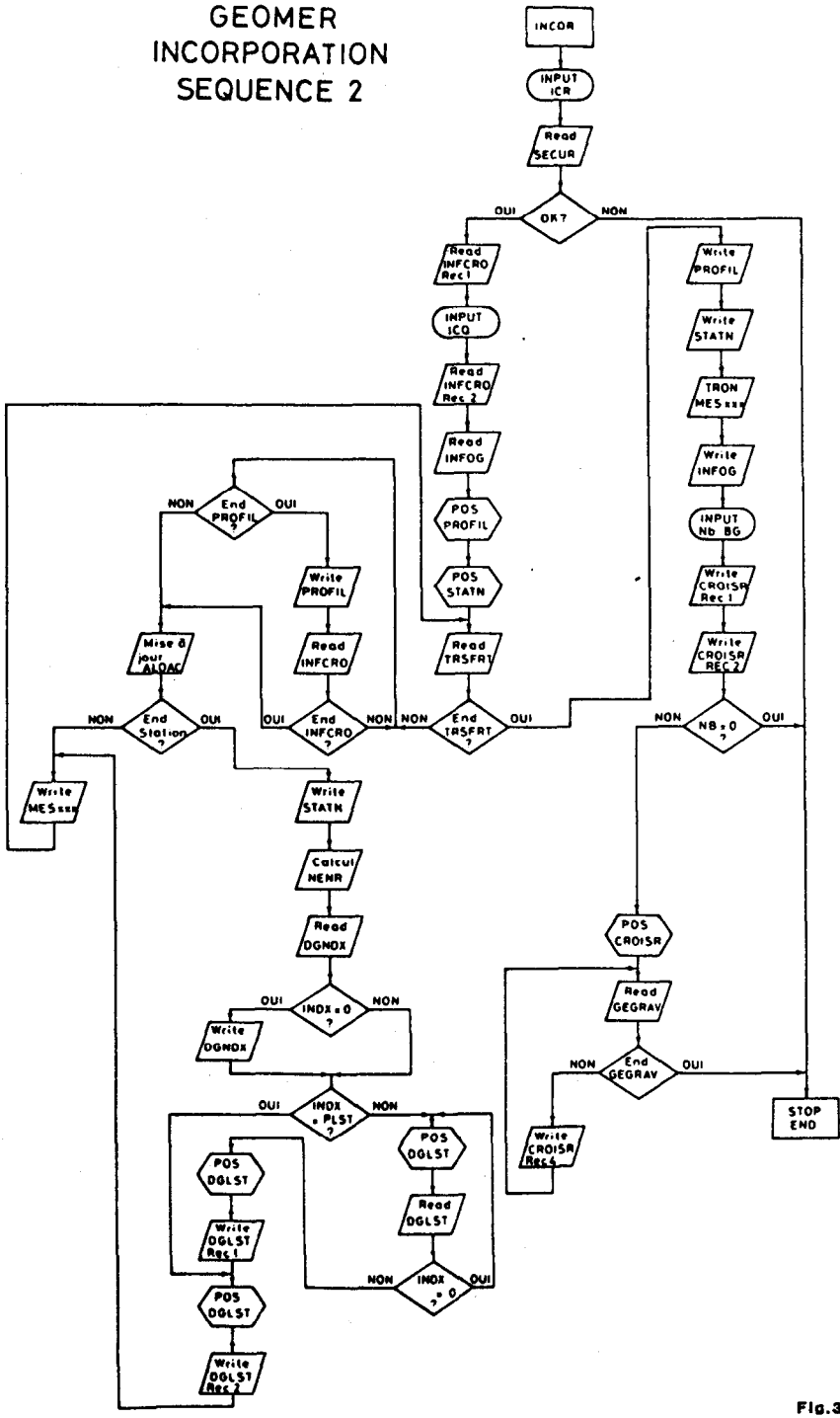


Fig. 3

- si $INDX$ est différent de 0, la station a déjà été traversée par une croisière et la valeur de $INDX$ indique l'adresse de $DGLST$ à laquelle est répertoriée la première traversée. Les séquences suivantes sont alors réalisées : identification de la nouvelle traversée et lecture de $DGLST$ à l'adresse $INDX$.

Si $PSTS=0$ il n'y a eu qu'une seule traversée : la valeur de $PLST$ est assignée à $PSTS$, la traversée est répertoriée à l'adresse $PLST$ avec mise à zéro du nouveau $PSTS$, $PLST$ est incrémenté et il y a passage au point suivant.

Si $PSTS$ est différent de zéro, il y a une autre traversée de répertoriée à l'adresse $PSTS$: $DGLST$ est lu à l'adresse $PSTS$. Si le nouveau $PSTS$ lu est différent de zéro, la lecture continue comme précédemment jusqu'à ce que l'enregistrement portant $PSTS=0$ soit trouvé, autrement dit, jusqu'à ce que toutes les traversées répertoriées aient été identifiées. Lorsque l'on a trouvé la dernière traversée de la station répertoriée, on attribue à $PSTS$ la valeur $PLST$, la traversée est répertoriée à l'adresse $PLST$ avec $PSTS=0$, $PLST$ est incrémenté et il y a passage au point suivant.

Quand le point de mesure suivant a été acquis, son appartenance à la même station est contrôlée. Si ce contrôle est positif le point se voit attribuer les mêmes identificateurs que le point précédent et il ne sera alors particularisé que par son temps au sein d'une même traversée. Si le contrôle est négatif, toutes les opérations précédentes sont à nouveau exécutées.

Extraction (Figure 4) La phase la plus importante de l'extraction est la recherche de toutes les traversées de station réalisées au cours d'une croisière ou d'un profil ou dans une fenêtre géographique.

Dans le cas d'une croisière ou d'un profil, la liste des traversées est réalisée par simple consultation des fichiers $CROISR$, $PROFIL$ et $STATN$.

Dans le cas d'une fenêtre géographique, il s'agit de rechercher toutes les traversées de station, quelles que soient les croisières au cours desquelles elles ont été exécutées. Autrement dit, il faut accéder à toutes les données existantes entrant dans le cadre de la fenêtre, sans aucune omission possible. Pour cela, un balayage de la fenêtre station par station est nécessaire. Pour chaque station entrant dans le cadre de la fenêtre les opérations suivantes sont exécutées : identification de la station et lecture de $DGNDX$ à l'adresse $NREC$.

Si $INDX=0$, station vierge et passage à la station suivante. Si $INDX$ est différent de 0, il y a au moins une traversée : lecture de $DGLST$ à l'adresse $INDX$ et extraction de la traversée. Si $PSTS=0$, il n'y a pas d'autre traversée (passage à la station suivante). Si

GEOMER
EXTRACTION
SEQUENCE 1

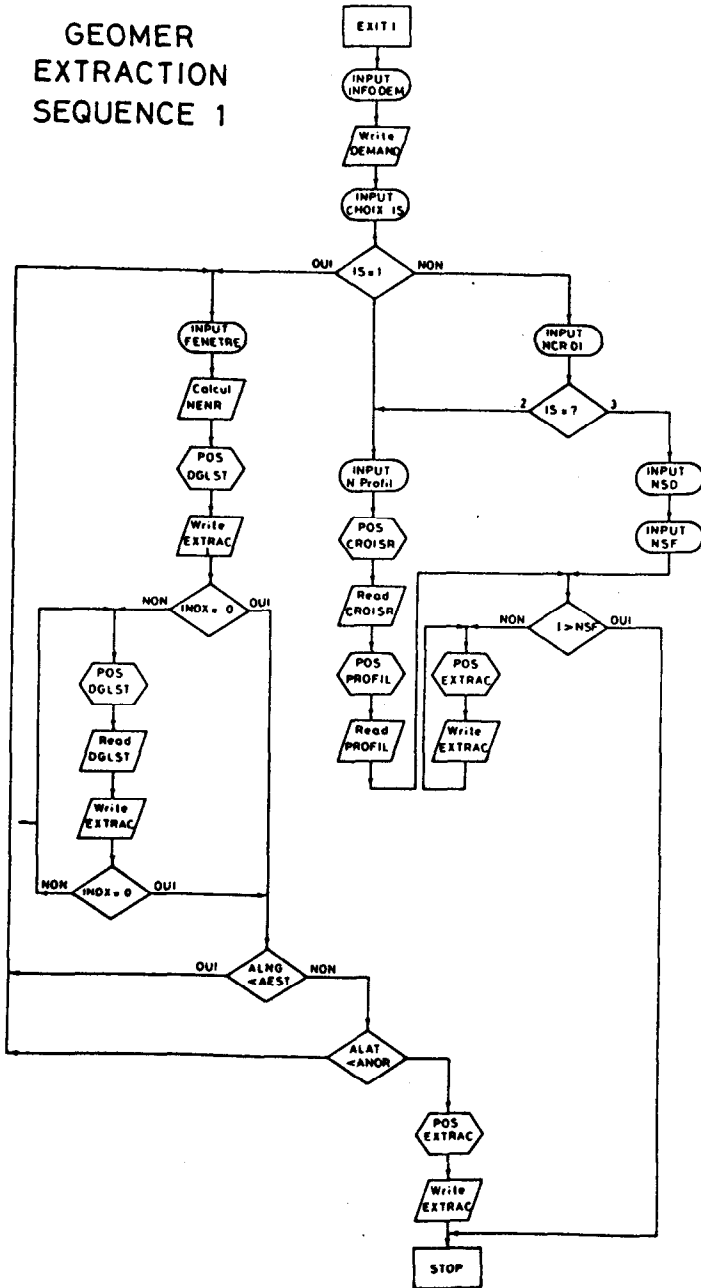


Fig. 4

GEOMER EXTRACTION SEQUENCES 2 & 3

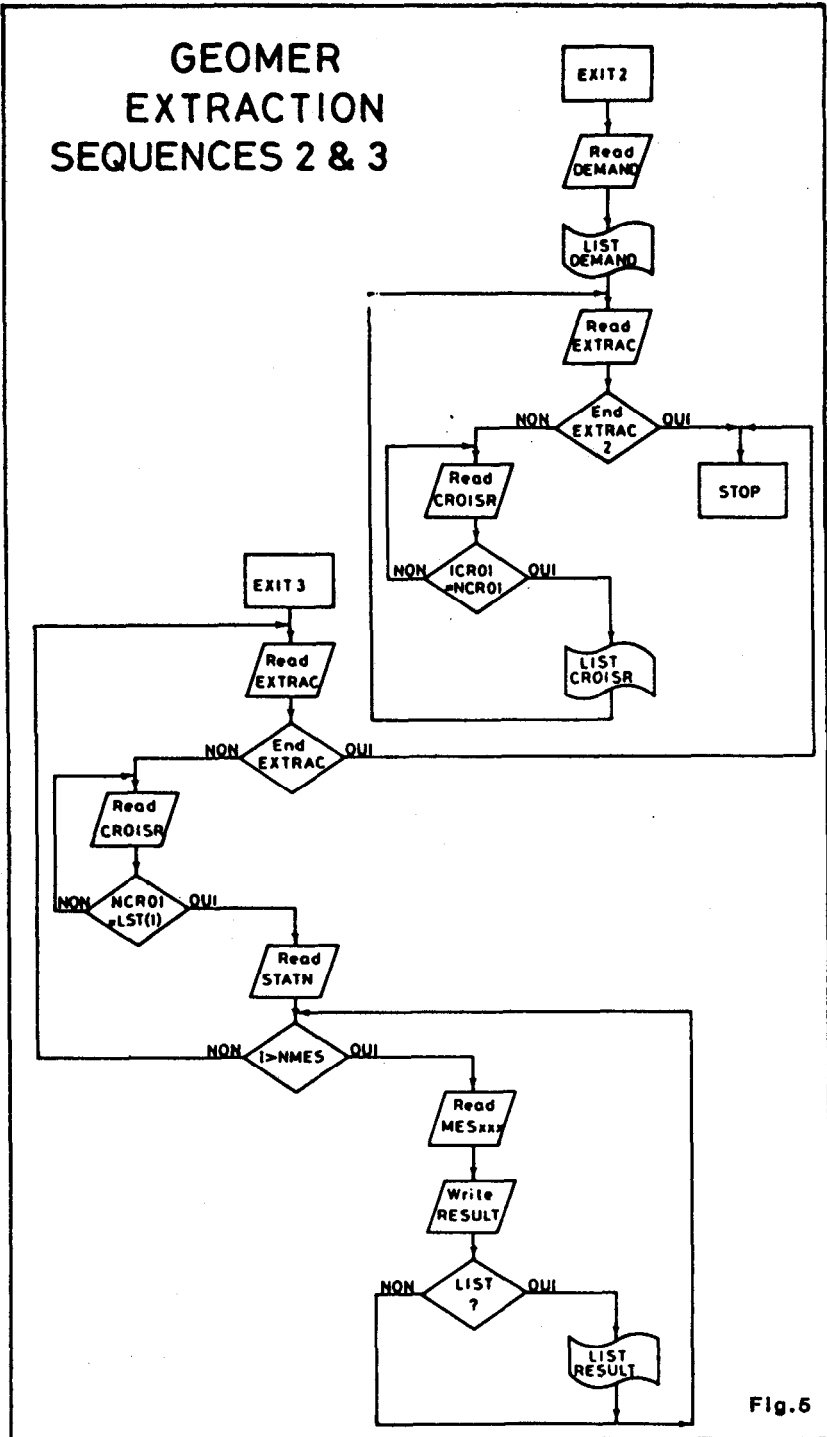


Fig.5

PSTS est différent de 0 il y a encore, au moins une traversée : accès à la traversée suivante à l'adresse PSTS et extraction de cette traversée ; répétition de cette séquence jusqu'à ce que toutes les traversées répertoriées aient été identifiées et extraites ; finalement passage à la station suivante.

Il n'est pas possible d'arrêter la répétition de cette dernière séquence tant que toutes les traversées n'ont pas été identifiées et extraites.

Lorsque la liste de toutes les traversées est établie, celle-ci est triée et les traversées sont classées par ordre chronologique des identificateurs de traversée.

Le premier terme de l'identificateur de traversée permet d'identifier le fichier MES_{xxx} dans lequel sont stockés les points de mesure à extraire.

Le second terme de l'identificateur donne l'adresse du fichier STATN où pourra être lu, d'une part, le pointeur d'accès au premier point de mesure à extraire et d'autre part le nombre de points de mesure répertoriés au cours de la traversée.

Quand tous les points de mesure d'une traversée de station sont extraits et répertoriés dans le fichier RESULT, on passe à la traversée suivante jusqu'à épuisement de la liste.

4. UTILISATION

Comme pour toutes les bases de données, l'accès en écriture est protégé et l'administrateur de la base ou des opérateurs agréés sont les seuls à pouvoir réaliser les opérations d'incorporation.

Les opérations d'extraction, telles qu'elles viennent d'être décrites sont totalement transparentes pour l'utilisateur. Celui-ci n'a qu'à suivre une procédure simple et limitée à la saisie des informations définissant l'extraction à réaliser.

Le transfert des données extraites, sur bande magnétique, est ensuite réalisé soit en format interne soit en format MGD77 si le demandeur est extérieur.

Les demandes d'extraction provenant de l'extérieur sont réalisées par un opérateur agréé.

Tous les échanges de données se font sous le format international d'échange de données de géophysique marine MGD77 mis au point par le *National Geophysical and Solar-terrestrial Data Center Boulder Colorado* dépendant de la *National Oceanic and Atmospheric Administration de l'U.S. Dpt of Commerce*.

Deux phases de développement sont actuellement à l'étude.

Phase 1 Dans un premier temps les logiciels gestionnaires de GEOMER vont être traduits en FTN77 MicroSoft version 3.31 afin de leur assurer une portabilité aussi grande que possible.

Un logiciel de création de base de données à la carte est en cours d'étude. Celui-ci va permettre à un éventuel acquéreur de ce type de base de données spatio-temporelle, de structurer sa propre base en fonction de la couverture géographique souhaitée, de ses besoins et du type de données à gérer.

L'implantation d'une telle base de données pourrait alors se faire sur des micro-systèmes du type compatible IBM PC-AT disposant d'une mémoire de masse de 30-40 MO, de 512 KO de mémoire vive et d'un système de sauvegarde du type steamer.

Phase 2 L'implantation sur un mini-système du type station Sun ou Apollo permettrait d'une part des accès en E/S par réseau type Ethernet et d'autre part des échanges de données avec les grandes bases internationales par connexion à un réseau international de télécommunications.

Certains types de données comme celles de géophysique marine, nécessiteraient alors la mise au point d'un système de protection équipé d'une mise à jour périodique, interdisant l'accès aux données frappées de confidentialité à un interrogateur extérieur non autorisé.