



DIENSTEN VAN DE EERSTE MINISTER  
PROGRAMMATIE VAN HET WETENSCHAPSBELEID  
Wetenschapsstraat 8  
1040 BRUSSEL  
BELGIE



SERVICES DU PREMIER MINISTRE  
PROGRAMMATION DE LA POLITIQUE SCIENTIFIQUE  
Rue de la Science, 8  
1040 BRUXELLES  
BELGIQUE

NATIONAAL ONDERZOEKS- EN  
ONTWIKKELINGSPROGRAMMA

LEEFMILIEU

WATER

PROGRAMME NATIONAL DE RECHERCHE  
ET DE DEVELOPPEMENT

ENVIRONNEMENT

EAU

**PROJEKT ZEE**

Eindverslag

Boekdeel 2

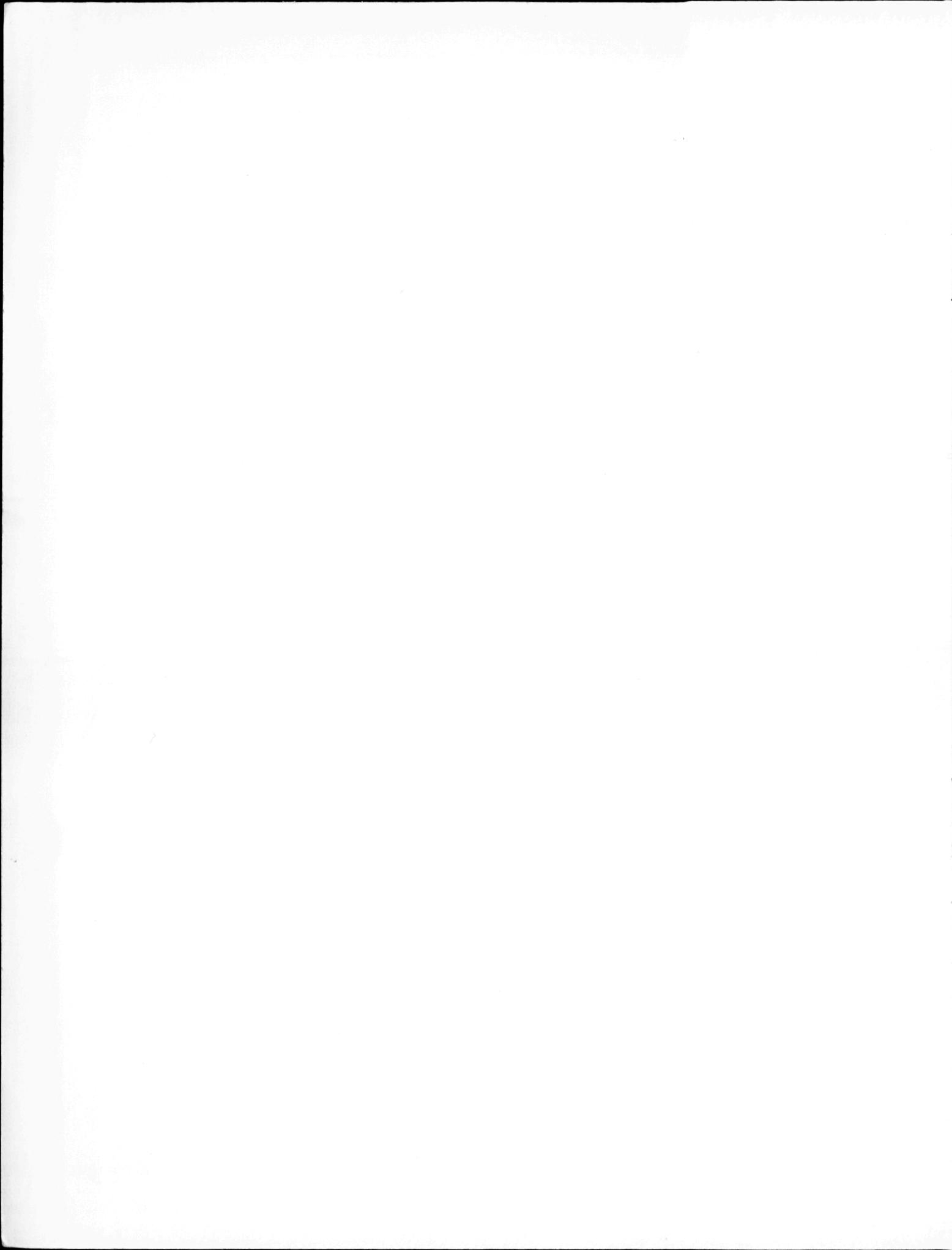
AANWINST EN COMPILATIE VAN GEGEVENS

**PROJET MER**

Rapport final

Volume 2

**ACQUISITION ET  
COMPILEATION DES DONNEES**

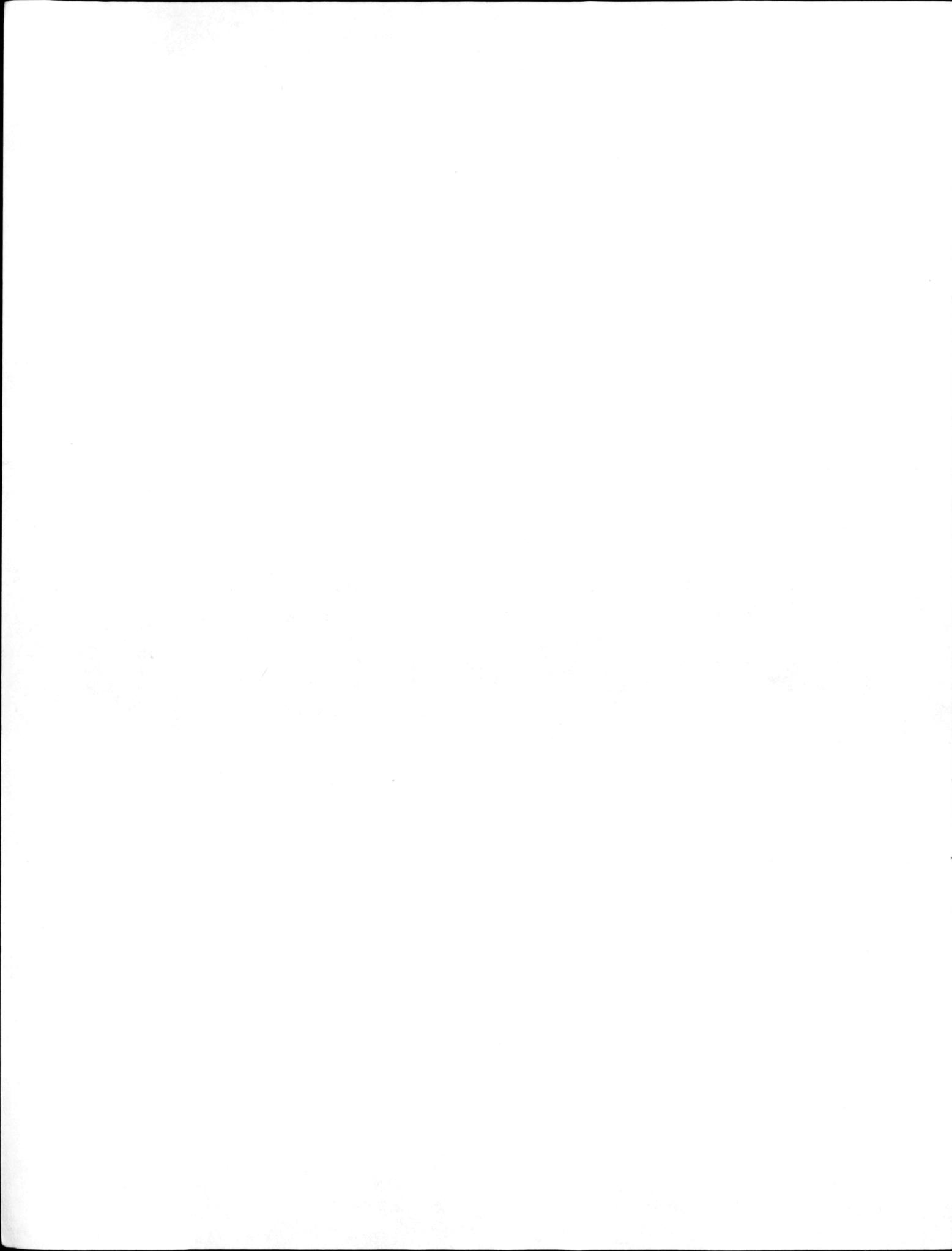


**PROJEKT ZEE**

Eindverslag

**PROJET MER**

Rapport final





DIENSTEN VAN DE EERSTE MINISTER  
PROGRAMMATIE VAN HET WETENSCHAPSBELEID  
Wetenschapsstraat 8  
1040 BRUSSEL  
BELGIE



SERVICES DU PREMIER MINISTRE  
PROGRAMMATION DE LA POLITIQUE SCIENTIFIQUE  
Rue de la Science, 8  
1040 BRUXELLES  
BELGIQUE

NATIONAAL ONDERZOEKS- EN  
ONTWIKKELINGSPROGRAMMA

LEEFMILIEU

WATER

PROGRAMME NATIONAL DE RECHERCHE  
ET DE DEVELOPPEMENT

ENVIRONNEMENT

EAU

**PROJEKT ZEE**

Eindverslag

**PROJET MER**

Rapport final

Boekdeel 2

Volume 2

**AANWINST EN COMPILATIE VAN GEGEVENS**

uitgevoerd door

**Jacques C.J. NIHOUL en G. PICHOT**

**ACQUISITION ET  
COMPILATION DES DONNEES**

édité par

**Jacques C.J. NIHOUL et G. PICHOT**

*Les modèles mathématiques requièrent, en phase préparatoire, l'acquisition de nombreuses séries temporelles de mesures et, en phase opérationnelle, la connaissance continue des conditions aux limites.*

*C'est pourquoi ont été préparés et mis en service des stations autonomes de courantométrie, deux bouées océanographiques et météorologiques ainsi qu'un piquet-laboratoire.*

*Les données issues de ces systèmes d'acquisition automatique, ainsi que celles qui ont été obtenues par échange entre institutions de recherche, forment la base de données. Le développement de programmes de traitement automatique et d'échange standardisé transforme cette base en banque de données; celle-ci constitue un précieux auxiliaire aux recherches sur l'écosystème, complément indispensable des modèles de simulation.*



## **Acquisition et compilation des données**

édité par

**Jacques C.J. NIHOUL et G. PICHOT**

TABLE DES MATIERES

---

ACQUISITION AUTOMATIQUE DE DONNEES EN MER

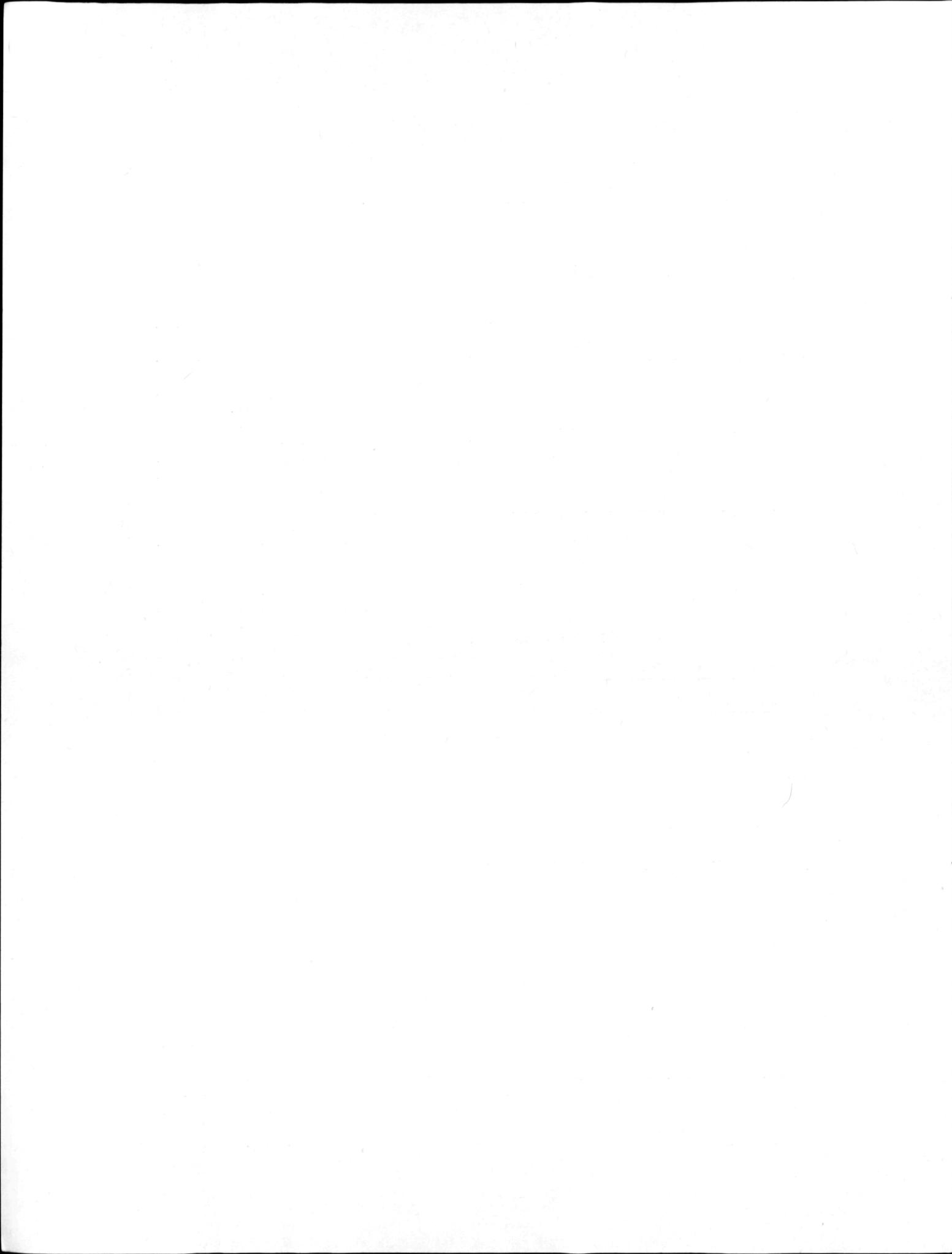
par Georges PICHOT, André POLLENTIER et Henri PICARD

<u>Introduction</u>	3
<u>Chapitre I - Courantomètres</u>	5
1.- Description des instruments utilisés	5
2.- Description du mouillage	6
3.- Inventaire et bilan des mouillages effectués	8
4.- Le décodage des données	13
5.- Le traitement des données	19
6.- Qualité des données	23
7.- Exemples	25
<u>Chapitre II - Les bouées océanographiques et météorologiques</u>	33
1.- Description de la bouée Sysna	34
2.- Le système d'acquisition de données (D.A.S.)	36
3.- Installation embarquée de transmission des données	45
4.- Enregistrement in situ sur bande magnétique	45
5.- Sensors océanographiques et météorologiques	48
6.- Alimentation en énergie	53
7.- La station de réception à terre	53
8.- Bilan des mouillages effectués	62
<u>Chapitre III - La station fixe</u>	69
<u>Références</u>	73

LA BANQUE DE DONNEES

par Yves ADAM, Huguette LAVAL, Pol CLOSSET, Jean-Pierre FOUGUENNE, Walter KEUTGEN

<u>Chapitre I - La gestion des données océanographiques</u>	77
<u>Chapitre II - Construction de la base de données</u>	79
1.- Etat présent de la base de données	79
2.- Données issues du Programme	80
3.- Données reçues par échange	91
<u>Chapitre III - Exploitation des données</u>	103
1.- Construction de la banque de données et du logiciel d'exploitation	103
2.- Conditions de portabilité du format de stockage et du logiciel	104
3.- Description générale du logiciel	106
4.- Stockage des données en format standard	107
5.- Etat actuel du logiciel d'exploitation	110
6.- Référence	114
<u>Chapitre IV - Echange des données</u>	115
1.- Introduction	115
2.- Description générale du GF2	116
<u>Conclusion</u>	125

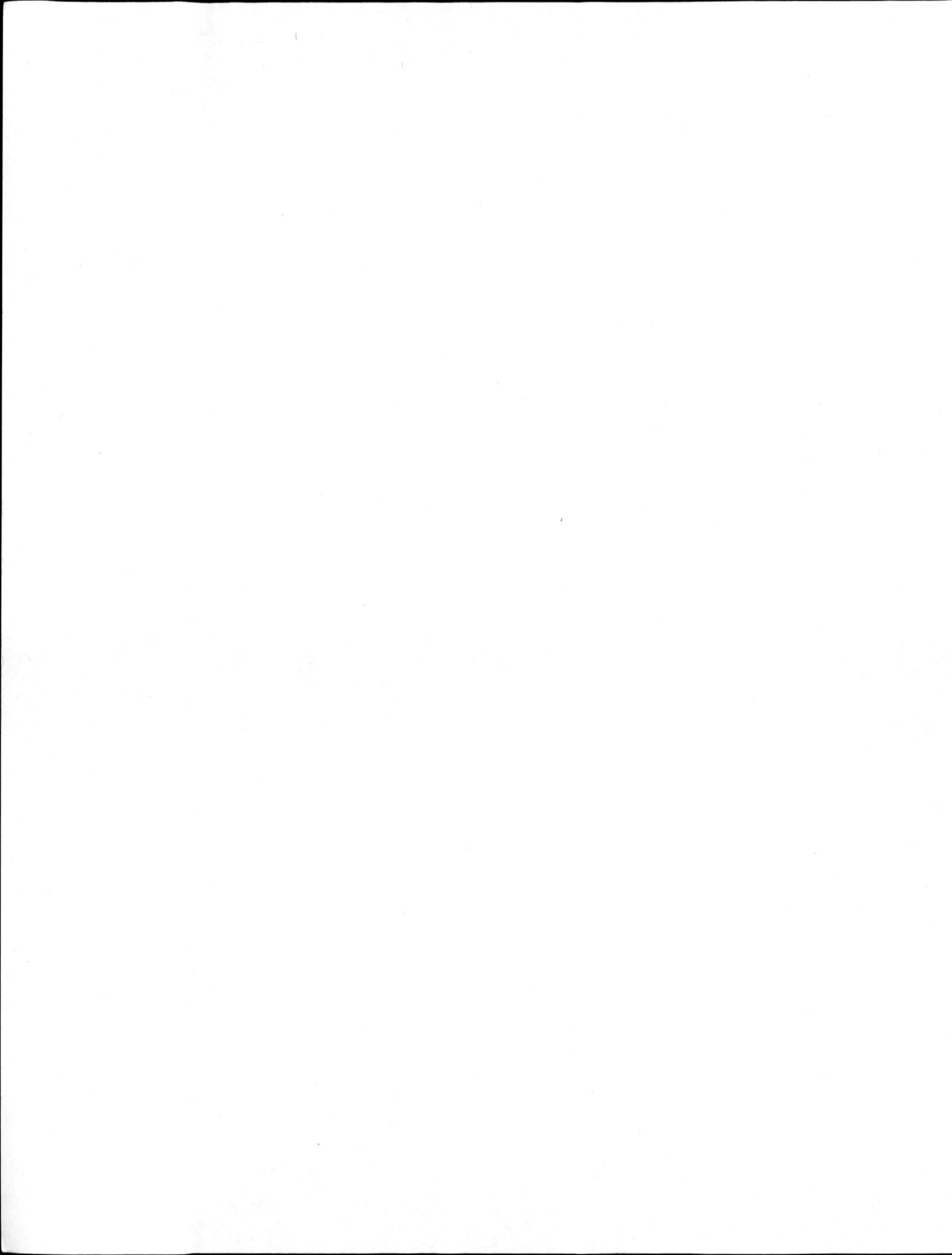


## ACQUISITION AUTOMATIQUE DE DONNEES EN MER

par

Georges PICHOT, André POLLENTIER et Henri PICARD

Les auteurs tiennent à exprimer à M. A. De Haen, au commandant J.P. Barbieux, au personnel de la Division Etudes et Recherches et à l'équipe de matelots miliciens universitaires, au commandant et à l'équipage de l'A 962 Mechelen, à M. A. Luyens, Nautisch Directeur van het Loodswezen et à son personnel, leurs vifs remerciements pour l'aide précieuse qu'ils leur ont apportée.



## Introduction

Pour établir le modèle mathématique qui décrit et prédit l'état d'un écosystème comme la mer du Nord, il est indispensable de pouvoir disposer de bonnes séries de mesures physiques, chimiques et biologiques.

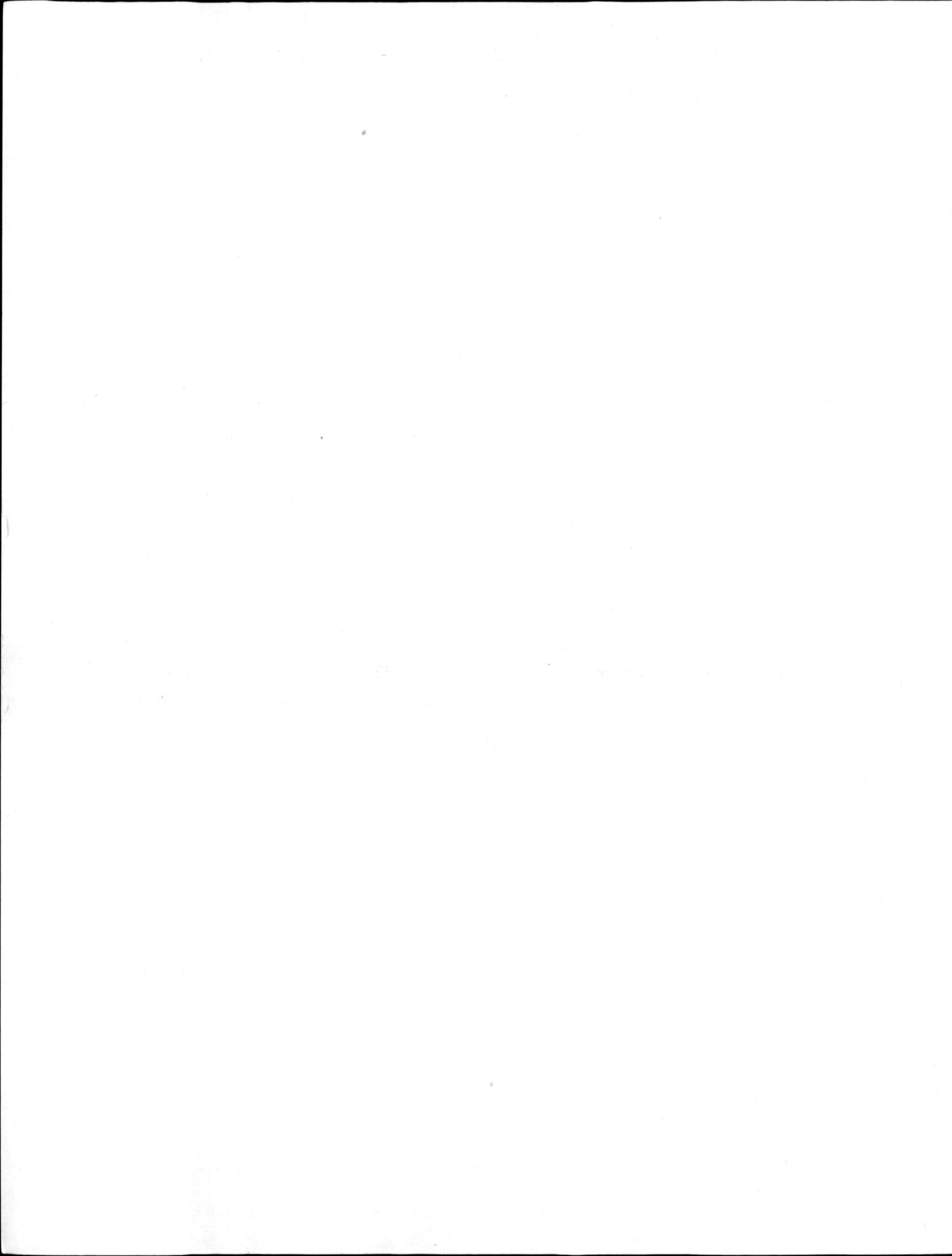
Dans la phase préparatoire du modèle, elles permettent d'ajuster les valeurs numériques de certains paramètres d'interactions et de vérifier la précision des premiers résultats. Dans la phase opérationnelle, elles fournissent les conditions aux limites requises pour la résolution des équations différentielles d'évolution.

Les moyens d'acquisition automatique de données mis en oeuvre dans le cadre du programme peuvent être divisés en trois groupes :

1. Les stations auto-enregistreuses de courantométrie : 22 mouillages d'instruments, effectués à partir de septembre 1973 ont procuré 317 jours de bonnes données.

2. Deux bouées océanographiques et météorologiques : mis en service en avril 1974, ces *prototypes* qui interrogent les différents senseurs et télétransmettent leurs informations à la station de réception et de traitement située à Ostende ont fourni quelque 80 jours de bonnes données.

3. Une station fixe : elle réalisera le même travail que les bouées mais, en outre, permettra la mise au point — sans limitation d'espace ou d'énergie — d'échantilleurs programmés et de senseurs chimiques et biologiques. Prête en juin 1975, elle n'a pu encore être mise en place (à 51° 21' 25" N , 3° 03' E) à cause de conditions météorologiques défavorables. Elle devrait l'être dans les premiers mois de 1976.



## Chapitre I

### Courantomètres

#### 1.- Description des instruments utilisés

La courantométrie est réalisée à l'aide de trois types d'instruments.

##### 1.1.- Le courantomètre AMF-VACM modèle 610

Cet instrument, dont nous possédons six exemplaires, est sans conteste le plus évolué de tous. Il possède un rotor Savonius et une vanne de direction couplée magnétiquement à un disque encodé optiquement et à un compas. Il effectue, sur un intervalle de temps fixé, les sommes des coordonnées cartésiennes des vecteurs unitaires de la vitesse du courant (le nombre des sommes étant proportionnel au module de cette vitesse) qu'il enregistre sur cassette digitale.

Il peut encore prendre la température moyenne de l'eau et la position instantanée du compas et de la vanne au moment de l'enregistrement.

L'AMF-VACM est robuste et peut avoir une autonomie d'un an. Malheureusement, il n'existe pas d'unité de contrôle acoustique des données. Il n'est donc pas possible d'effectuer un test de l'enregistrement extérieur à l'appareil et ainsi de vérifier son bon fonctionnement, une fois qu'il est ancré.

### 1.2.- Le courantomètre NBA modèle DNC-2A

Cet instrument dont nous possédons douze exemplaires est le plus utilisé dans nos campagnes de mesures.

Un rotor, entraîné par le mouvement de l'eau, envoie des impulsions qui sont intégrées pendant une période fixée. La direction du courant, dans laquelle l'instrument est maintenu par un assemblage d'ailerons, est donnée par la sommation de 128 lectures du compas optique. La température moyenne de l'eau peut également être mesurée. L'enregistrement se fait sur une bande digitale.

L'indication "read after write" et l'unité de contrôle acoustique existent et permettent, en toute circonstance, la vérification du bon fonctionnement de l'appareil et de la validité de l'enregistrement.

### 1.3.- Le courantomètre Plessey modèle M021

Trois courantomètres Plessey ont été employés lors des premiers mouillages. La façon de mesurer la vitesse du courant est pratiquement identique à celle du N.B.A. La direction du courant, dans laquelle l'instrument est maintenu par une queue, est donnée par trois lectures instantanées du compas. L'enregistrement se fait sur bande magnétique.

Il convient enfin d'ajouter que le système d'ancre du Plessey n'est pas aisément à manipuler et que les délais de livraison des instruments et des pièces de rechange ne sont pas souvent respectés.

## 2.- Description du mouillage

La figure 1 donne le schéma d'un mouillage de courantomètres. Il comprend :

- une bouée profilée, sous la surface, d'une flottabilité d'environ 200 kg qui maintient le courantomètre aussi vertical que possible,
- les courantomètres avec leur bloc d'ancre,
- la bouée de marquage (Selco type 7 spar buoy) [qui, suivant la convention de l'IMCO, est peinte en larges bandes jaunes et rouges] et son système d'ancre,

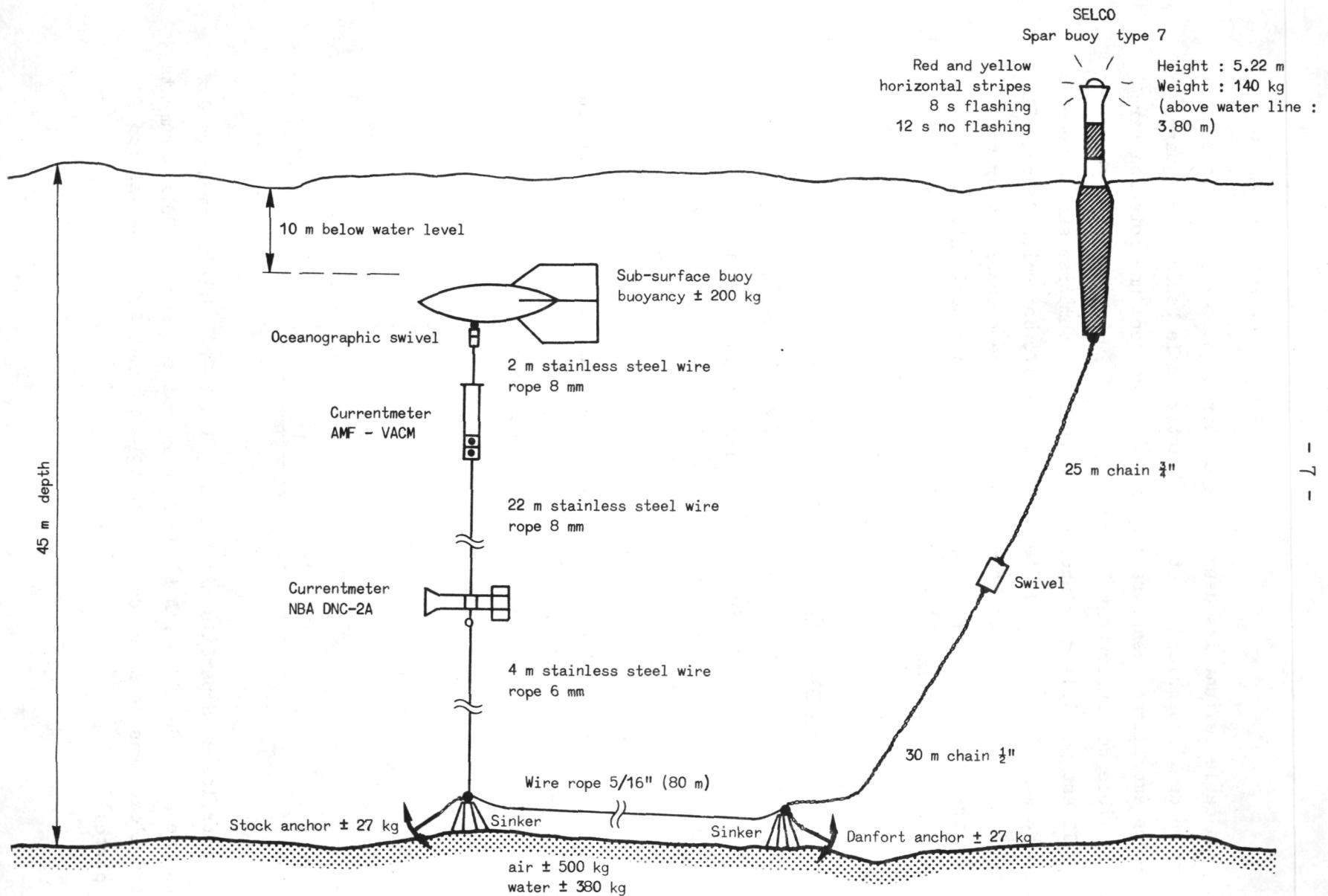


fig. 1.  
Schéma d'un ancrage-type

- un cable reliant les deux blocs d'ancrage et permettant la récupération de l'ensemble du système, à partir de la bouée de marquage,
- des émerillons océanographiques qui permettent une rotation verticale libre des instruments.

Il est possible de placer plusieurs courantomètres sur la même ligne de mouillage. Si l'un d'eux est un N.B.A., son unité de contrôle acoustique permet, dans de bonnes conditions, de décoder ses propres enregistrements mais aussi, dans des conditions d'accident, de mettre en évidence l'existence même de la ligne.

### 3.- Inventaire et bilan des mouillages effectués

Le tableau 1 donne l'inventaire des différents mouillages effectués dont la distribution géographique est donnée à la figure 2.

Un examen rapide du tableau 1 montre clairement que la durée de vie d'un mouillage dépend moins de l'instrument de mesure lui-même et de l'énergie dont il peut disposer que du risque d'un accident grave entraînant la perte du courantomètre et des données qu'il contient ou la crevaison du flotteur sous la surface.

Le bilan des pertes d'appareil s'établit comme suit :

Type	Mouillés	Récupérés
AMF-VACM	5	5
Plessey-M021	4	2
NBA-DNC2A	16	13

Sur les 25 appareils mis à l'eau, 5 ont été définitivement perdus : d'où une perte de 20 %. A titre de comparaison, sur les 90 instruments mis à l'eau dans le cadre de JONSDAP 73, 70 ont été récupérés (22 % de pertes).

fig. 2.

Distribution géographique des mouillages de courantomètres effectués de septembre 1973 à décembre 1975. Les numéros des mouillages font référence au tableau 1.

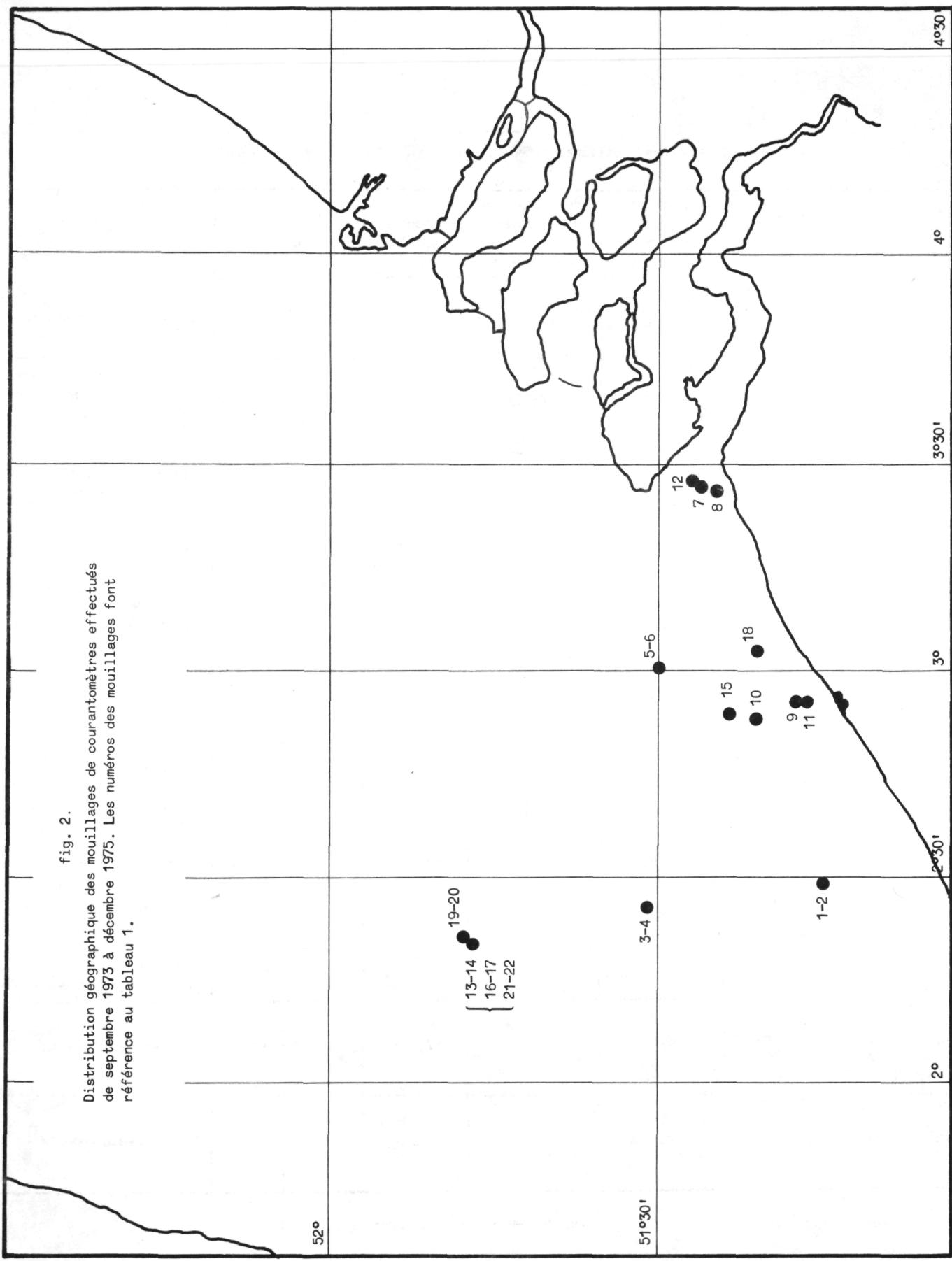


Tableau 1

Inventaire des mouillages effectués (septembre 1973 - décembre 1975)

Numéro du mouillage	Position	Données récupérées	Instrument et numéro de série	Profondeur de l'instrument à partir du fond de la mer	Profondeur totale	Intervalle d'enregistrement	Remarque
*	52°41' N 04°48' E	Néant	Plessey 085	12,5 m			Ensemble du mouillage (comprisant le Plessey 085 et un CM du K.N.M.I.) perdu
*	52°47' N 04°48' E	07-09-73 18-10-73	Plessey 234	13 m			Batteries vides à la récupération. Données reçues du K.N.M.I. Bandes perforées, code PL.
*	52°44' N 03°50' E	28-09-73 18-10-73	AMF-VACM	22,5 m		15 min	CM sur le fond lors de la récupération. Deux autres CM du même mouillage (appartenant au K.N.M.I.) perdus
1	51°14'55" N 02°28'58" E	07-09-73 13-10-73	Plessey 700	10,5 m	25 m	5 min	Quelques données non valides, défaut dans le système d'encodage. Rupture du mouillage et CM en dérive le 13-10-73
	same rig	"	NBA 212	18,5 m		15 min	3800 enregistrements
3	51°31'00" N 02°25'30" E	Néant	NBA 219	29 m	35 m	15 min	Mouillage et CM perdus. Bouée de surface récupérée. Un chasseur de mines, appelé à l'aide, retrouve le mouillage et découvre une mine à proximité. Fait sauter la mine et par conséquent les CM.
	same rig	Néant	NBA 223	11,5 m		15 min	
5	51°29'58" N 03°00'02" E	Néant	Plessey 699	-	-	5 min	CM perdu : bras de suspension cassé. Le reste du mouillage récupéré
	same rig	22-11-73 30-11-73	NBA 222	-		10 min	Bandes magnétiques déchirées après 1200 enregistrements
7	51°26'45" N 03°26'50" E	19-03-74 02-04-74	NBA 221	3 m	7,8 m	10 min	2200 enregistrements
	51°25'10" N 03°26'07" E	21-06-74 28-06-74	NBA 222	3 m	15 m	6 min	2500 enregistrements

\* Courantomètres mouillés par le K.N.M.I. (De Bilt, Pays-Bas) dans le cadre de JONSDAP 73.

Tableau 1

(suite)

Numéro du mouillage	Position	Données récupérées	Instrument et numéro de série	Profondeur de l'instrument à partir du fond de la mer	Profondeur totale	Intervalle d'enregistrement	Remarque
9	51°17'12" N 02°55'20" E	16-07-74 30-07-74	NBA 222	3 m	-	3 min	7000 enregistrements
10	51°21'15" N 02°52'40" E	17-07-74 31-07-74	VACM 188	4 m	-	15 min	1200 enregistrements
11	51°16'50" N 02°55'10" E	Néant	NBA 230	3 m	7,8 m	8 min	CM perdu, bras de suspension cassé. Mouillage récupéré, à l'exception du flotteur subsurface.
12	51°26'40" N 02°26'40" E	14-04-75 14-05-75	NBA 227	3 m	8 m	10 min	4400 enregistrements
13	51°47'00" N 02°20'00" E	Néant	NBA 228	3 m	46 m	10 min	Subsurface coulé. NBA : 6 jours de données utilisables.
14	same rig	Néant	VACM 285	25 m		7,5 min	VACM : avancement de bande bloqué, pas d'enregistrement
15	51°23'38" N 02°53'20" E	Néant	NBA 226	5 m	20 m	10 min	Bandé magnétique cassée après 24 h de fonctionnement
16	51°47'00" N 02°20'00" E	Néant	NBA 229	3 m	46 m	10 min	Subsurface coulé. NBA : batterie vide à la récupération. Aucune donnée enregistrée sur la bande.
17	same rig	28-07-75 (22-09-75)	VACM 284	25 m		7,5 min	VACM : données non encore reçues du W.H.O.I.
18	51°21'22" N 03°02'34" E	02-09-75 16-09-75	NBA 226	2 m	8 m	5 min	4000 enregistrements. Subsurface perdu lors de la récupération
19	51°47'50" N 02°20'50" E	22-09-75 24-10-75	NBA 212	5 m	46 m	10 min	Subsurface coulé. N.B.A.: 8 jours de données valides.
20	same rig	22-09-75 24-10-75	VACM 286	27 m		7,5 min	VACM : données non encore reçues du W.H.O.I.

Tableau 1

(fin)

Numéro du mouillage	Position	Données récupérées	Instrument et numéro de série	Profondeur de l'instrument à partir du fond de la mer	Profondeur totale	Intervalle d'enregistrement	Remarque
21	51°47'00" N 02°20'00" E	11-12-75	NBA 224	25 m	46 m	10 min	Nouveau flotteur subsurface. Bouée de marquage perdue. CM non encore récupérés.
22	same rig	11-12-75	NBA 222	5 m		10 min	

Si on exprime ce même bilan en nombre de jours, sur les 610 jours d'opération planifiés, 193 ont été perdus à cause de la disparition des lignes de mouillage et/ou des courantomètres.

Ceci est dû principalement aux activités de pêches à proximité des ancrages. A plusieurs reprises, lors des récupérations, des morceaux de filets de pêche ont été retrouvés, emmêlés aux mouillages accidentés. Une autre cause est aussi la météorologie car, pour fournir des données caractéristiques aussi bien des tempêtes que de temps calme, le programme des opérations a toujours été établi indépendamment des saisons et des périodes d'éventuel mauvais temps. Le problème posé par ces lourdes pertes fait l'objet d'un examen attentif au sein du groupe de travail "Recovery of Equipements" de l'Action Cost 43 des Communautés Européennes.

Pour les mouillages restés intacts, les pertes dues aux défectuosités des instruments eux-mêmes s'établissent comme suit :

Type d'instrument	Nombre de jours de mouillage intact	Défectuosité des instruments	Pourcentage des pertes
AMF-VACM	125	42	33 %
Plessey-M021	84	8	9 %
NBA - DNC2A	208	79	38 %
TOTAL	417	129	31 %

Ces pourcentages sont du même ordre de grandeur que ceux estimés par d'autres instituts océanographiques travaillant en mer du Nord [Hill et Ramster (1974)]. De l'expérience acquise au fil des années, il semble que la façon optimale de réaliser la courantométrie consiste :

- à planifier des mouillages pour une durée de trois semaines,
  - à effectuer, après une dizaine de jours et à l'aide de l'unité de contrôle acoustique N.B.A., une vérification de la présence même de la ligne et du bon fonctionnement des N.B.A.,
  - à prévoir un intervalle d'interrogation de telle sorte que les instruments aient une autonomie de 6 semaines et qu'ils puissent rester opérationnels même si de mauvaises conditions météorologiques retardent la procédure de récupération.

#### 4.- Le décodage des données

Une fois les courantomètres récupérés, il faut d'abord effectuer le décodage des données.

4.1.- Données fournies par le courantomètre N.B.A.-DNC2A

Les données fournies par un courantomètre N.B.A.-DNC2A se trouvent sur une bande magnétique de format  $\frac{1}{4}$  pouce, 75 BPI. Un enregistrement consiste en 38 bits répartis en 4 mots de la façon suivante :

MOT 1 (8 bits) : consistant en deux demi-mots de 4 bits placés au début et à la fin de l'enregistrement. Le numéro de série (SER.#) est fourni par une résistance fixe et l'élément sensible à la température est un thermistor.

MOT 2 (10 bits) : les impulsions du rotor sont sommées pendant un intervalle de temps fixé.

MOT 3 (8 bits) : l'enregistrement de la direction est la moyenne de 128 lectures d'un disque encodé optiquement suivant le code Gray. Originellement, 7 bits étaient disponibles pour la direction. A notre demande, le constructeur a ajouté un 8ème bit (qui correspondait avant au 12ème bit de numérotation de l'enregistrement). Ce qui porte la résolution de l'instrument à  $\pm 1,4^\circ$ .

MOT 4 (11 bits) : partant de 0 et jusqu'à un maximum de 2047, la numérotation de l'enregistrement se fait par incrémentation de 1 à chaque interrogation.

Un bit de parité est enregistré à la fin de chaque enregistrement pour engendrer une parité positive. La traduction de cette bande magnétique se fait *off line* au moyen d'un décodeur couplé à un perforateur de ruban (fig. 3). Ce travail est actuellement réalisé par la N.B.A. *Data Handling Division* (U.K.) à qui nous ne demandons, parmi les nombreux services de traitement des données qu'elle propose, que la traduction sur bande perforée en ASCII, avec contrôle de la parité et de la présence des 40 impulsions de référence. Cette bande se présente de la façon suivante :

1	0001	$\wedge$ 123 020	$\wedge$ 357	$\wedge$ 0624	Format 9A2
VD	READ	SER # TEMP	DIR	VEL	

avec

- VD : validité de la donnée. Si  $VD = 1$ , la donnée est inscrite dans un fichier dont l'entête est "NBAF\*\*" (où \*\* est le numéro de

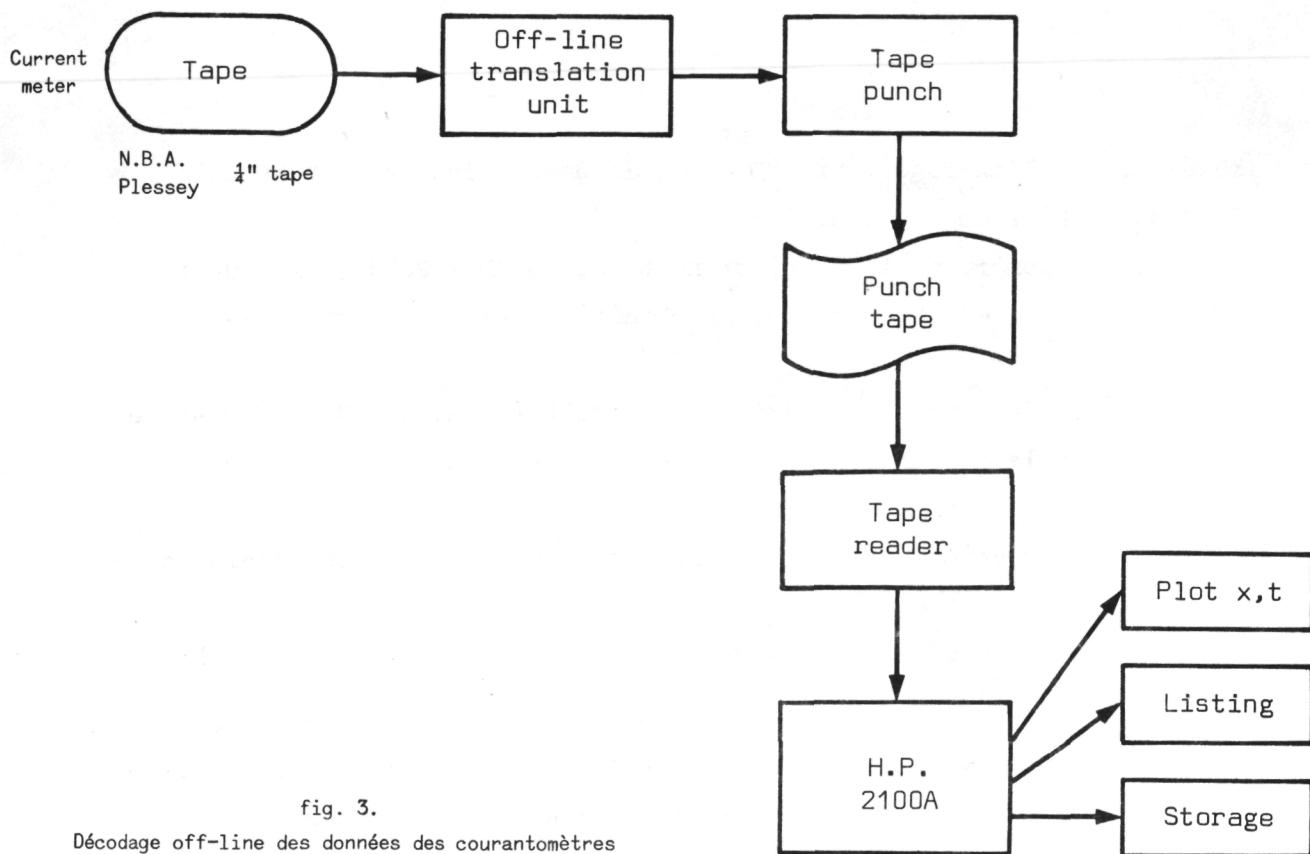


fig. 3.

Décodage off-line des données des courantomètres

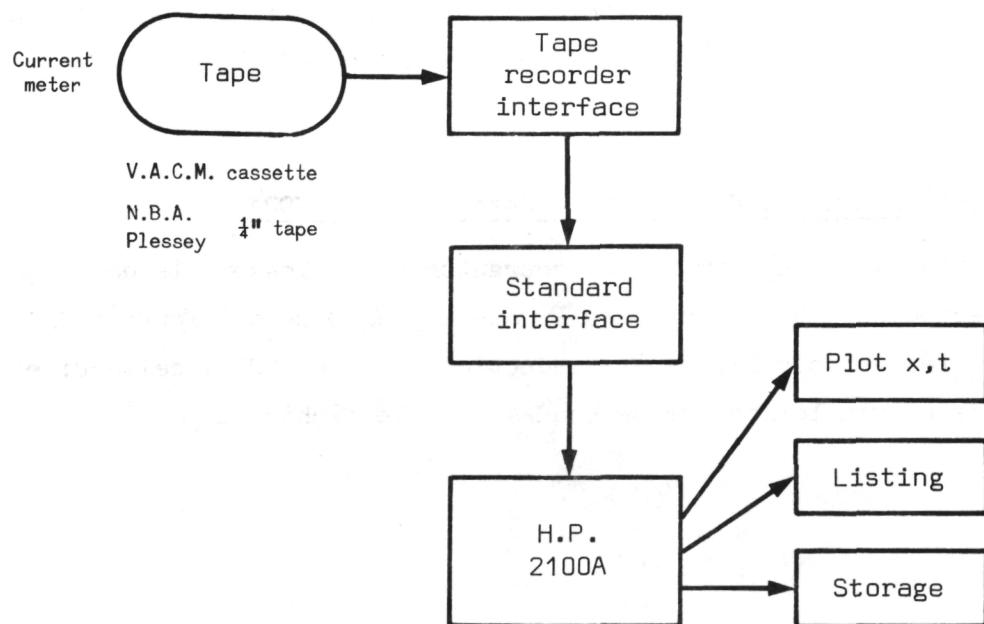


fig. 4.

Décodage on-line des données des courantomètres

référence de l'ancrage). Si  $VD = 0$ , la donnée fautive est stockée dans un fichier de réserve.

- READ : numéro de l'enregistrement. Si  $READ > 2048$ , il faut retrancher 2048 et dans ce cas, la précision sur la direction est  $\pm 1,4^\circ C$ .

- SER<sup>#</sup> TEMP : SER, numéro de série fixé, s'alterne avec TEMP, température moyenne dont la valeur réelle s'obtient via une table d'étalonnage fournie par le constructeur.

- DIR : direction moyenne exprimée en degrés et avec une précision de  $\pm 1,4^\circ C$ .

- VEL : nombre d'impulsions du rotor par intervalle de temps. La vitesse est donnée par

$$V = \frac{VEL \times SCAF}{60 \times RECI} \times 32,1 + 0,978 \text{ en cm/s où SCAF est le facteur d'échelle et RECI , l'intervalle d'enregistrement.}$$

Les enregistrements sur bandes magnétiques peuvent également être introduits dans le calculateur via une interface appropriée (fig. 4). Pour cela, une installation hardware a été développée par l'unité d'Ostende (fig. 5). D'autre part, le logiciel (*driver DVR 77 et programme NBA 77*) est actuellement en développement si bien que, sous peu, la traduction des données des courantomètres NBA-DNC2A pourra être réalisée par le calculateur à Ostende.

#### 4.2.- Les données des courantomètres Plessey-M021

Vu l'emploi limité des courantomètres Plessey, le décodage est effectué par le Plessey *Data Processing Service* qui fournit une bande perforée en code PL. Après traduction en code ASCII, celle-ci est stockée, via le lecteur de bandes, sur le fichier d'un disque.

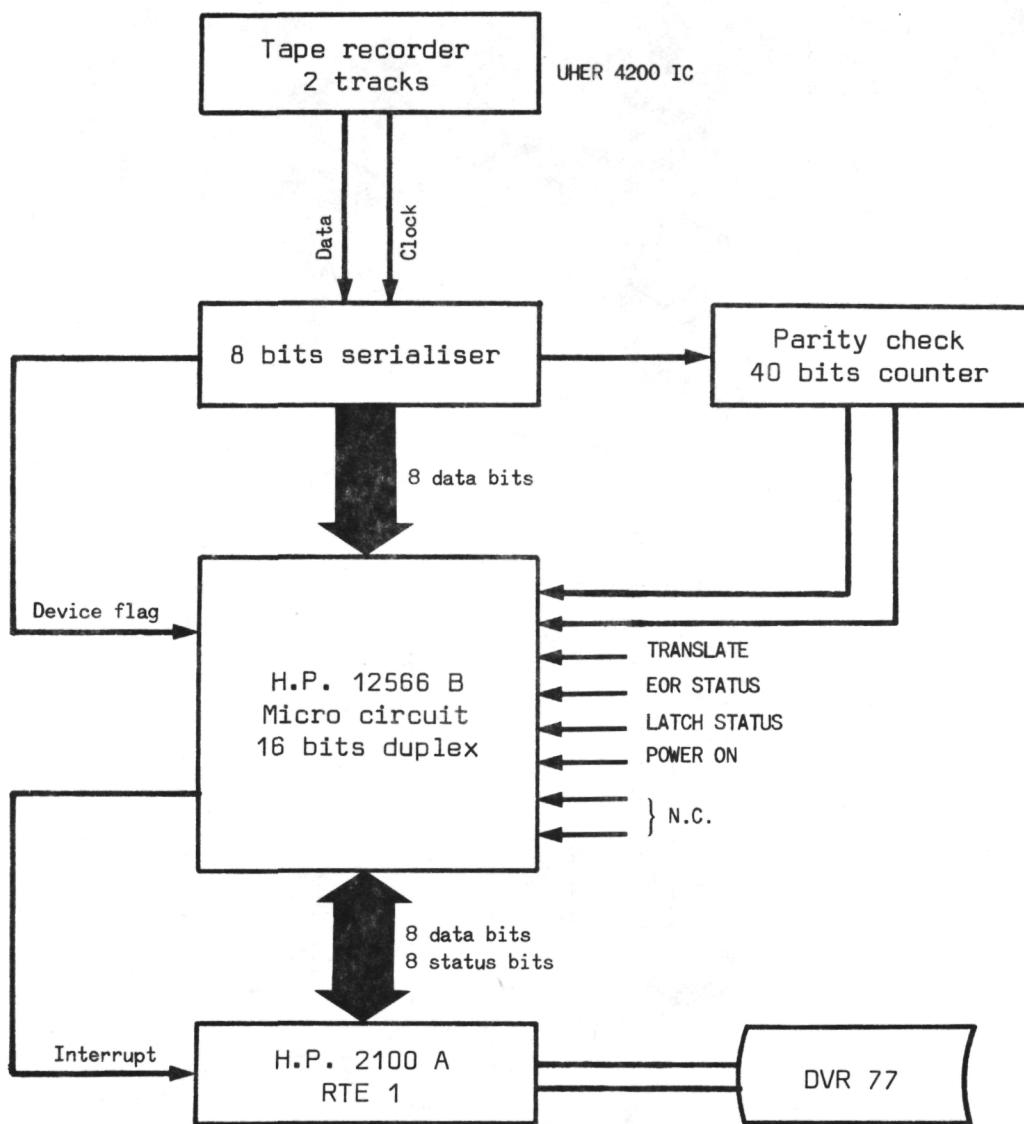


fig. 5.  
On line decoder for N.B.A. - DNC2A magnetic tape

#### 4.3.- Les données des courantomètres VACM-AMF

La bande magnétique du VACM-AMF est une cassette de style Philips (300 ft) 1600 fci . Le code est de 800 BPI , avec encodage de phase (ECMA standard) sur quatre pistes. Le format est de 3200 bits par inch à un taux d'enregistrement de 100 Hz par piste. La figure 6 donne la

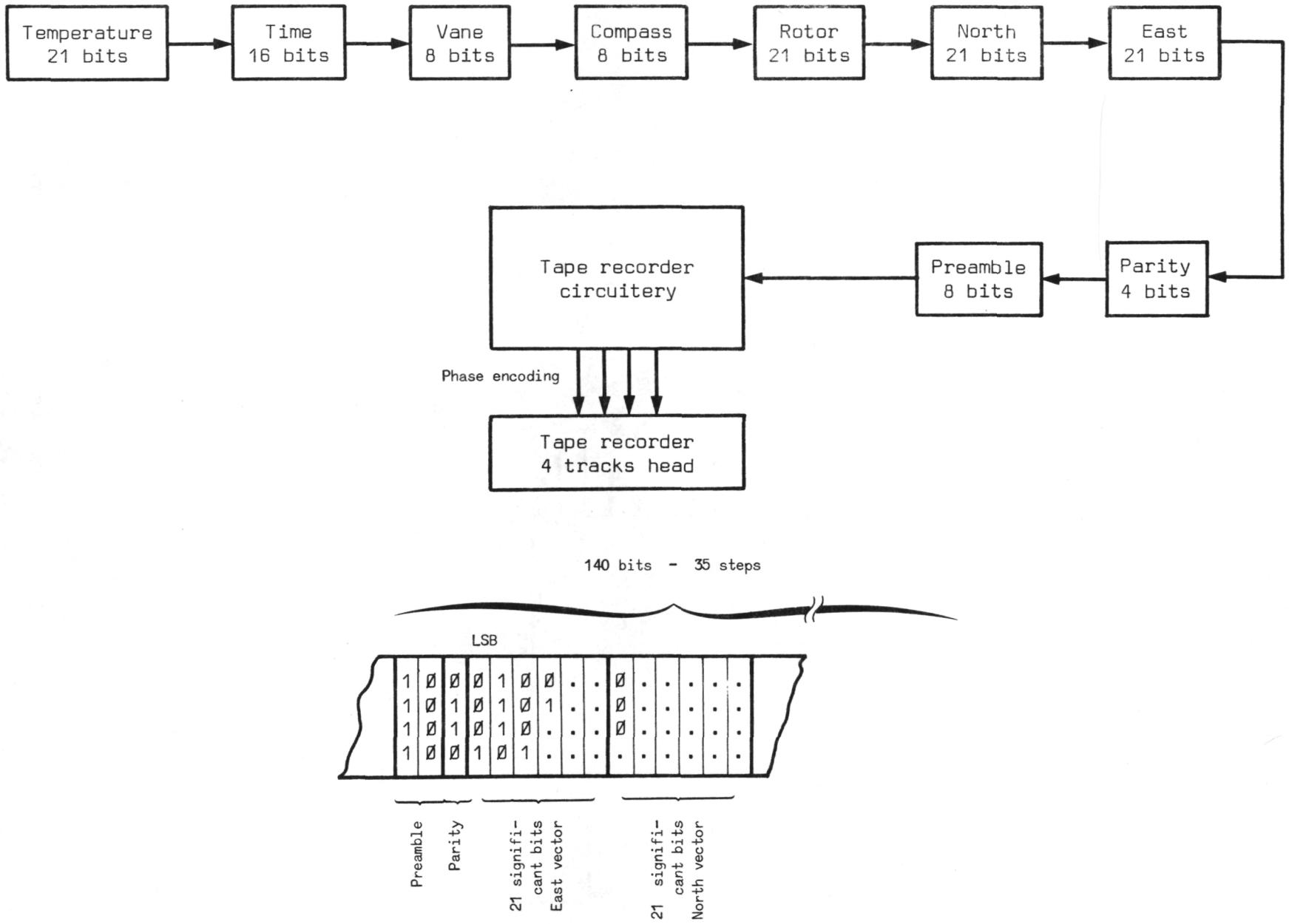


fig. 6.

structure de l'enregistrement sur la bande magnétique. Actuellement, la traduction de ces cassettes est exclusivement réalisée par le Woods Hole Oceanographic Institution. Celles-ci sont transférées sur bande de 9 pistes et dans le format Maltais à l'aide d'un lecteur de cassettes développé par le W.H.O.I. et d'un calculateur HP 2100 . Pendant ce transfert, différents tests sont réalisés (erreurs de parité, enregistrements longs et courts, etc.).

## 5.- Le traitement des données

### 5.1.- Données du N.B.A.-DNC2A (fig. 7)

- CRFNW : la bande perforée fournie par la N.B.A *Data Handling Division* est stockée sur un fichier du disque NBA\*\*. Les données fautives sont placées dans un fichier de réserve.

- DUR77 (en développement) : les données sont transférées directement de la bande enregistreuse à un fichier du disque, après classement des bits et séparation des données fautives.

- CRNBA : les enregistrements-tests réalisés au début de la mise en service de l'instrument sont éliminés; les données utiles converties en valeurs réelles sont stockées séparément en température (NBAT\*\*), direction (NBAD\*\*), vitesse (NBAV\*\*) et coordonnées cartésiennes du courant (MOUS\*\*).

- GMNBA : calcul des moyennes horaires et journalières (24 h 50').

- LSNBA : effectue la liste de toutes les données avec leurs moyennes horaires et journalières.

- TBVAT : contrôle des valeurs temporelles des entités établies par CRNBA.

- PLOMO : sortie graphique cumulée du contenu de MOUS\*\* (*progressive vector diagram*).

- PLOMT : sortie graphique, jour après jour du contenu de MOUS\*\* (*ellipse de marées*).

- GRAFC : sortie graphique des différentes grandeurs en fonction du temps avec, comme départ, l'heure et le jour de l'ancreage.

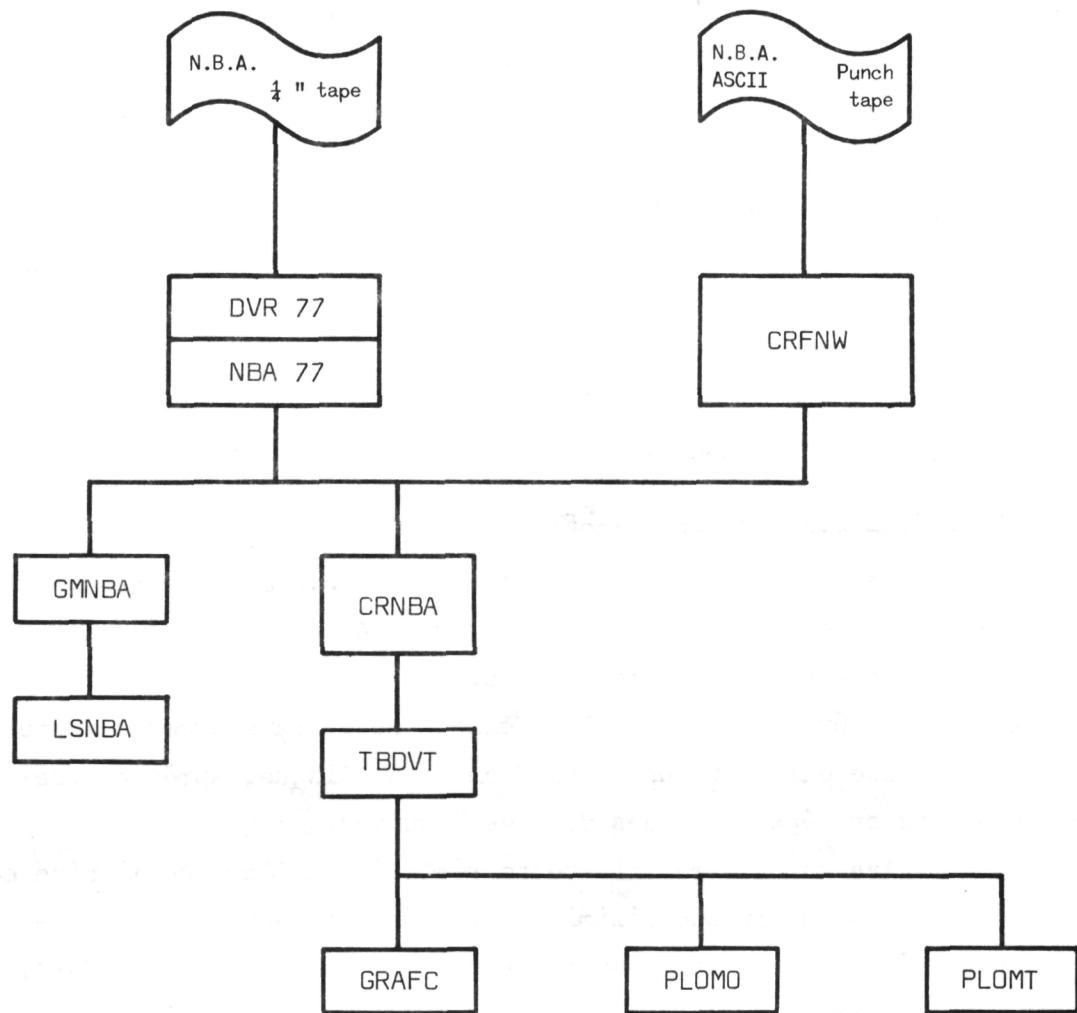


fig. 7.

Traitemen t des données des courantomètres N.B.A. - DNC-2A

La mise au point de ce système de programmes (qui, d'ailleurs, pourrait être adapté au traitement des données VACM-BCD FORTRAN en provenance du W.H.O.I.) implique l'emploi des jours juliens ainsi que la détection et le remplacement des valeurs erronées. Il semble urgent de compléter ce système par un programme d'histogrammes (très utile pour détecter certaines erreurs) et par un programme de stockage.

### 5.2.- Données du Plessey

Après une indispensable traduction en code ASCII, la bande perforée fournie par le *Plessey Data Processing Service* peut être stockée sur un fichier du disque, via un programme pratiquement identique à CRFNW.

### 5.3.- Données du VACM-AMF traitées par le W.H.O.I.

Le W.H.O.I., qui a conçu et développé le VACM et qui assure le décodage des cassettes, propose encore le traitement des données par le système de programmes suivant (fig. 8) :

- CARP : transfert de données de la cassette vers une bande 9 pistes, avec détection de certaines fautes (parité, enregistrements longs et courts, ...),
- XTAL : contrôle du fonctionnement de l'horloge-cristal et stockage des valeurs correspondant à des données fautives dans un fichier d'attente,
- CASDEC : transformation des données brutes en valeurs réelles (m/s, °C, etc.).

A partir d'ici, les données peuvent être transformées du format Maltais en BCD pour leur traitement ultérieur en FORTRAN.

- HISTO : production d'histogrammes de toutes les grandeurs pour le contrôle de la qualité des données et le choix des facteurs d'échelle pour les programmes suivants.

- DISPL0 : production des diagrammes analogiques des variables en fonction du temps, éventuellement après TAPDIS.

- THISTO : production d'histogrammes des variables enregistrées pendant des périodes de deux ou quatre heures successives, ainsi que des graphiques, sur imprimante, des séries temporelles des variables.

- SCRUB : élimination des mauvaises mesures détectées par THISTO.

- TAPDIS : remplacement des valeurs erronées par interpolation linéaire.

- TIMSAN : calcul des spectres de la vitesse et de la température, dessin de ces spectres à échelle log-log ou log-linéaire.

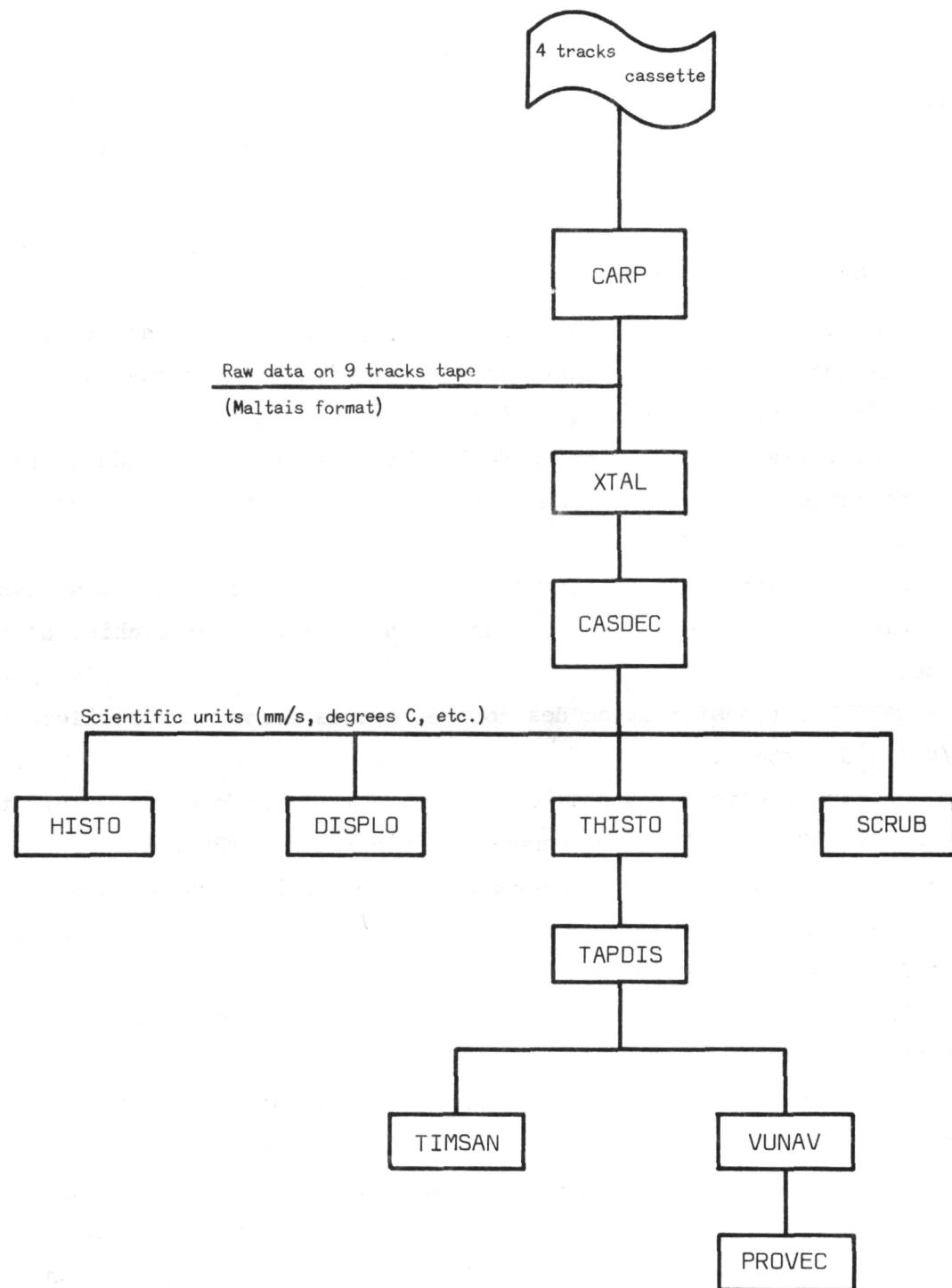


fig. 8.  
Traitement des données des courantomètres VACM-AMF effectué par le W.H.O.I.

- VUNAV : calcul des moyennes vectorielles horaires destinées au stockage et aux autres représentations graphiques.
- PROVEC : diagramme progressif vectoriel des moyennes vectorielles horaires.

L'exécution complète de cet ensemble de programmes de traitement des données des courantomètres VACM est assez coûteuse et peut prendre un certain temps car nos données ne sont traitées par les spécialistes du W.H.O.I. que lorsqu'ils ont achevé de traiter les leurs.

Comme plusieurs programmes assez semblables existent déjà à l'unité d'Ostende, ne sera plus dorénavant exécuté au W.H.O.I. que le décodage vers une bande 9 pistes en valeurs réelles, c'est-à-dire les programmes CARP, XTAL et CASDEC. Ces données décodées pourront ensuite être branchées sur les programmes développés pour le traitement des données NBA. D'autre part, l'étude de la réalisation de notre propre interface pour la traduction des cassettes a été envisagée.

#### 6.- Qualité des données

Pour pouvoir fournir des données dont la qualité est garantie, il faut :

- éliminer toutes les erreurs en effectuant, par exemple, les tests de parité, le contrôle des longueurs d'enregistrements et la détection d'autres erreurs évidentes dues à l'instrument lui-même ou à l'appareil de décodage.

- vérifier les données de vitesse, de température et de direction compte tenu de ce qui est déjà connu sur l'endroit où les mesures sont effectuées et sur les résultats escomptés.

Cela suppose un choix de certaines bornes supérieures et inférieures ainsi que d'un taux de variation des mesures.

Ces vérifications faites, il faut encore assurer une calibration précise des senseurs de vitesse, de direction et de température des instruments utilisés. Comme les études entreprises dans le cadre du Projet Mer attachent beaucoup d'importance aux circulations résiduelles

dont l'ordre de grandeur correspond aux erreurs instrumentales, il faut calibrer, avec un soin tout particulier, le compas magnétique (éventuellement la vanne directionnelle des VACM) et le rotor de chaque instrument.

Pour chaque compas fourni, le constructeur fournit une courbe de calibration qui ne semble pas très reproductible. C'est pourquoi il s'est avéré indispensable de calibrer la direction de chaque instrument, avant et après mouillage. Cela est réalisé à Ostende sur table tournante magnétique et avec rotor en fonctionnement pour annuler tout couplage magnétique [Picard (1974)] (fig. 9). D'autre part, nous participons aux programmes d'intercalibration de compas de courantomètres menés à un niveau européen par le *Fisheries Laboratory de Lowestoft* (U.K.).

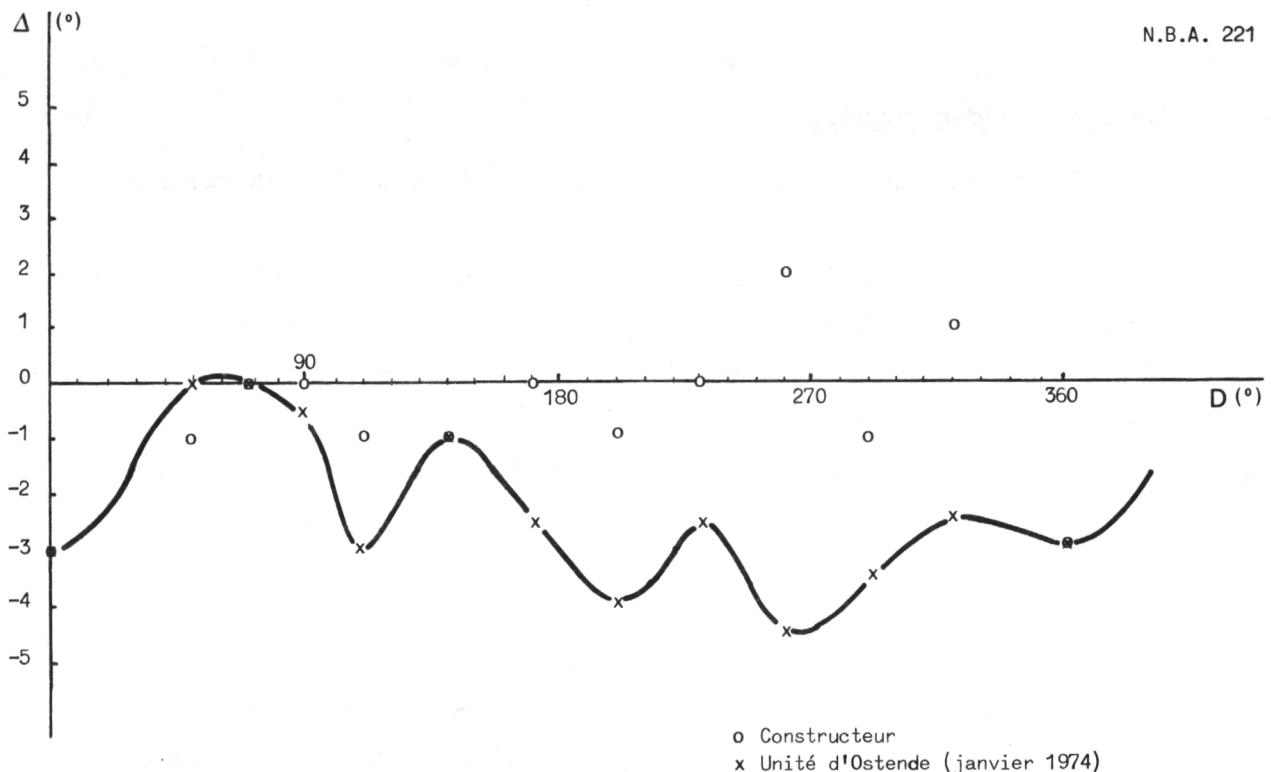


fig. 9.  
Exemple de calibration

Pour le rotor, il conviendrait de calibrer, avec un soin identique, sa constante et sa vitesse de démarrage. Il serait souhaitable que ce travail soit réalisé avec l'aide du Laboratoire de Recherches Hydrauliques de Borgerhout qui possède l'infrastructure nécessaire.

#### 7.- Exemples

A titre d'exemple, les figures 10 et 11 montrent quelques graphiques caractéristiques produits en routine par l'unité d'Ostende et mis à la disposition de l'unité hydrodynamique de notre Projet Mer. La figure 10 concerne un mouillage situé à  $51^{\circ} 26' 40''$  N  $3^{\circ} 26' 40''$  E qui a fonctionné du 14 avril au 14 mai 1975 et la figure 11 un mouillage qui a travaillé du 2 au 16 septembre 1975 à la position  $51^{\circ} 21' 22''$  N  $3^{\circ} 02' 34''$ . Les figures a sont les graphiques de la vitesse, de la direction et de la température de l'eau en fonction du temps. Les figures b représentent les diagrammes vectoriels progressifs du courant pour toute la période de fonctionnement. Les figures c enfin sont les doubles ellipses de marée correspondant à des jours caractéristiques des mouillages.

N.B.A. - DNC 2A      pos: 51° 26' 40" N      03° 26' 40" E  
Period: 14. 04. to 14. 05. 1975      recording interval: 10 min  
depth from seabottom: 3 m      water depth: 8 m

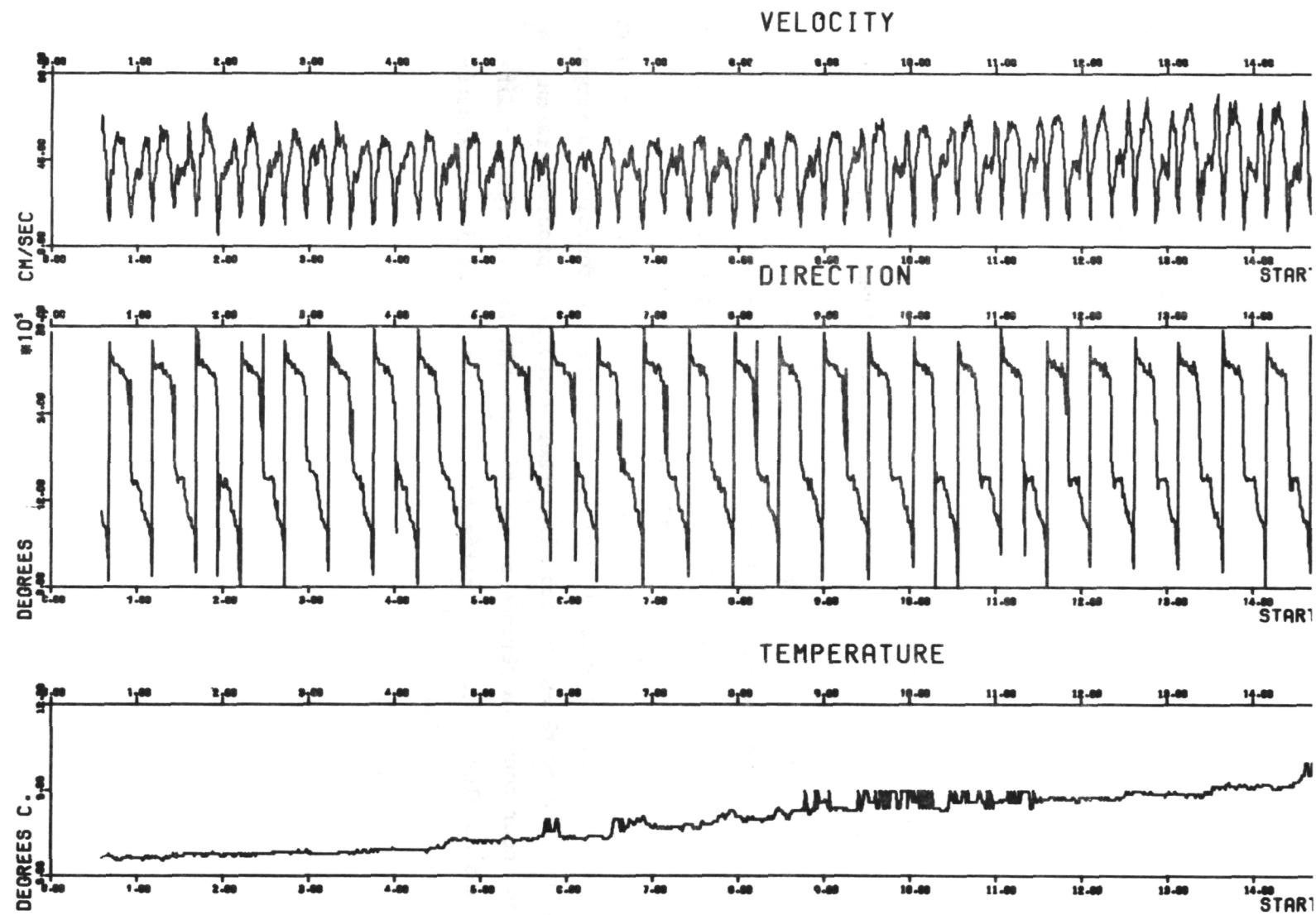
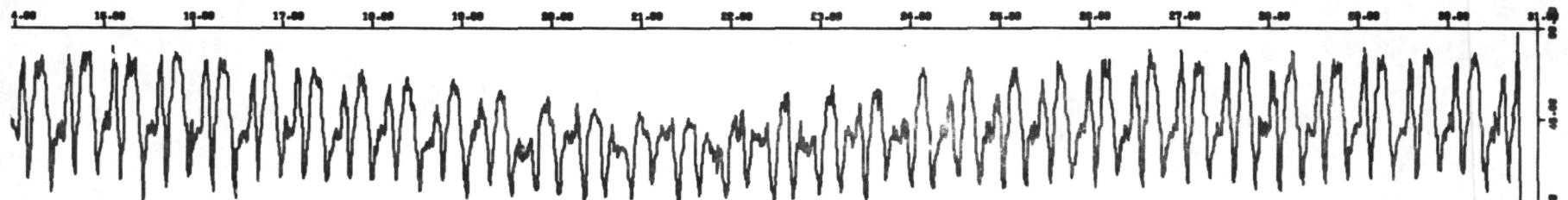
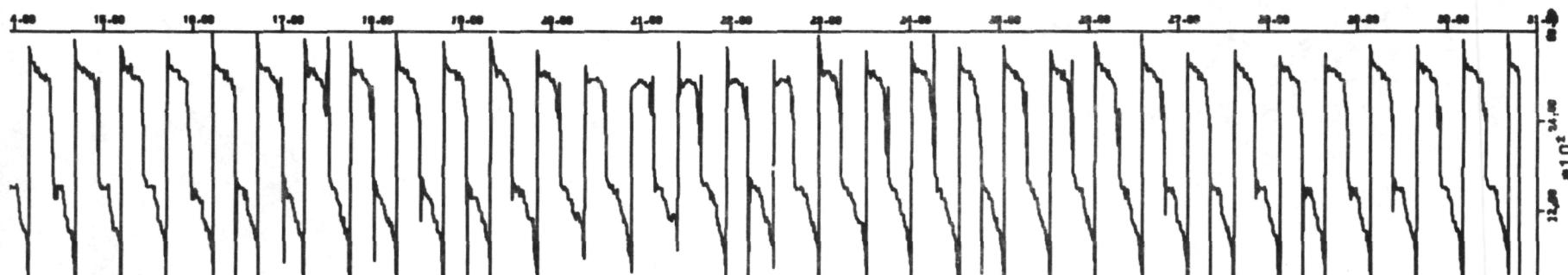


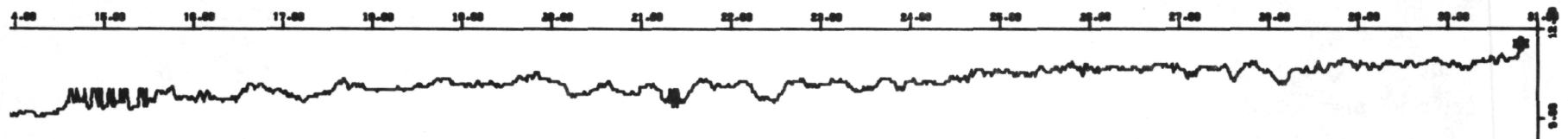
fig. 10a. (début)



START-DAY : 14-4-75



START-DAY : 14-4-75



START-DAY : 14-4-75

fig. 10a. (fin)

N.B.A. DNC2A PROGRESSIVE VECTOR DIAGRAM

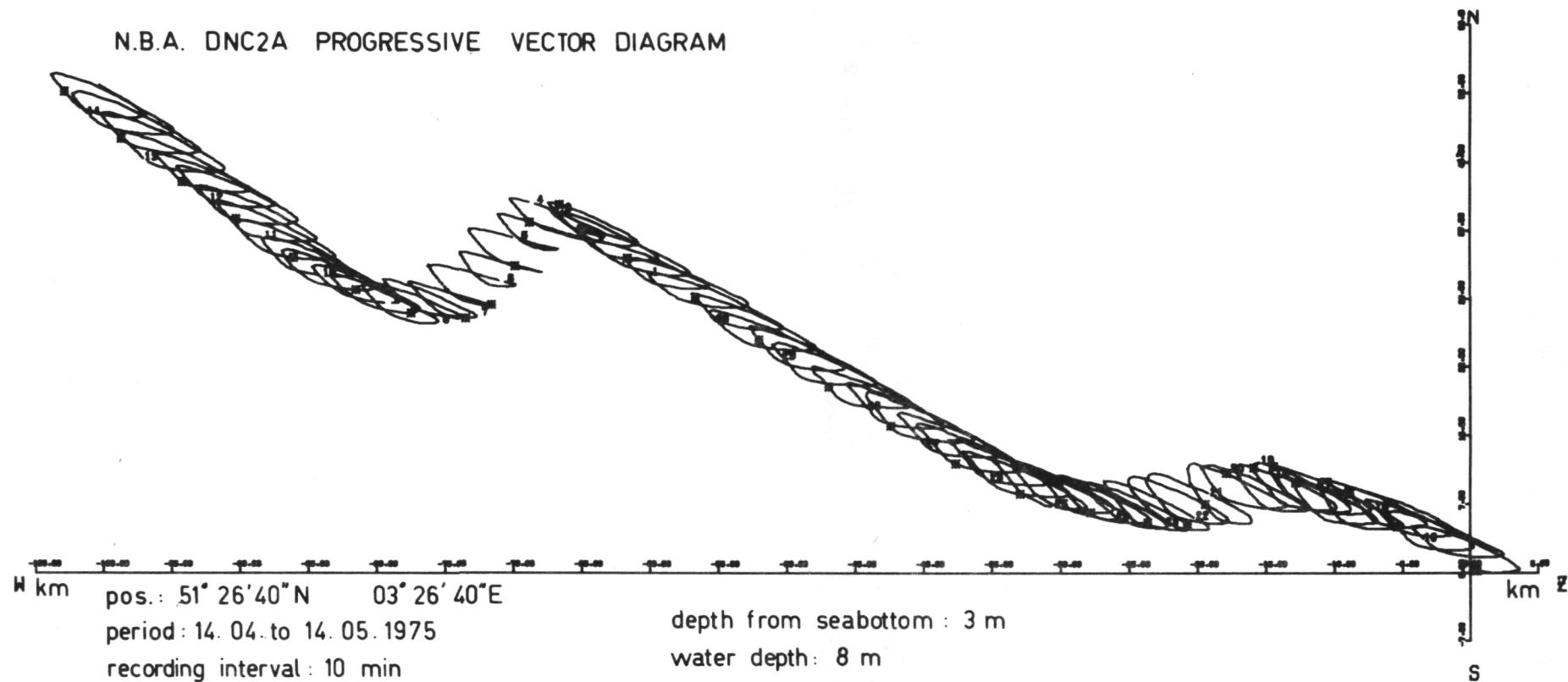


fig. 10b.

DATE : 15. 04. 1975

recording interval : 10 min

N.B.A. DNC2A TIDAL ELLIPSE

pos.: 51° 26' 40"N  
03° 26' 40"E

depth from seabottom : 3 m  
water depth: 8 m

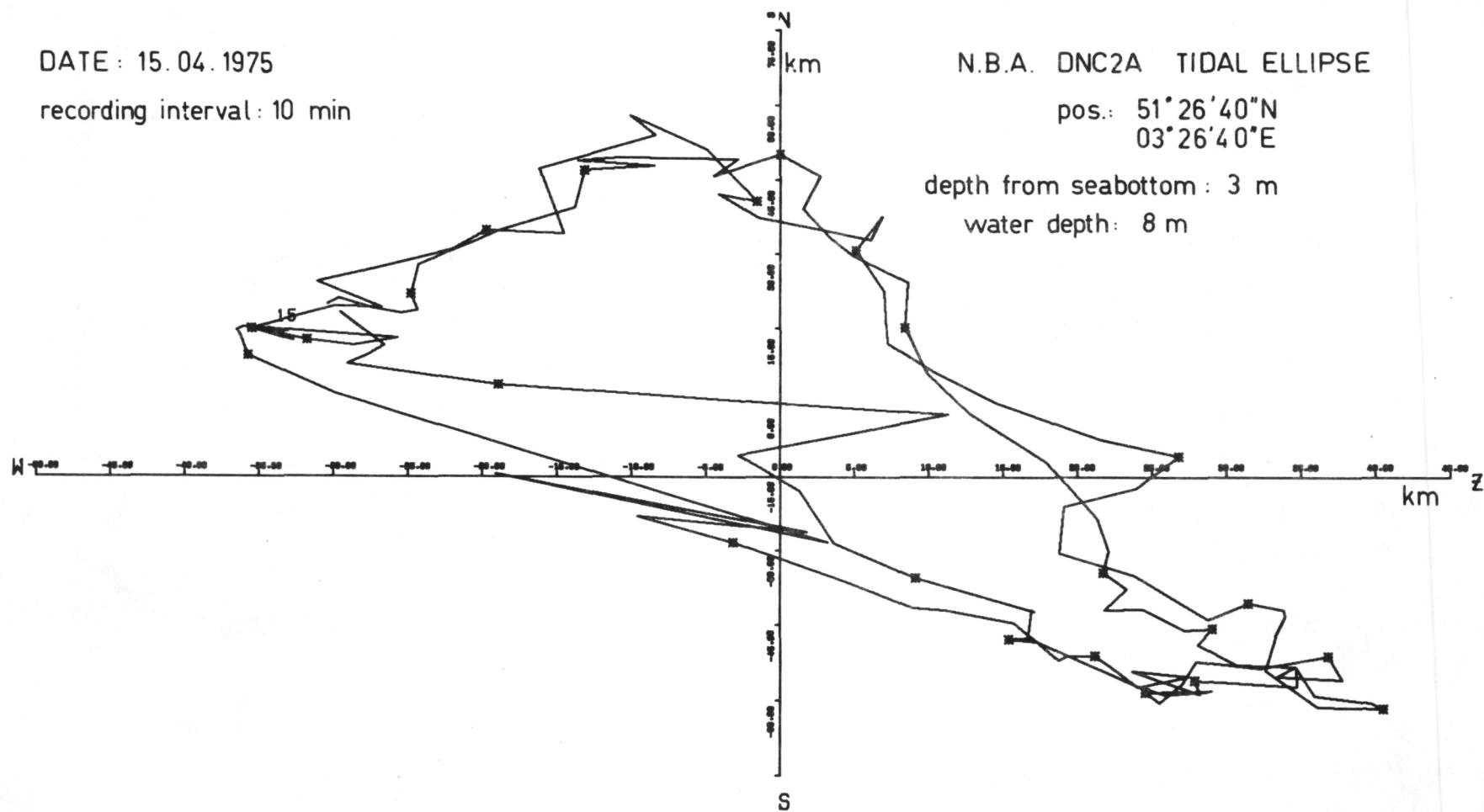


fig. 10c.

N.B.A. - DNC 2A pos.: 51° 21' 22" N 03° 02' 34" E  
period: 02.09. to 16.09.1975 recording interval: 5 min  
depth from sea bottom: 2 m water depth: 8 m

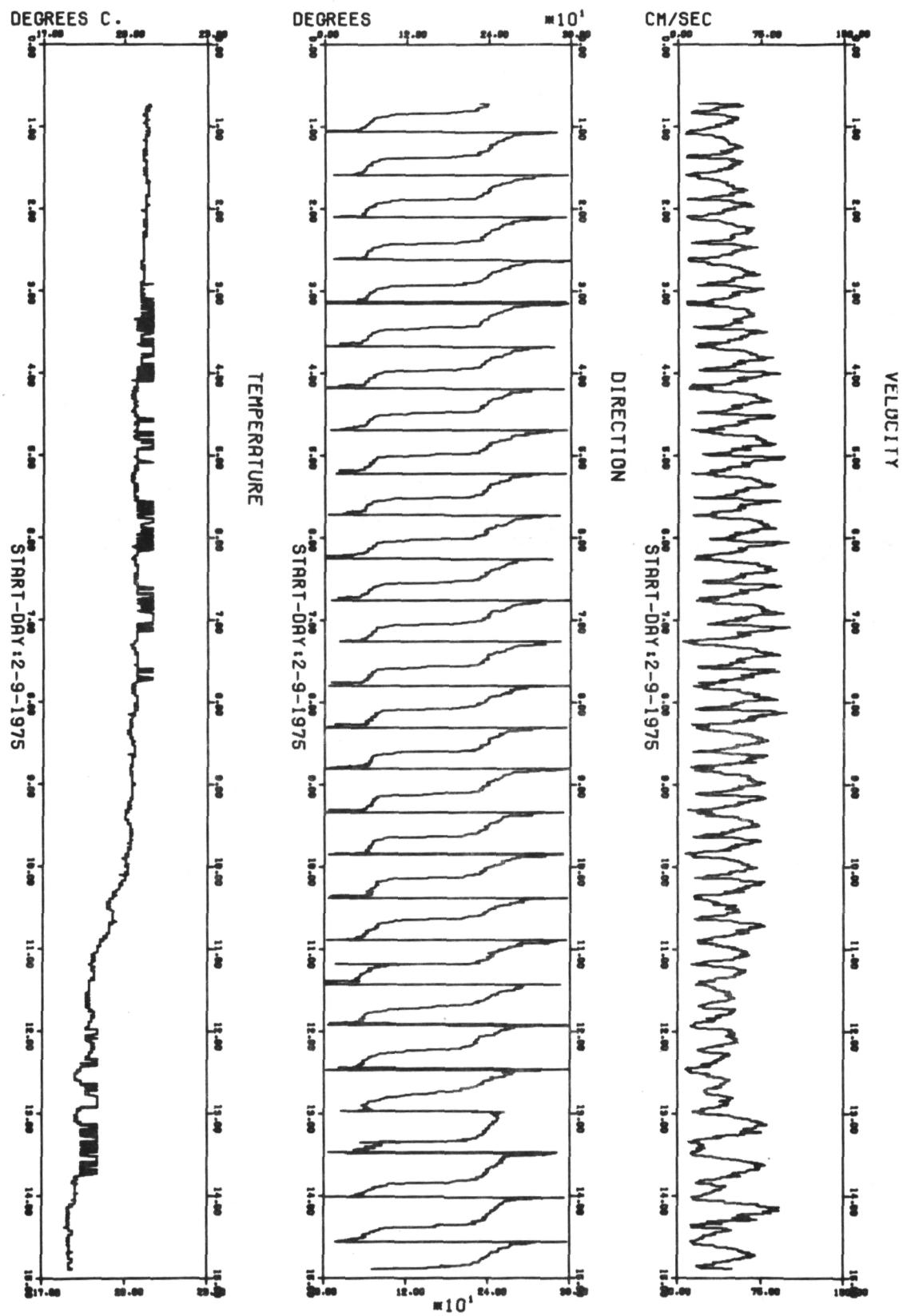


fig. 11a.

N.B.A. DNC 2A PROGRESSIVE VECTOR DIAGRAM

pos. 51°21' 22" N      period: 02.09. to 16.09. 1975  
03°02' 34" E      recording interval: 5 min

depth from sea bottom: 2 m

water depth : 8m

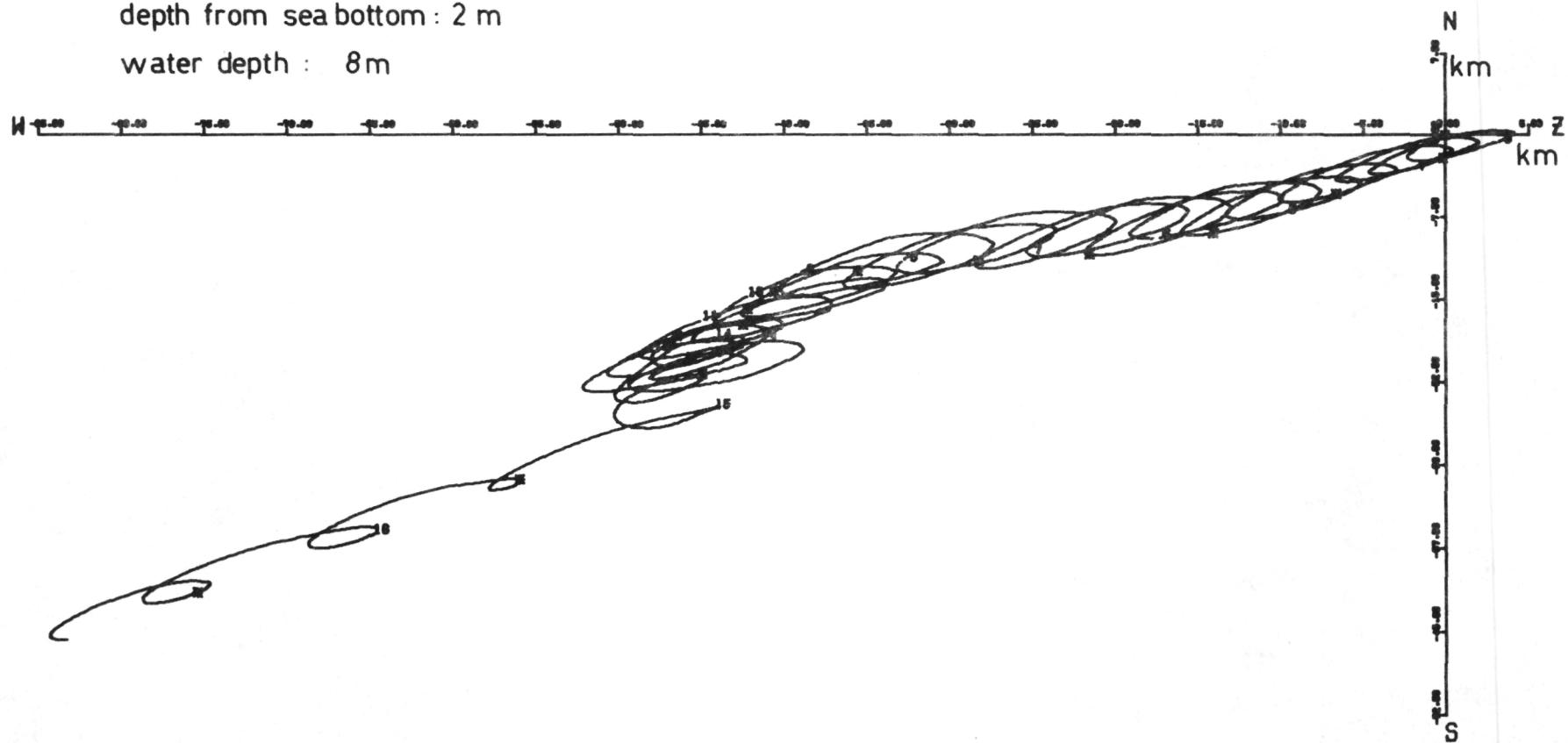


fig. 11b.

DATE 03.09.1975

recording interval 5 min

N.B.A. DNC 2A TIDAL ELLIPSE

pos.  $51^{\circ} 21' 22'' N$   
 $03^{\circ} 02' 34'' E$

depth from seabottom : 2 m

water depth : 8 m

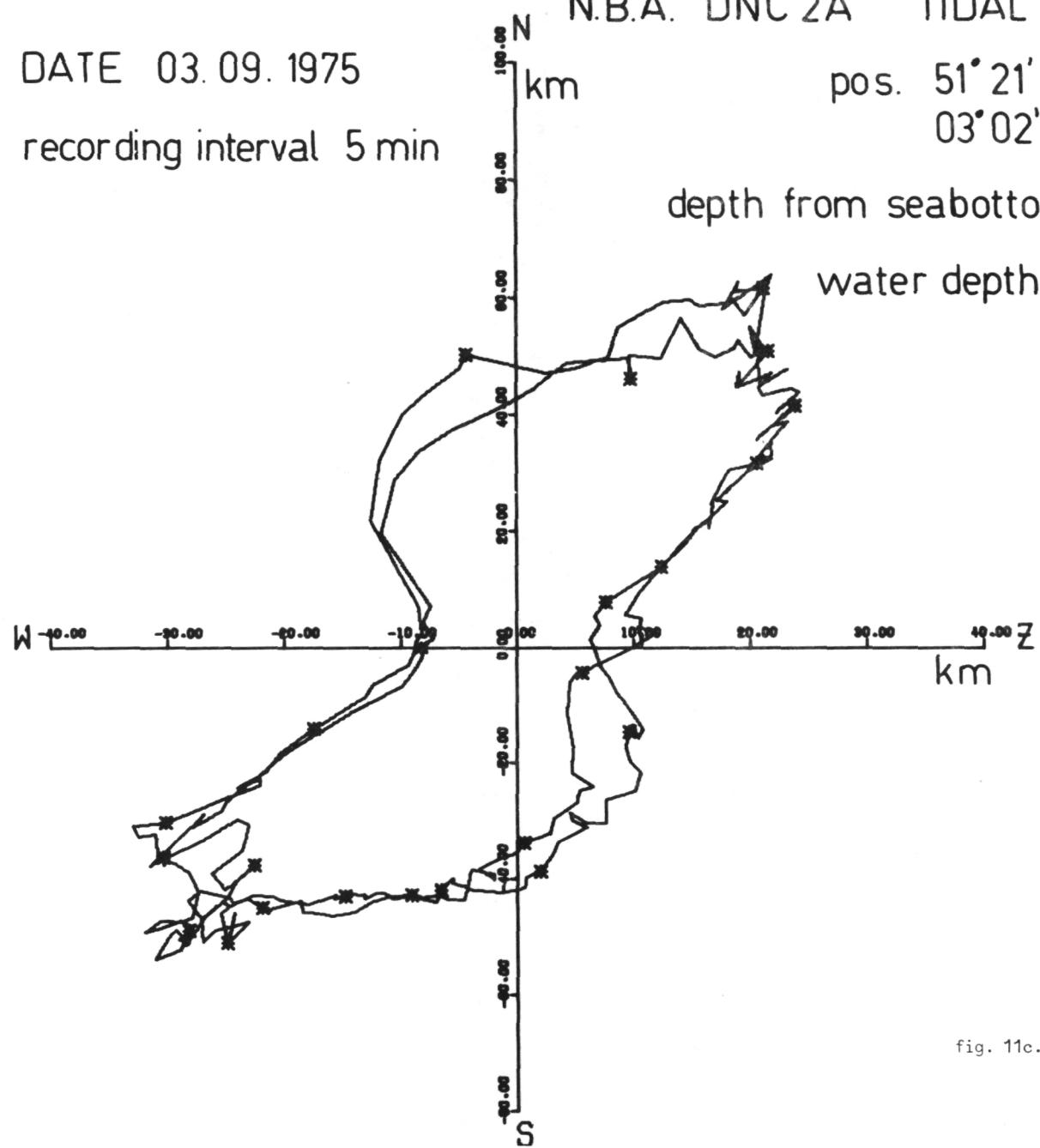


fig. 11c.

## Chapitre II

### Les bouées océanographiques et météorologiques

La mesure continue des variables océanographiques et météorologiques a nécessité la mise en service de stations de mesures automatiques.

Dès 1971, un groupe de travail créé pour effectuer une étude de marché a examiné les bouées suivantes : Comex (Telem 150, HSB/1 et Borem), Sysna, Selco (type 6, 7, 12 et 17), Simrad (SBI, MET et *satellite commanded*), Hagenuk (UBA et Stamob), LCT (L55), Dornier, General Oceanics et Saclantcen. Il tire les deux principales conclusions suivantes :

- 1) La demande en bouées scientifiques est rare, irrégulière et géographiquement très dispersée, ce qui va à l'encontre des conditions requises pour le développement industriel, la compétition effective et l'existence d'un marché sain.
- 2) En dépit d'une publicité considérable, la plupart des bouées n'ont jamais été parfaitement testées.

La bouée que nous souhaitions devait être suffisamment grande et robuste pour contenir en toute sécurité un grand nombre de senseurs très divers, être facilement remarquée par le trafic maritime et pouvoir fonctionner de façon autonome pendant de longues périodes et dans toutes les conditions météorologiques. Le choix s'est porté sur la bouée française SYSNA qui semblait le meilleur compromis entre nos souhaits et le budget qui nous était alloué.

### 1.- Description de la bouée SYSNA (fig. 12)

Le corps de la bouée Sysna est un disque de 3 mètres de diamètre et de 1,5 mètre de haut. Sa bonne stabilité est encore augmentée par deux ballasts. Le premier, de 900 kg , est fixé au bas de la bouée sous forme de courte ellipse profilée. Le second, de 1500 kg , est suspendu, à 8 mètres sous la bouée, par une ligne de fibre synthétique en forme de V . La chaîne d'ancrage lui est fixée.

Quatre capots, disposés symétriquement sur le pont de la bouée, couvrent les puits cylindriques contenant les *racks* de batteries. Un cinquième capot donne accès au compartiment principal où sont fixés deux *racks* standards (19 pouces) d'un mètre de haut pour l'instrumentation électronique. Bien que cette chambre contienne toute l'électronique, l'émetteur, la panneau de commande et de distribution et quelques batteries sèches, beaucoup d'espace reste encore disponible pour des équipements supplémentaires. Enfin, plusieurs enceintes plus petites mais complètement étanches assurent une réserve de flottabilité en cas d'inondation du compartiment principal.

Le mât tubulaire central de 8 mètres de haut est haubanné par 4 cables en acier inoxydable. Il supporte une structure tubulaire où sont fixés la lampe flash, le réflecteur radar, les antennes et les senseurs météorologiques. Deux aérogénérateurs dont un seul suffit pour garder les batteries à pleines charges, sont fixés à mi-hauteur du mat, pour ne pas perturber les mesures météorologiques.

Le tableau 2 résume les caractéristiques mécaniques. Toutes les précautions ont été prises dans le choix des matériaux pour éviter la corrosion galvanique des parties submergées.

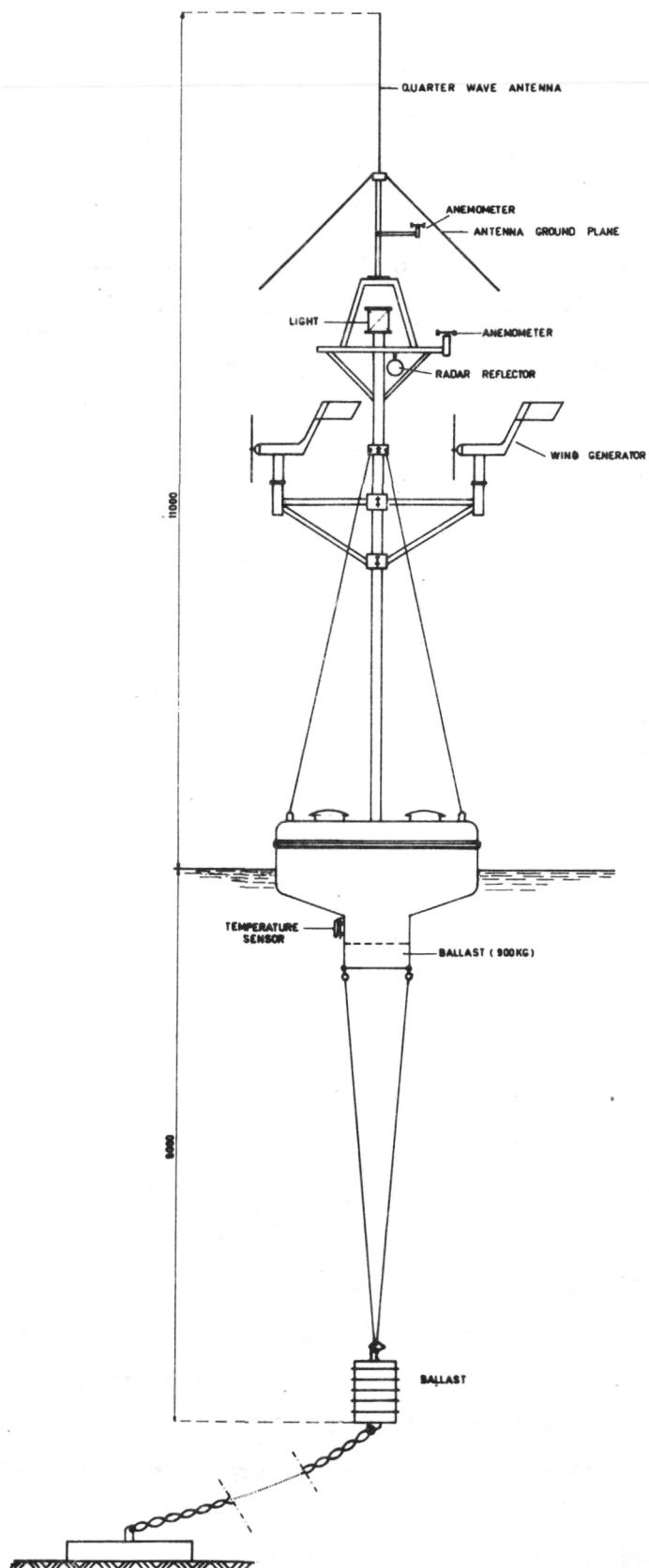


fig. 12.  
La bouée Sysna

Tableau 2

Caractéristiques mécaniques de la bouée Sysna

<u>Mechanical characteristics</u>	
Overall diameter of the float	3,021 mm
Height of the float (including small ballast tail)	2,180 mm
Maximum height inside	1,580 mm
Mast length	8,100 mm
Mast diameter	150 mm
Displacement	2,000 kg
Float volume	10 m <sup>3</sup>
Stability	2 t.m
Rolling period	2.8 s
Heave period	1 s
Spare buoyancy (with 1 ton of batteries and equipment)	7 tons
<u>Materials used</u>	
Buoys : AG4MC aluminium alloy, specially resistant to marine corrosion.	
Tubes : AG3 . Cast parts : AS 13 74 . Guys and guy fixings : stainless steel 316 (18/12 Mo).	

## 2.- Le système d'acquisition de données (D.A.S.)

Il était convenu originellement que la bouée Sysna devait être livrée avec, en outre, certains senseurs océanographiques et météorologiques, le système d'acquisition de données et l'appareil de décodage. Le constructeur s'est finalement désisté et l'électronique désirée n'existant pas ailleurs sur le marché, l'unité d'Ostende a été obligée de développer et de construire, elle-même, un système polyvalent d'acquisition de données (appelé D.A.S. dans la suite du texte).

### 2.1.- Caractéristiques générales du D.A.S.

La quantité de données, la résolution et la précision exigées impliquent que le D.A.S. placé dans les bouées et la station de réception à terre soient des appareils digitaux. Cette digitalisation des signaux permet une collecte fidèle et un traitement ultérieur efficace des données (fig. 13).

Les senseurs disponibles sur le marché produisent une grande variété de signaux de sortie. Aussi, le D.A.S. doit-il servir de système multiplexeur capable de gérer les signaux, qu'ils soient

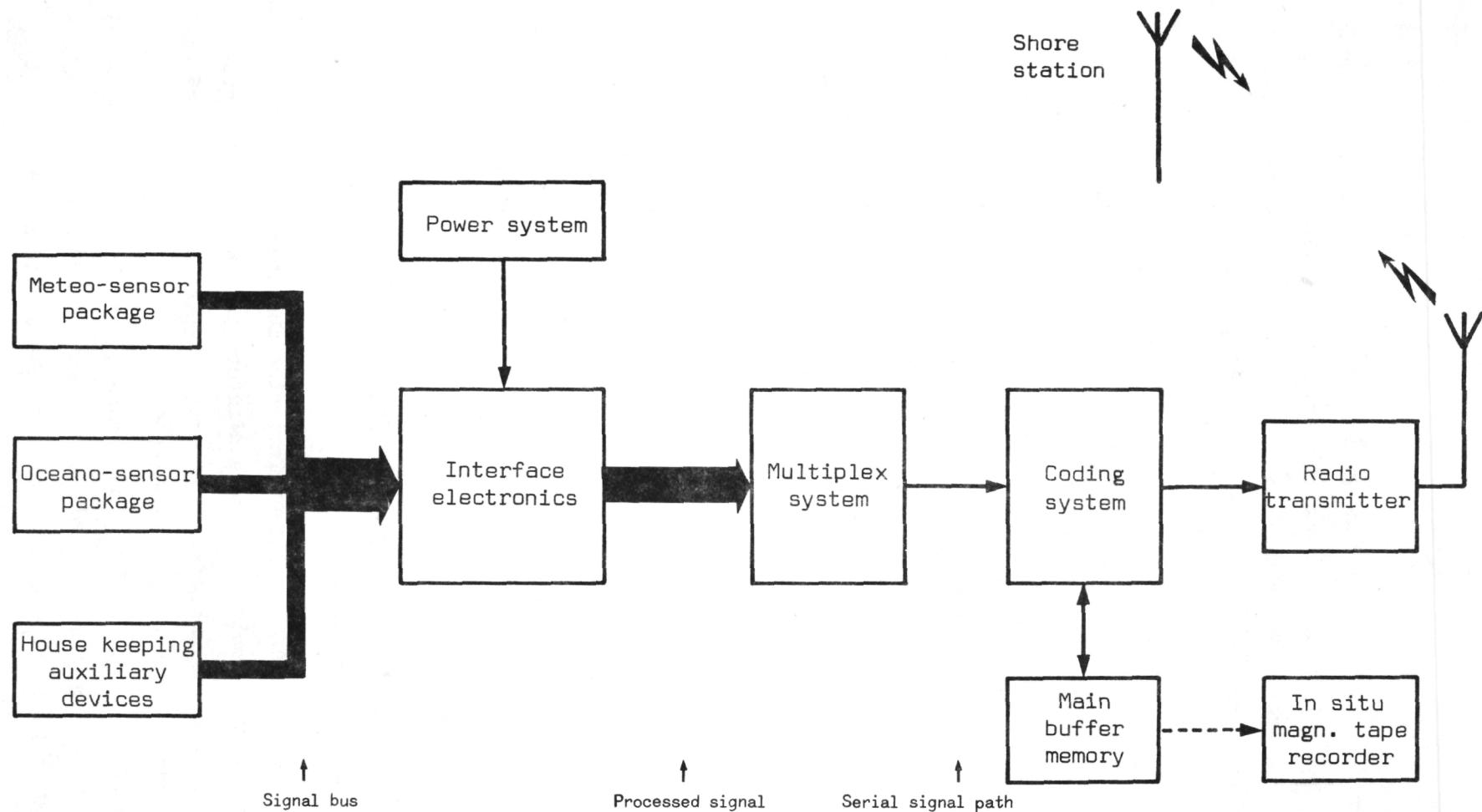


fig. 13.  
Le système d'acquisition des données (D.A.S.)

analogiques ou digitaux. L'interrogation des signaux analogiques se fait en utilisant les classiques convertisseurs analogiques-digitaux. Mais cette conversion se fait aussi près que possible de la source des signaux, pour éviter tout problème de bruit et réduire au maximum les parasites.

Avant d'être multiplexés, les signaux analogiques sont d'abord traités de telle façon qu'ils correspondent aux possibilités du multiplexeur et du convertisseur digital. Les senseurs dont la sortie est déjà digitale peuvent éventuellement subir une conversion de code (BCD, Binary, Gray, 2's complement, ...) ou une adaptation à un système logique (TTL, CMOS, ...).

Le temps disponible entre deux mesures successives (notamment pour les séries temporelles de mesures de houle) correspond au taux minimum de transfert des données tandis que le maximum est fixé par le transmetteur (actuellement un émetteur-récepteur VHF) et l'enregistreur *in situ* (Kennedy model 330 cartridge recorder). La transmission des données et leur stockage *in situ* requièrent le choix d'un code efficace et fiable.

Le D.A.S. doit être un instrument "de terrain" robuste et fiable à consommation limitée, compte tenu des problèmes d'alimentation. La précision des mesures ne doit pas être liée aux variations de la température ambiante. Des précautions particulières doivent être prises à l'encontre de l'humidité et du milieu salin.

## 2.2.- Les circuits intégrés CMOS du D.A.S.

### 2.2.1.- Description générale

La figure 14 schématisse les principales fonctions du D.A.S. et la figure 15 donne un exemple d'un cycle d'interrogation. L'emploi quasi généralisé de circuits intégrés CMOS conduit à une haute fiabilité et réduit fortement la consommation d'énergie.

Les différents senseurs océanographiques et météorologiques ainsi que les unités de contrôle (*housekeeping*) sont interrogés par un multiplexeur extensible à 64 canaux. La version actuelle en possède 32,

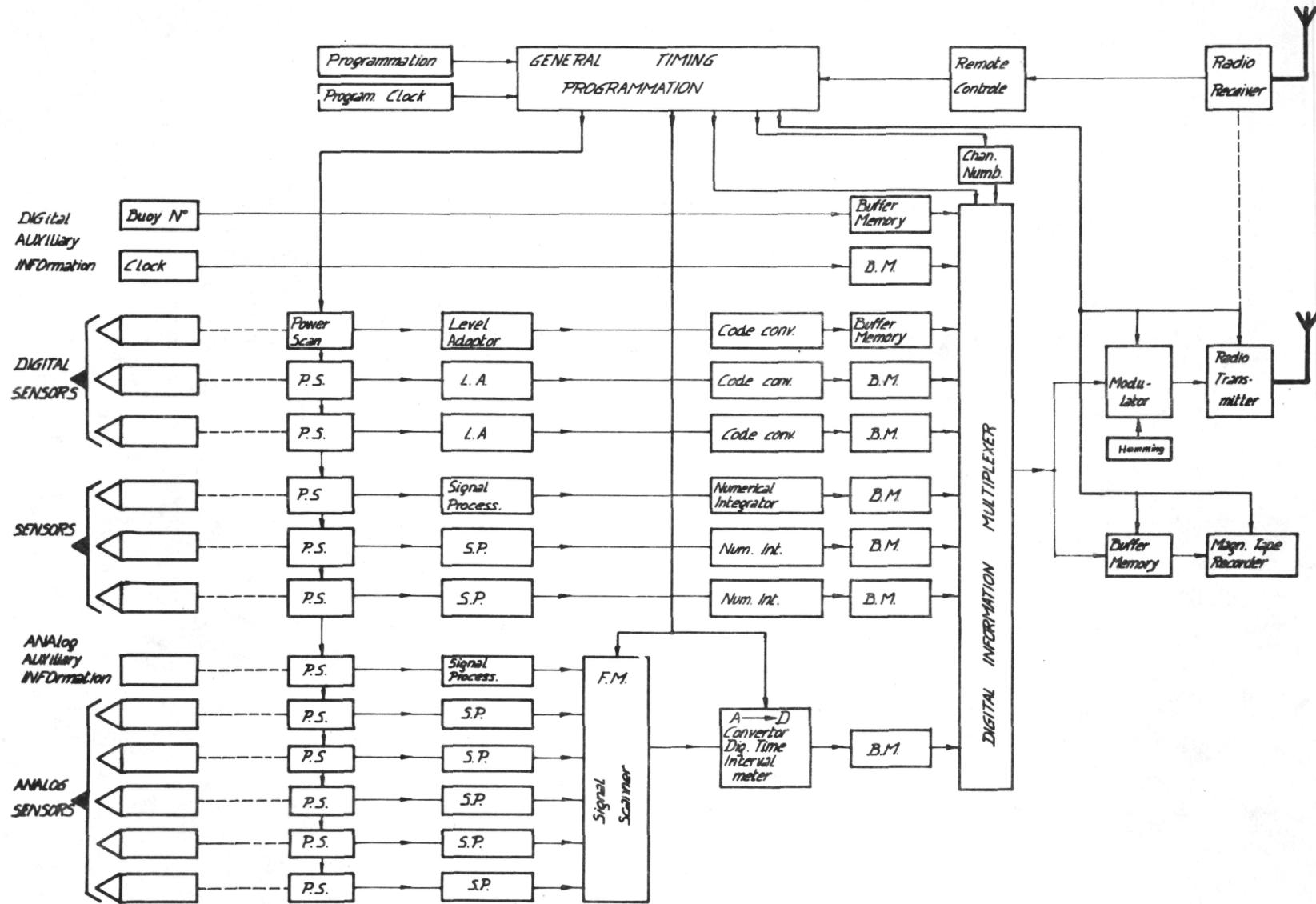


fig. 14.

Système d'acquisition des données CMOS

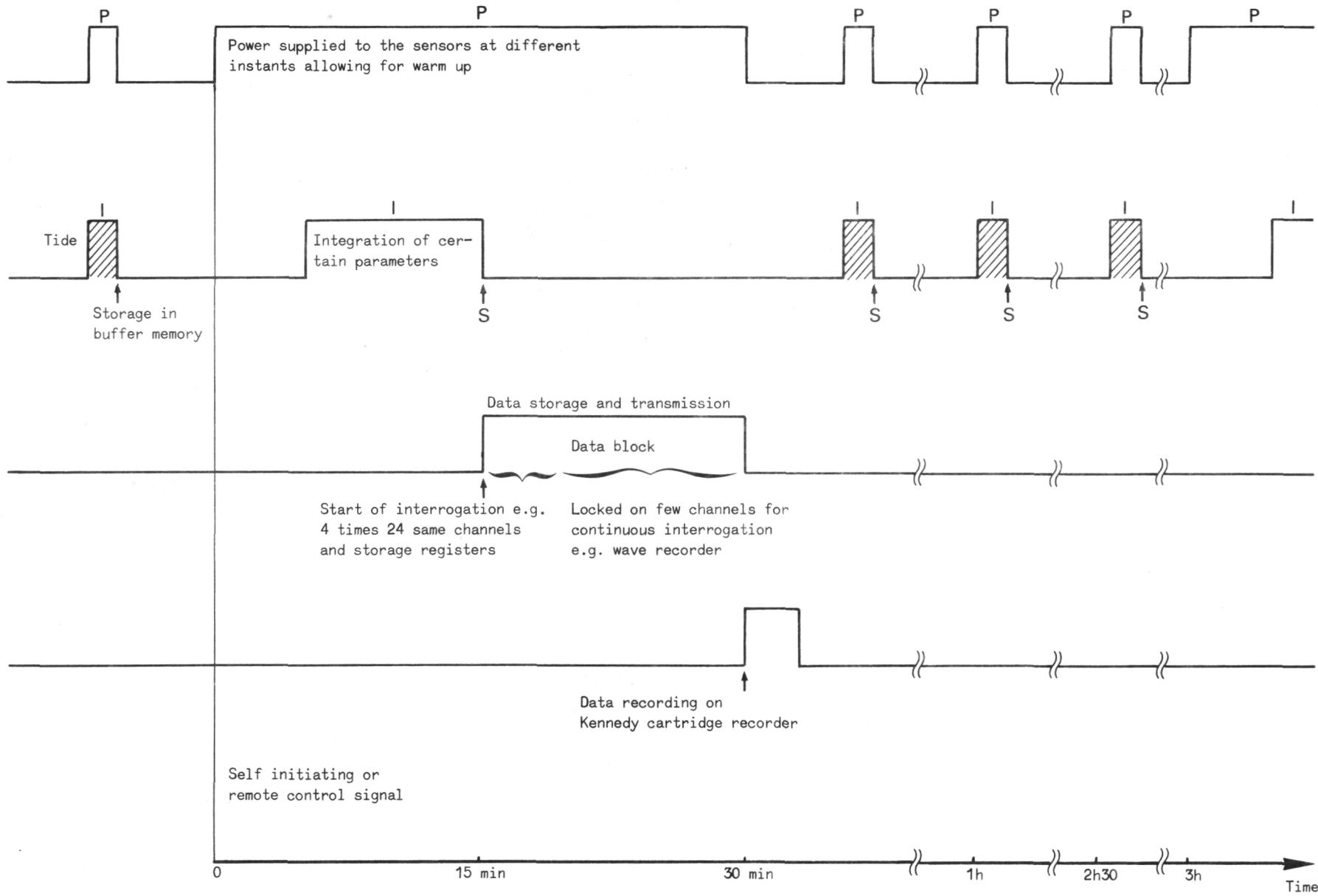


fig. 15.  
 Cycle de mesure du D.A.S.

répartis en deux groupes : le multiplexeur analogique à 16 canaux suivi d'un convertisseur A-D (*Time Interval Meter*) et le multiplexeur digital à 16 canaux. Une programmation générale fixant le rythme et la fréquence journalière des interrogations contrôle le multiplexeur. Vu l'énergie disponible, la collecte et l'émission continues de données n'est pas possible et l'utilisateur doit choisir un rapport *ON-OFF* d'interrogation accommodant au mieux la consommation d'énergie et le besoin minimum de données précises. L'interrogation peut se faire de deux façons différentes dépendant de l'excitation initiale :

- suivant un mode synoptique : où la collecte de données est enclenchée par une horloge et où leurs acquisition, décodage, émission et stockage s'effectuent à des intervalles de temps réguliers et préalablement fixés;

- suivant un mode télécommandé, où la collecte est mise en route par un signal émis de la station à terre.

De plus, le système d'acquisition permet, entre deux émissions, d'intégrer certains signaux, d'effectuer certaines mesures continues ou de stocker certaines valeurs caractéristiques jusqu'au moment où elles sont envoyées en bloc à la station côtière.

Un *scanner* de puissance, consistant en une série de relais couplés optiquement est chargé de fournir aux senseurs la tension d'alimentation adéquate. La séquence de *scanner* dépend de la durée d'interrogation du senseur et du temps de chauffe nécessaire. Cela permet de réduire fortement la consommation d'énergie, le *cross talk* sur les câbles et la dérive due au *self-heating*.

Des interfaces appropriés adaptent les différents signaux d'entrée, effectuent l'intégration numérique de certains paramètres, enregistrent le temps des mesures, etc. Les informations digitales sont transformées par le modulateur en code "*Pulse Length*" puis envoyées directement à l'entrée de l'émetteur à modulation de phases.

Dans certains cas (par exemple lorsque la station de réception est très éloignée), les données peuvent être stockées dans une mémoire tampon CMOS et ensuite stockées sur un enregistreur digital "Kennedy

cartridge". Enfin, par un signal d'excitation émis de la station de réception, cette mémoire tampon peut être interrogée et son contenu transmis.

Le tableau 3 résume les caractéristiques techniques du D.A.S.

Tableau 3

Caractéristiques techniques du D.A.S.

- 1) Multiplexer : 16 analog and 16 digital channels
  - Single channel repeat selector for testing and calibration
  - Interrogation speed : .1 , .2 , .5 , 1 , 2 , 5 , 10 channels/s
  - Interval between cycles : continuous , 20<sup>f</sup> , 1 h , 3 h , 6 h , 1 day
- 2) Analog to Digital Convertor (Time Interval Meter)
  - Accuracy : 10<sup>-5</sup> Hz
  - Full scale resolution : 10<sup>-6</sup>
- 3) Input Signals
  - Frequency modulated a.c. or pulses (standard 1 V min. amplitude)
  - D.C. : 0-10 mV , 100 mV, ± 1 V , ± 5 V , ± 10 V
    - Input impedance > 100 kohm (1 M)
    - Accuracy > .1 %
  - Digital : Standard B.C.D. coded parallel or serial inputs
- 4) Output Signals
  - Serial B.C.D. "8.4.2.1." code 40 bits/channel
  - Digit 1 : word separation, 2 : sign, 3 & 4 : channel # ,  
5-10 : measurement
  - Modulator : Hamming code, 40 parity bits added (80 bits/channel)
- 5) Power Requirements (D.A.S. only, without CMOS memory, transmitter, recorder)
  - Main electronics + 15 V : 100 mA , - 15 mA : 20 mA
  - Power scan + 15 V : 10 mA/relay
  - Voltage to period convertor : ± 15 V , 10 mA
- 6) Dimensions
  - Main electronics and standard interfaces : two 19" racks, 4 units high each
- 7) Environmental
  - Operating temperature range : - 5 °C to 40 °C
  - Humidity : a completely sealed compartment for 4 racks is used

#### 2.2.2.- Traitemennt des signaux

Suivant la nature d'un senseur donné, le signal de sortie est analogique ou digital.

S'il est digital, l'information est stockée dans une petite mémoire tampon, sous forme B.C.D. éventuellement, après une adaptation de niveau et une conversion de code. Ce stockage peut se passer durant

le cycle de mesures, ou avant, fournissant de cette manière des mesures instantanées ou intégrées.

S'il n'est pas présent sous forme digitale, le signal est transformé en un signal modulé en fréquence au moyen d'une électronique classique dont les principales qualités requises sont la fiabilité, la précision et la faible consommation d'énergie. Une fois présent sous forme F.M. le signal peut être digitalisé de deux façons différentes suivant que la mesure doit être intégrée ou non. Si la donnée doit être intégrée, la conversion analogique-digitale se fait par une mesure de fréquences dont la résolution dans le temps dépend des désiderata de l'utilisateur. Chaque résultat est ensuite stocké dans une mémoire tampon identique à celle des données digitales. Cependant, la souplesse du système permet de garder cette donnée disponible pour l'interrogation instantanée, par exemple pour la production de séries temporelles. Si la donnée ne doit pas être intégrée, le signal passe par un multiplexeur analogique à 16 canaux, immédiatement suivi par un mesureur d'intervalle de temps. Le résultat est momentanément stocké dans une mémoire tampon. Le taux d'échantillonnage de ce mesureur d'intervalle de temps est synchrone avec la fréquence de *scanning*.

L'information de chaque canal analogique est maintenant disponible dans la mémoire tampon. Une fois que l'horloge met en route un cycle de mesures, le contenu des mémoires tampon est sous contrôle du multiplexeur digital, transféré en séries vers l'encodeur Hamming et le modulateur.

Si des canaux analogiques doivent être interrogés, le multiplexeur digital reste bloqué sur la mémoire tampon relative au mesureur d'intervalle de temps et passe donc le contrôle du *scanning* au multiplexeur analogique. En plus, le multiplexeur analogique lui-même peut être bloqué sur certains canaux pendant une durée programmable. Les données disponibles en BCD série après le multiplexeur sont affectées d'un numéro de canal puis stockées dans une mémoire tampon CMOS. A la fin de l'interrogation, les données sont alors stockées en bloc sur l'enregistreur Kennedy.

Avant ou après enregistrement *in situ*, la série de données est envoyée dans un encodeur Hamming qui est un générateur de bits de parités permettant la détection et la correction automatiques des fautes à la station de réception. La sortie de l'encodeur Hamming est envoyée au modulateur où la série d'informations est transformée en code P.C.M. puis connectée à l'entrée de l'émetteur.

Le bloc de données enregistrées et/ou transmises, correspondant à un cycle de mesures consiste, dans la version actuelle, en 4 interrogations successives des 32 canaux suivies de l'interrogation continue de 1 à 4 canaux fixés (par exemple la houle et la vitesse du vent) pendant un temps déterminé (par exemple 10 minutes). Le tableau 4 donne à titre d'exemple un aperçu des blocs de données émises lors des mouillages Sysna II (mars 1975) et Sysna III (août 1975).

Tableau 4

Blocs de données émises pendant les mouillages Sysna II et III

Digital channels 1-8 , 25-32

Analog channels 9-16

Continuous interrogated channels I-IV analog

Channel 1	Data station #	Channel 17	Water temperature I
2	Data block #	18	Water temperature II
3	# of days	19	Barometric pressure
4	Elapsed time clock	20	Idem
5	Integration time	21	not used
6	Conductivity (1' INT)	22	not used
7	Incident radiation (1' INT)	23	V-P mercury test
8	not used	24	N.B.A. wind speed inst.
9	Flood alarm	25	Friedrich wind speed (10' INT)
10	Aerogenerator current I	26	N.B.A. wind speed (10' INT)
11	Aerogenerator current II	27	VAWM COSINE (10' INT)
12	Battery voltage	28	VAWM SINE (10' INT)
13	DC/DC voltage	29	VAWM rotor count (10' INT)
14	Buoy temperature	30	not used
15	Air temperature I	31	not used
16	Air temperature II	32	not used

Channel I 128 continuous interrogations of N.B.A. anemometer

Channel II-IV 128 continuous interrogations of Datawell wave recorder

### 3.- Installation embarquée de transmission des données

Comme, dans la phase actuelle de préparation, la distance entre station de réception et station de mesure est relativement réduite (de l'ordre de 50 milles maximum), la transmission se fait par un émetteur transistorisé de faible puissance (10 à 20 W) travaillant dans la bande V.H.F.

Le tableau 5 résume les caractéristiques techniques des émetteurs-récepteurs.

Tableau 5

#### 4.- Enregistrement *in situ* sur bande magnétique

Il peut arriver que les expériences scientifiques exigent des mesures continues, impossibles à transmettre pour des raisons énergétiques ou que la transmission elle-même soit mauvaise (parasites, ...) ou impossible (pannes des émetteurs-récepteurs, bris d'antennes, ...). Il faut donc pouvoir enregistrer les données *in situ* sur un enregistreur à bandes magnétiques. Ces données obtenues en très grande quantité compte tenu de la mesure continue de la houle et de l'autonomie de la bouée (six semaines minimum) requièrent un enregistreur digital fiable à grande capacité. Le choix s'est porté sur l'enregistreur Kennedy cartridge modèle 330 à taux de transfert de 40K par seconde, à densité d'enregistrement de 1600 BPI, et dont les spécifications sont résumées au tableau 6.

Tableau 6

Spécifications de l'enregistreur Kennedy modèle 330

Cartridge type	3M, DC300A Isoelastic Data Cartridge
Tape width	0.25 inch, 1.0 mil
Tape length	300 feet
Recording density	1600 bpi
Number of tracks	1, 2 or 4
Recording head	Dual gap read-after-write with separate erase bar for each track
Record format	Single track, phase encoded
Record mode	Single track serial
Normal tape speed	25 ips
Rewind/search tape speed	90 ips nominal
Start/stop time	at 25 ips, 30 ms $\pm$ 1 ms; at 90 ips, 120 ms nominal
Start/stop displacement	at 25 ips, 0.38 inch; at 90 ips, 5.4 inches nominal
Transfer rate	40,000 bits per second
Instantaneous speed variation	$\pm$ 3 percent
Long term speed variation	$\pm$ 1 percent
Write	Serial - forward direction only
Read	Bidirectional
Power requirements	+ 5 V dc at 1.0 amp $\pm$ 15 V dc at 0.2 amp $\pm$ 24 V dc at 6 amps peak
Interface	DTL/TTL low true
Weight	7 pounds
Dimensions	6.5 inches wide, 4-3/4 inches high, 12 inches deep

Après chaque interrogation, les données sont à la fois transmises par radio et stockées dans une grosse mémoire tampon pour être ensuite enregistrées en blocs sur le Kennedy. Cette mémoire donne la possibilité de transmettre, sous l'impulsion d'un signal émis par la station à terre, le dernier bloc de données enregistrées.

Les interfaces nécessaires entre le D.A.S., la mémoire tampon et l'enregistreur (fig. 16) ont été développés en coopération avec le bureau d'études électroniques Countach. Un microprocesseur assure la supervision des différentes fonctions. Un circuit *standby* permet de mémoriser en toutes circonstances certaines fonctions comme la numérotation des pistes et des blocs. La mémoire centrale comprend deux parties : la mémoire centrale d'une contenance maximum de 4 k bytes de 48 bits par byte et le contrôleur de mémoire qui règle automatiquement les fonctions lecture, écriture dans la mémoire.

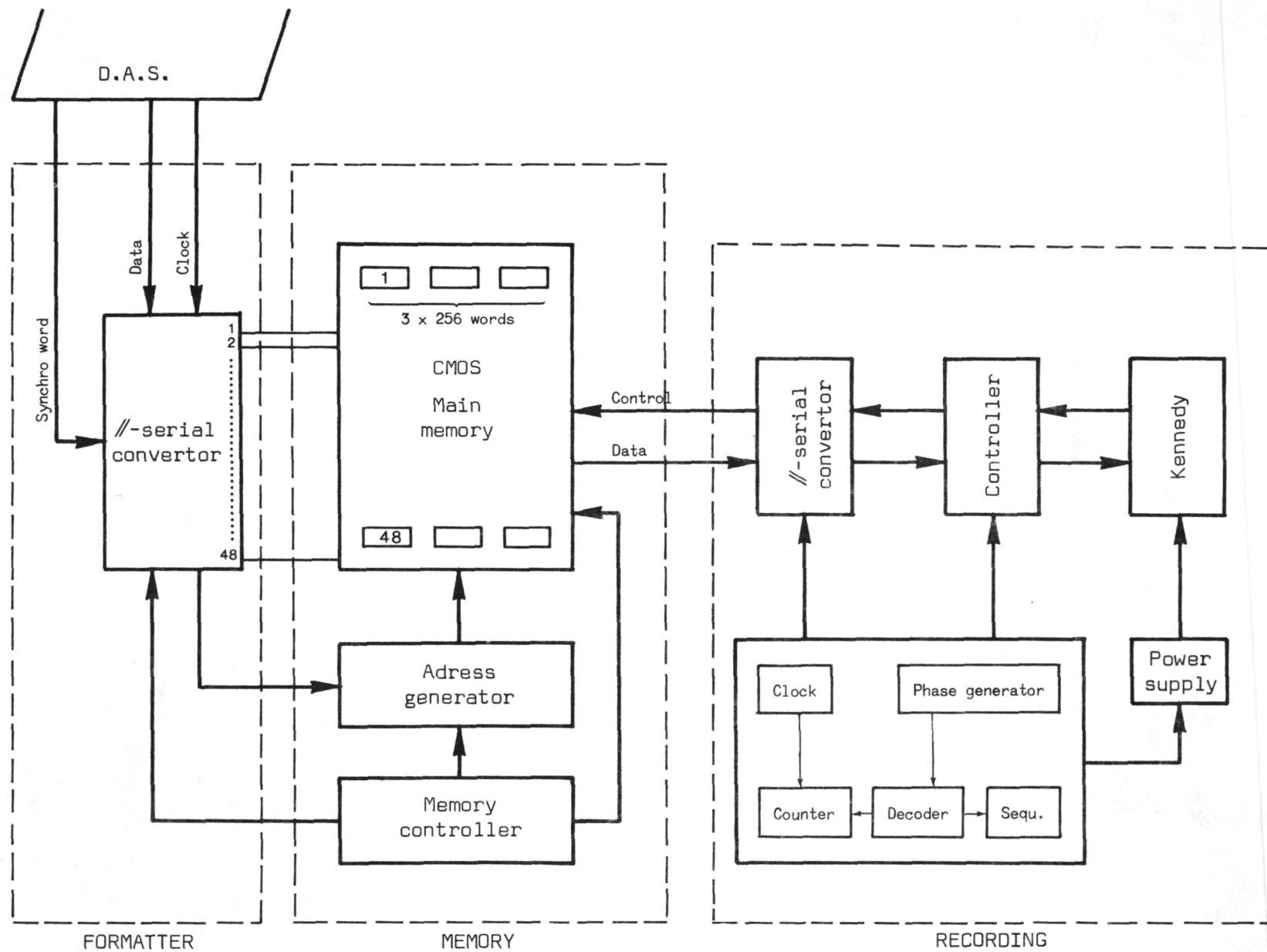


fig. 16.  
L'interface Kennedy - D.A.S.

L'unité de commande de transmission règle l'émission des données à partir de la mémoire centrale soit par télécommande soit après enregistrement sur le Kennedy. L'interface *tape drive* commande les différents signaux, les adaptant à l'enregistrement. Le tout tient dans deux *racks* de 19 pouces.

La consommation, fortement réduite par l'emploi de la mémoire CMOS, dépend du stockage des données et des éventuelles retransmissions.

## 5.- Senseurs océanographiques et météorologiques

### 5.1.- Critères requis des senseurs

Le choix des senseurs océanographiques et météorologiques est lié à l'obligation de réaliser des mesures automatiquement à partir de stations autonomes en mer et de rendre ces données rapidement disponibles.

De dimension relativement réduite, ces senseurs doivent encore être robustes, étanches et résistants à la corrosion. Travaillant automatiquement, ils ont à fournir une grandeur électrique dont les qualités requises sont la dérive nulle, la stabilité à long terme et l'échauffement nul.

Comme l'énergie est fournie par des batteries rechargeables, ils doivent être alimentés, si possible, en tension continue et consommer le moins possible.

La confection d'une interface économique avec le D.A.S. dépend du niveau du signal, de la précision et la résolution désirées, de l'impédance d'entrée, de la variation à fond d'échelle et des spécifications de l'alimentation.

### 5.2.- Les senseurs météorologiques (tab. 7)

#### 5.2.1.- Le vent

La mesure de la vitesse du vent se fait au moyen d'anémomètres "3 cups" de deux types différents : le Friedrich 40311 et le Munro

Tableau 7

Senseurs météorologiques

Parameter	Manufacture	Principle of operation	Range	Accuracy	Output signal	Interface	Remarks
Wind speed 10 m above sea-level	Friedrich	3 cups assembly coupled to an inductive interrupter	.3 - 60 m/s	.2 m/s	Square pulses	Pulse shaper	$10^1$ integration for mean value
	Munro N.B.A.	3 cups assembly coupled to an optically interrupter	2 - 90 kn	$\pm 1$ kn	Square pulses	Pulse shaper	$10^1$ integration and vector averaging
Wind direction	Munro N.B.A.	Vane coupled to a 7 bits Gray optically encoded disc	360°	$\pm 2.8^\circ$	7 bits Gray code	Hatch code convertor vector summation	$10^1$ integration with R.O.M. addressing for sine and cosine values
Compass	N.B.A. (A.M.F.)	Magnetic 7 bits Gray optically en- coded disc	360°	$\pm 2.8^\circ$	7 bits Gray code	Summation with wind vane angle	Idem as above
Barometric pressure	M.B. Gilmore	Strain gage full bridge	900 - 1100 mb	$\pm .25\%$ $\pm .02\%/\text{°C}$	(10 mV) 5 V	(chopped amplifier) V-P convertor	Calibrated at the pressure room of the Royal Meteorological Institute, Brussels
Air temperature 1 & 10 m above sea-level	Selfmade	UJT oscillator with ther- mistor	- 10 °C to + 30 °C	estimated $\pm .1$ °C	F.M. pulses	Pulse shaper	Calibration curve necessary
Incident radiation	Kipp & Zonen	Photosensitive cell	$3 \times 10^{-7}$ m to $2 \times 10^{-6}$ m	depending on cleaning of glass hemisphere	15 mV	Chopped amplifier V-P convertor	$1^1$ integration minimum (Photo diode sensor in de- velopment)

IM 204 MK4. Pour le premier, les impulsions sont formées par une rotation dans un champ inductif mécaniquement couplée à celle des trois "cups". Le second, amélioré par la firme N.B.A. consiste en un système d'impulsions optiques.

La direction du vent est donnée par une vanne Munro IM 205 MK4 couplée à un compas magnétique NBA encodé optiquement.

Les valeurs instantanées de la vitesse et la direction du vent ne sont pas très significatives, à cause du mouvement de la bouée elle-même. Il est donc nécessaire non seulement de prendre les valeurs moyennes de la direction et la vitesse sur une période de données mais encore de calculer leurs moyennes vectorielles. Les détails techniques de cette procédure ont été exposés par Pollentier *et al.* (1974). En résumé, la direction du vent est échantillonnée chaque fois que l'anémomètre a compté 20 impulsions, ce qui correspond à un déplacement d'air de 3,08 m. A chaque échantillonnage, ce vecteur unitaire est projeté sur les axes nord et est et à la fin de l'intégration, il est aisé d'exprimer le déplacement de l'air et la direction en fonction de la somme des composantes nord et est de tous ces vecteurs unitaires.

#### 5.2.2.- La pression atmosphérique

La pression atmosphérique est prise par un senseur de pression M.B. Gilmore *full bridge strain gage*. Lors des premiers mouillages, le senseur employé avait une sortie de 10 m V.F.S. Malheureusement, la dérive en température de l'interface était considérable et donc les données fournies non valides.

Depuis, on a développé une interface de 10 mV et acquis un senseur à sortie de 5 V.F.S. L'interface avec le multiplexeur analogique consiste en un convertisseur "V to P".

#### 5.2.3.- La température de l'air

Le senseur de température, construit à l'unité d'Ostende, est simplement un circuit oscillateur U.J.T. dont l'élément sensible est un thermistor (Fenwal). Le signal de sortie est une suite d'impulsions dont la période varie avec la température.

Comme le signal n'est pas linéaire, une courbe de calibration correspondant à chaque senseur est stockée sur ordinateur et est utilisée par les programmes de transformations des données en valeurs réelles ( $^{\circ}\text{C}$ ) .

#### 5.2.4.- La radiation solaire

La dérive des senseurs de température de l'air observée lors des premiers mouillages était probablement due à la radiation solaire. Pour tester cette hypothèse, un senseur très simple constitué d'une LDR fut monté sur la bouée. Comme ce senseur atteignait rapidement son plateau de saturation et que les expériences scientifiques exigeaient des mesures très précises de la lumière, notamment pour la photosynthèse, il fut remplacé par un albedomètre KIPP dont la cellule photoélectrique mesure la lumière incidente et réfléchie dans un intervalle de longueur d'onde allant de  $3 \times 10^{-7}$  à  $2,5 \times 10^{-6}$  m. Le signal de sortie de 10 m volt F.S. est renforcé par un amplificateur "chopped" et ensuite transformé en un train d'impulsion F.M. par un convertisseur standard.

### 5.3.- Les senseurs océanographiques (tab. 8)

#### 5.3.1.- Les vagues

La hauteur et la période des vagues est mesurée sur la bouée Sysna par un *Datawell Heave Sensor*. La forme et les dimensions de la bouée Sysna sont telles qu'elle peut convenablement suivre les mouvements de la surface de l'eau. Aussi, la mesure des vagues effectuée par le Datawell consiste en une mesure de l'accélération verticale qui est ensuite intégrée 2 fois électroniquement. Les séries continues de mesures prises à une fréquence minimale d'un hertz permettent de calculer aisément le spectre des vagues.

#### 5.3.2.- La température de l'eau

La mesure de la température de l'eau s'effectue avec un senseur identique à celui utilisé pour la température de l'air.

Tableau 8  
Senseurs océanographiques

Parameter	Manufacture	Principle of operation	Range	Accuracy	Output signal	Interface	Remarks
Waves	Datawell Heave Sysna buoy	Twice integrated accelerometer	± 10 m	2.5 % ± 3 cm	DC ± 10 V	Voltage to period conv.	5' integration interrogation frequency 1 Hz minimum
	Baylor wave staff	2 wire rope inductive	50 feet	.5 %	DC 5 V	Idem	Idem as above
Water temperature	Selfmade	UJT oscillator with thermistor	0 - 20 °C	± .1 °C	F.M. pulses	Pulse shaper	Calibration curve necessary
Salinity	Plessey	Measure of seawater conductivity by induction	10 - 40 %	± .03 %	F.M. 4995 - 7900 Hz	Pulse shaper	Intercalibration with a Beckman lab salinometer

### 5.3.3.- La salinité

La salinité est donnée par une mesure de la conductivité de l'eau de mer. Le senseur employé est un conductivimètre Plessey dont le signal de sortie est modulé F.M. et intégré pendant une minute.

## 6.- Alimentation en énergie

Les balises, le système d'acquisition de données, les senseurs et leurs interfaces, l'installation d'émetteurs-récepteurs embarqués doivent être alimentés par un système central. Les tensions nécessaires sont  $\pm 24 \text{ V} \pm 20\%$ , ainsi qu'une série de tensions dérivées. L'énergie nécessaire est fournie par deux aéro-générateurs Aérowatt type 24FP7 (24 V, 1 Ampère). A cause de leur production irrégulière, ces générateurs sont tamponnés dans un ensemble de batteries de  $\pm 24 \text{ V}$ . Ces batteries sont de type cadmium-nickel, et ont une puissance de 185 Ampères/heure (KAP 18).

Une certaine quantité d'Ampères/heure de ces batteries est consommée en fonction des besoins totaux en énergie et compte tenu d'une possibilité d'une période de temps calme pendant plusieurs jours. La tension des batteries peut varier au cours du temps, si bien qu'il faut prendre des précautions contre la surcharge (électrolyse) et la sous-charge (énergie insuffisante pour l'alimentation des balises). Le D.A.S. est alimenté en  $\pm 18 \text{ V}$  via un ensemble de convertisseurs DC/DC. Un schéma général de l'alimentation en énergie est donné (fig. 17).

## 7.- La station de réception à terre (fig. 18)

### 7.1.- La station d'émission-réception

La station de réception située à Ostende démodule les signaux reçus et les décode dans la forme digitale originale, tout en détectant les fautes de transmission.

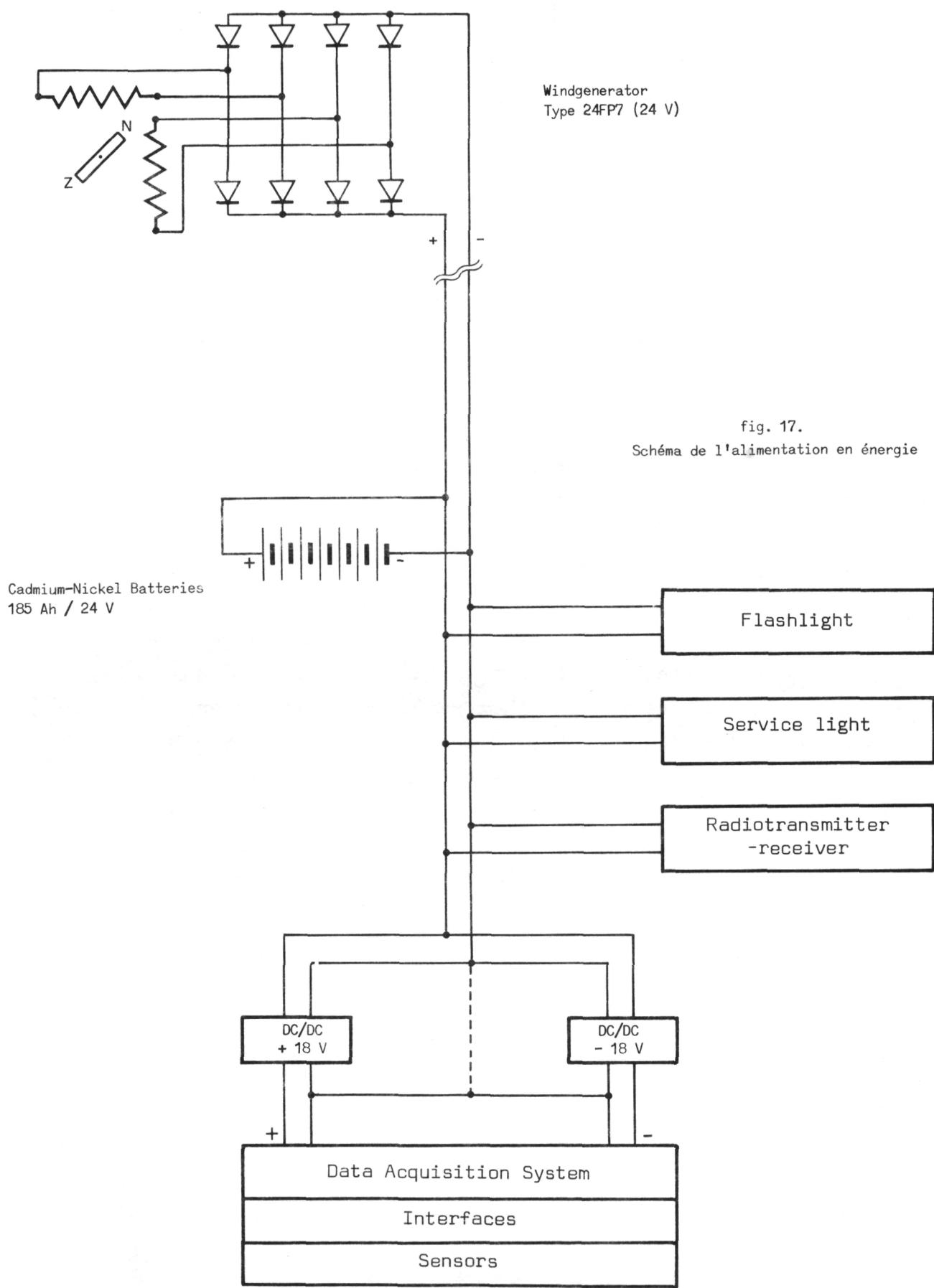


fig. 17.

Schéma de l'alimentation en énergie

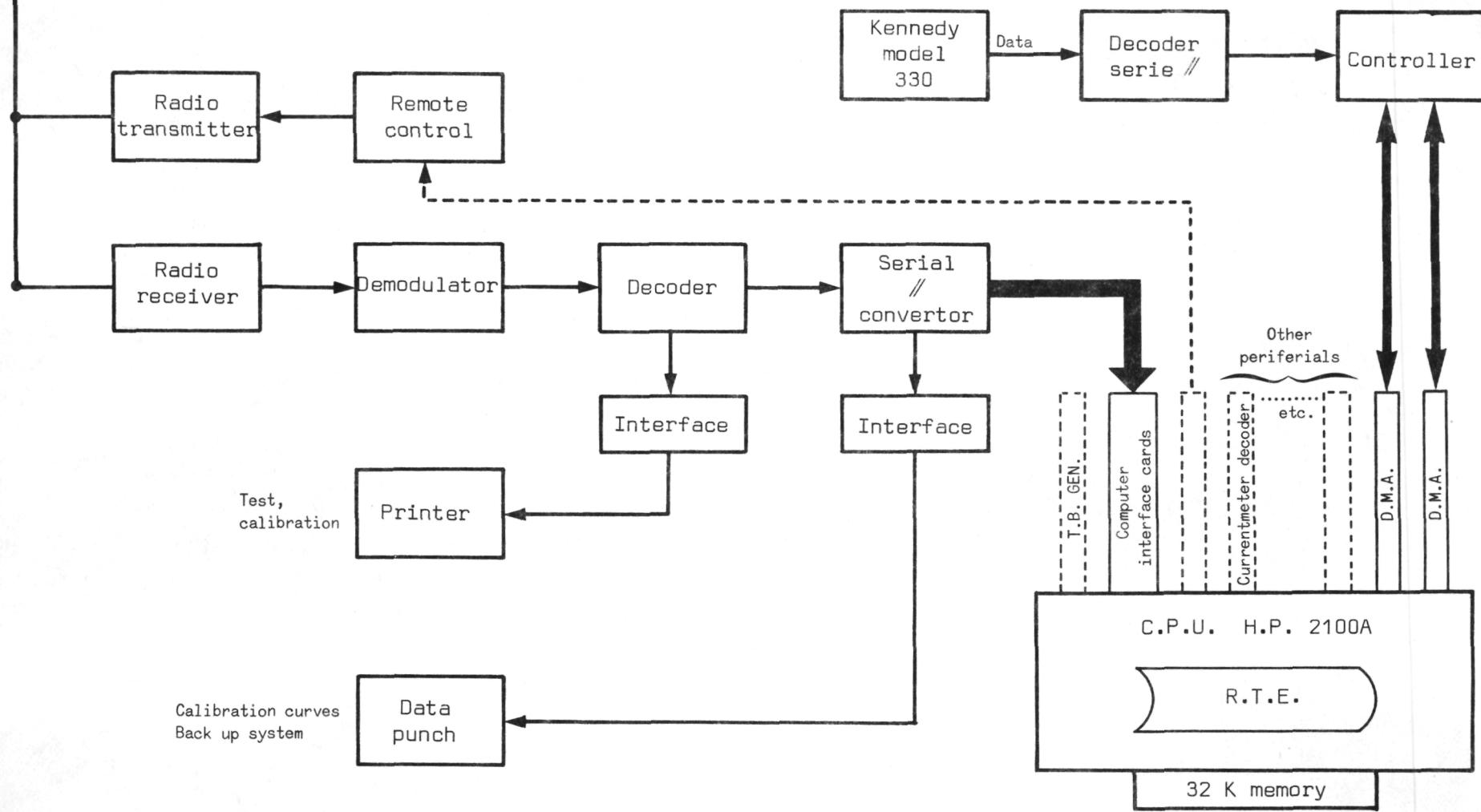


fig. 18.  
Station de réception à terre

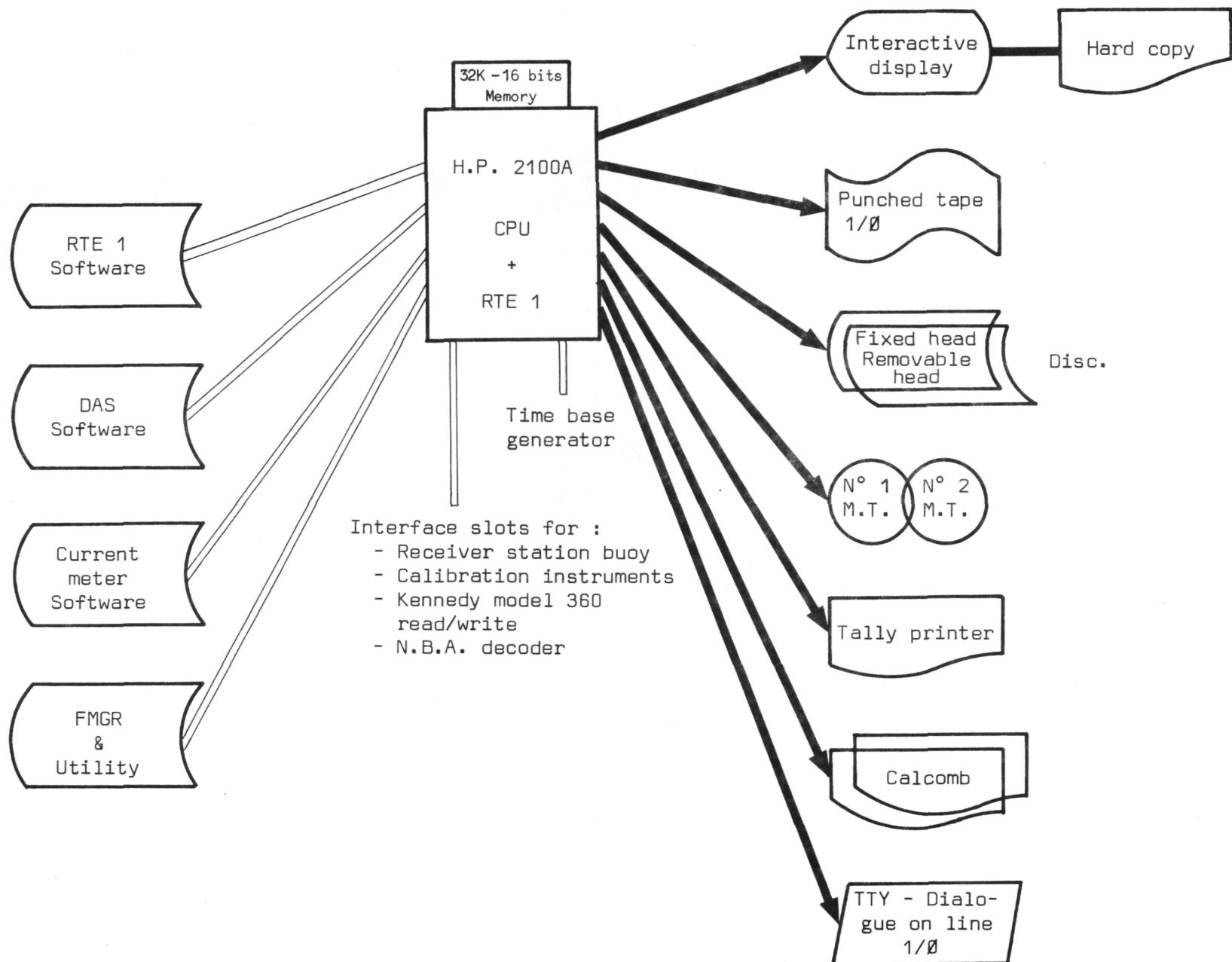


fig. 19.

H.P. 2100A Data processing computer

Après une transformation série-parallèle, les données sont conduites, via une interface standard H.P. à un ordinateur travaillant en temps réel (*Real Time Executive* : R.T.E.) et décrit par la figure 19.

Pour les travaux de tests et de calibration, une imprimante et un perforateur sont disponibles avec interfaces *offline*. Ce perforateur sert également de système de réserve, lorsque le stockage *on line* n'est pas possible (entretiens, défauts *hardware* du R.T.E., manipulations du logiciel, ...).

#### 7.2.- L'interface Kennedy-HP2100A (fig. 20)

Cette interface permet de lire et d'écrire des blocs d'informations sur une cartouche 3M avec un format adapté qui permet d'écrire, dans une bouée, des blocs d'informations avec une grande fiabilité de recouvrement d'erreur lors de la lecture. L'interface est raccordé d'une part avec l'ordinateur H.P. 2100A et d'autre part avec l'enregistreur Kennedy modèle 330. Le raccordement avec l'H.P. se fait au moyen de deux lignes d'entrée-sortie. Chaque ligne possède 16 bits d'informations en entrée et 16 bits d'informations en sortie. Toutes les informations se présentent en parallèles. Les modules H.P. réalisant cette fonction sont du type microcircuit type 12566. Le premier circuit est raccordé sur le bus d'entrée-sortie programmé, c'est-à-dire qu'il faut une instruction spécifique pour entrer ou sortir des données sur ce microcircuit. Le modèle d'entrée-sortie parallèle est employé pour commander le contrôleur Kennedy et pour lire les *status* du Kennedy.

Le second microcircuit est raccordé au D.M.A. Cet organe permet d'entrer ou de sortir des données sans intervention du programme; le D.M.A. est uniquement initialisé par programme. Le dernier microcircuit sert à entrer ou sortir des informations se trouvant ou à écrire sur la bande magnétique.

L'ordinateur H.P. et le programme associé au Kennedy pilotent l'interface et l'enregistreur. Le programme initialise le transfert de données par D.M.A. et le microcircuit associé; ensuite il commande

KENNEDY

H.P. 12566B-001

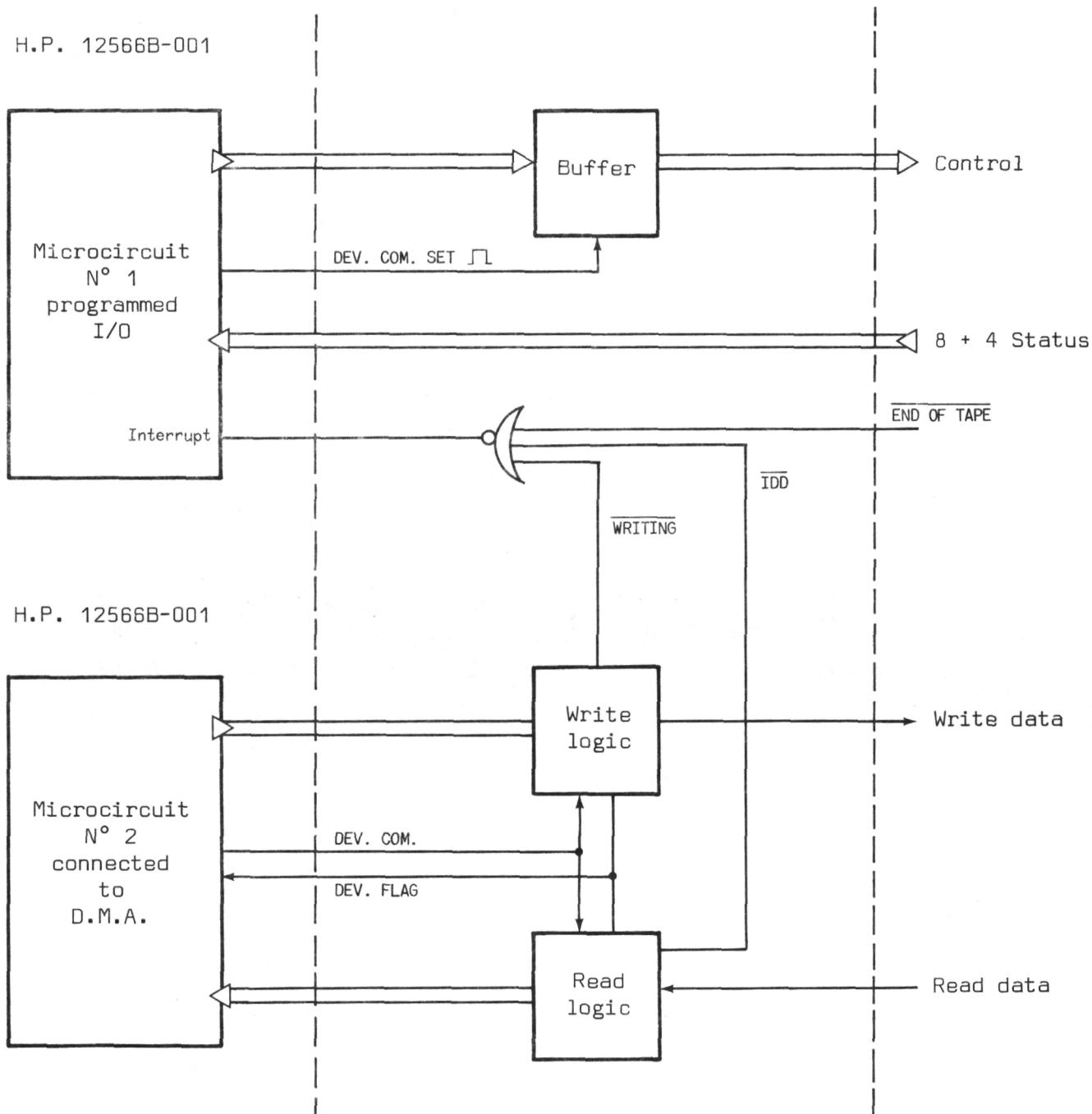


fig. 20.  
Interface Kennedy - H.P. 2100A

le mouvement du moteur. L'interface se charge de la mise en forme des données et de leur transfert vers le microcircuit. Lors de la fin d'écriture ou de la lecture, une interruption avertit l'ordinateur de la fin de transfert.

#### 7.3.- L'installation R.T.E. de traitement des données

L'installation R.T.E. de traitement des données comprend les périphériques standards (disques fixes et amovibles, bandes magnétiques, téléimprimantes, perforateur et lecteur de rubans, imprimante par lignes, tables traçantes Calcomp de 11 et 31 pouces, etc.) ainsi que différents interfaces à buts spécifiques (décodage des bandes magnétiques des courantomètres, lecture du Kennedy, différentes calibrations, etc.).

Les principales phases du traitement sont :

- le transcodage des données binaires reçues,
- la mise en ordre des *digits*, des nombres et enfin des cycles de mesures,
- la répartition des données arrivant de différentes stations,
- la synchronisation et l'introduction ou la correction de l'échelle de temps,
- la transformation des nombres reçus en paramètres physiques au moyen des équations de calibration,
- l'inspection des données : listes, aperçu sur video, dessin automatique, méthode statistique.
- l'édition, après détection automatique ou manuelle des erreurs,
- concentration des données : moyennes pondérées, filtres numériques, méthodes statistiques, corrélations, spectres de puissance, etc.

Il est très important que la plupart de ces opérations puissent être menées *on line* et que les autres soient faites aussitôt que possible après la réception des données pour d'abord détecter tout mauvais fonctionnement des instruments et ensuite rendre les données disponibles à un usage direct, aisément et efficacement dans les modèles mathématiques.

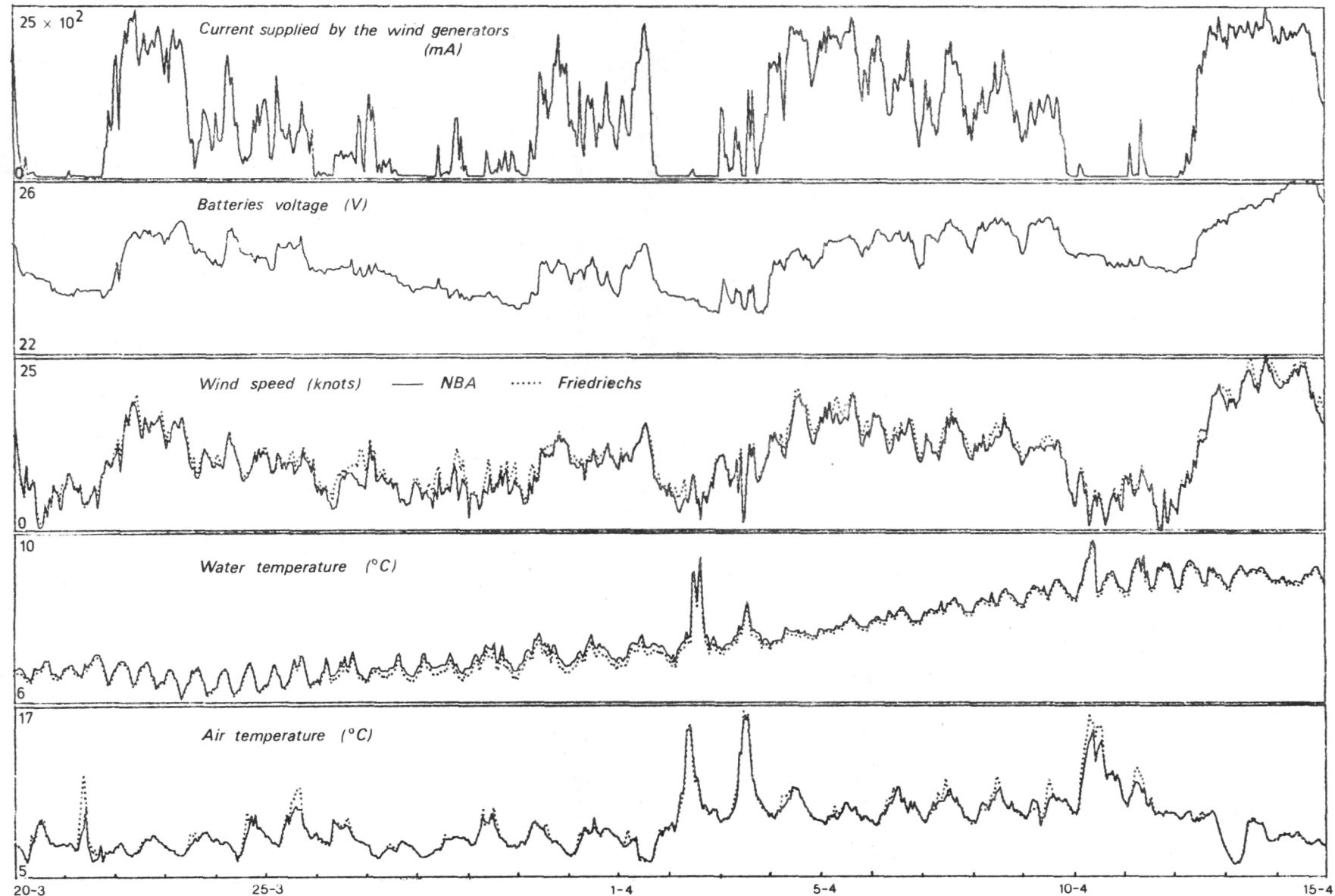
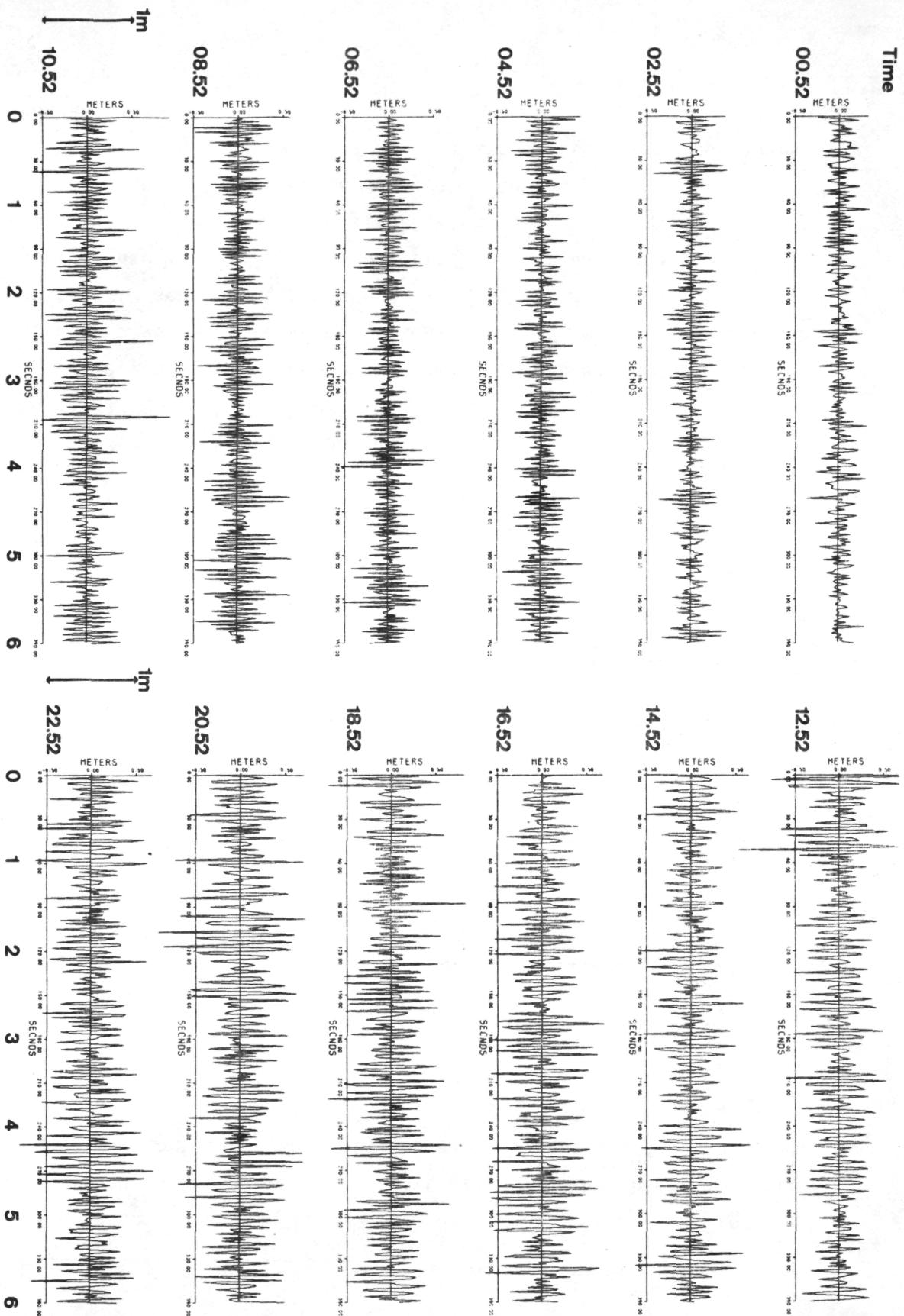


fig. 21a.

Sysna I Mooring 19-3-1974 15-4-1974  
Position 51 $^{\circ}$ 20'45" N 20 $^{\circ}$ 53'40" E

Time



Data well  
WAVE DATA

10 min continuous interrogation at 1 s interval

fig. 21b.

Sysna I Mooring 19-3-1974 15-4-1974  
Position 51°20'45" N 20°53'40" E

#### 8.- Bilan des mouillages effectués

Le tableau 9 résume les mouillages effectués ainsi que les défauts, pannes ou accidents qui se sont produits. Le tableau 10 donne, mouillage par mouillage, la liste des senseurs qui ont fonctionné et des données qu'ils ont fournies.

Tableau 9  
Mouillages effectués par la bouée Sysna

Databuoy	Mooring period	Position	Principle cause of failure
Sysna I	19-02-1974 15-03-1974	51°20'45" N 02°53'40" E	Anchoring system (Electrolytic corrosion)
Sysna II	24-02-1975 18-03-1975	51°21'30" N 02°55'14" E	Some sensors not working Collision with unidentified ship
Sysna III	22-08-1975 02-09-1975	51°21'25" N 03°02'35" E	Wind generator defective causing troubles in power system
	19-09-1975 10-10-1975	51°21'25" N 03°02'35" E	Defect in D.A.S. followed by failing anchoring system

Il faut noter que ces opérations ont été planifiées, sans tenir compte de la période de l'année et d'éventuelles mauvaises conditions météorologiques.

##### 8.1.- Mouillage Sysna I

Le premier mouillage eut lieu du 19 mars au 15 avril 1974 à la position 51° 20' 45" N , 2° 53' 40" E.

Outre une série de données de service, ont été mesurées les grandeurs suivantes : température de l'air à 8 mètres au-dessus de la surface (2 senseurs), vitesse du vent à 10 mètres au-dessus de l'eau (2 senseurs différents : un Friedrich et un Munro), température de l'eau à un mètre sous la surface (2 senseurs) et vagues chaque fois suivies toutes les secondes pendant 10 minutes.

Une rupture d'ancre mit fin à l'expérience après 26 jours de bonnes opérations.

Tableau 10

Bilan du fonctionnement des senseurs lors des mouillages de la bouée Sysna

Housekeeping : aerogenerator current, battery voltage, day of mooring, elapsed time clock, transmission number, integration time, buoy #, flood alarm, buoy temperature, etc.			
---	--	--	--

Sysna I parameters	Sensors	Data recovered	Remarks
Air temperature	2 self made	26 days	Some difference between the two sensors due to solar energie
Water temperature	2 self made	26 days	
Wind speed	Friedrich Munro	26 days 26 days	10' integrated values 10' integrated values
Waves	Datawell	26 days	10' continuous interrogation at 1 s interval

Sysna II parameters	Sensors	Data recovered	Remarks
Conductivity	Plessey	23 days	
Air temperature	2 self made	23 days	
Water temperature	2 self made	1 : 3 days 2 : 23 days	sensor 2 : right calibration coefficient have to be fitted after 11 days
Wind speed	Friedrich Munro	23 days 23 days	10' integrated values " and 10' vector averaging
Wind direction (compass)	Munro NBA	23 days 23 days	10' vector averaging for mean value For wind direction angle only
Waves	Datawell	-	Error in cable system

Sysna III parameters	Sensors	Data recovered	Remarks
Conductivity	Plessey	28 days	
Air temperature	2 self made	28 days	
Water temperature	2 self made	28 days	Some second quality data due to computer break down at the receiver end. Ventilation of computer room insufficient
Wind speed	Friedrich Munro	28 days 28 days	Idem as above "
Wind direction (compass)	Munro NBA	28 days 28 days	" "
Waves	Datawell	28 days	Good data
Barometric pressure	M.B. Gilmore	-	Prototype interface, mass problems and temperature drift

Les données (fig. 21 a, b), transmises à la station côtière, ont été stockées sur ruban perforé et ensuite traitées *off line*.

#### 8.2.- Mouillage Sysna II

Le second mouillage eut lieu du 24 février au 18 mars 1975 à la position  $51^{\circ} 21' 30''$  N,  $2^{\circ} 55' 40''$  E. Furent mesurées la conductivité, la température de l'air et de l'eau, la vitesse du vent intégrée sur 10 minutes (Friedrich et Munro), ainsi que la moyenne vectorielle de la direction et de la vitesse du vent sur 10 minutes.

Des difficultés dues au mauvais câblage des senseurs de vagues et de température de l'eau nécessitèrent plusieurs interventions et le changement de ces senseurs en mer.

Malheureusement, cette bouée a été accrochée par un bateau non identifié, le 18 mars 1975 entre 13 h 09 et 15 h 59. Le corps de la bouée accidentée fut ramené à Ostende par le Service du Pilotage tandis que son mât, coulé, fut retrouvé et récupéré par une unité de la Force Navale équipée d'un sonar.

Les données (fig. 22) furent décodées à la station de réception et une fois encore, stockées sur bandes perforées.

#### 8.3.- Mouillage Sysna III

Comme la première bouée était immobilisée, pour un long moment, par les travaux d'expertise et de réparation relatifs à l'accident, il fut décidé d'équiper la seconde bouée, livrée entretemps, avec le même ensemble de senseurs auxquels s'ajoutait celui de la pression atmosphérique dont l'interface était disponible.

Cette bouée fut mouillée le 22 août 1975 à la position  $51^{\circ} 21' 25''$  N,  $3^{\circ} 02' 35''$  E. Un léger déséquilibrage des pales d'un aérogénérateur provoqua par vibration un court-circuit dans le circuit de balisage et par conséquent un épuisement des batteries.

Comme cette panne était difficilement réparable en mer, la bouée fut ramenée à Ostende le 2 septembre 1975 et remise en service le 19,

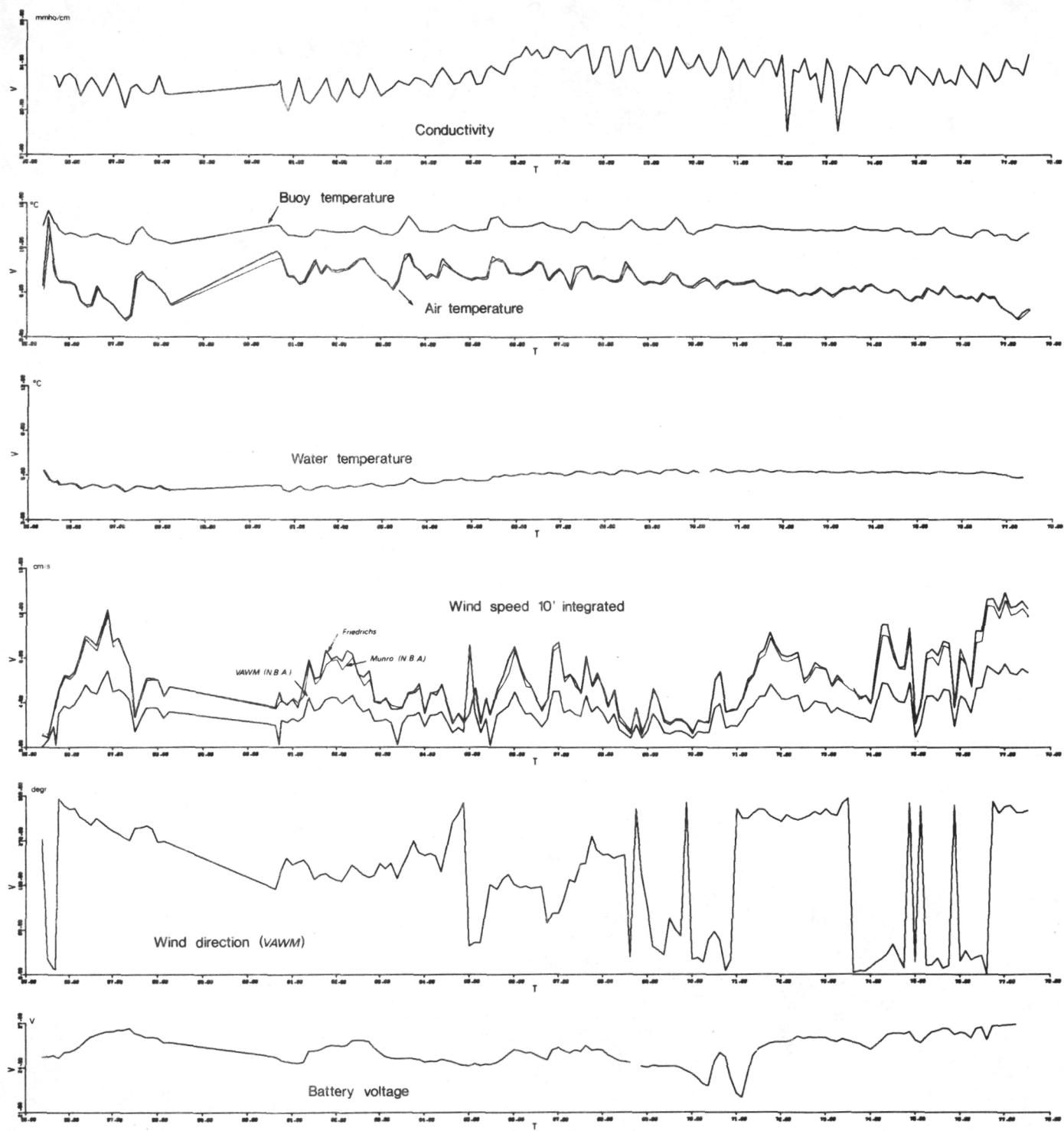
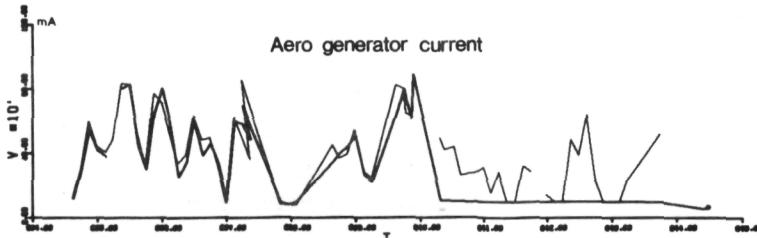
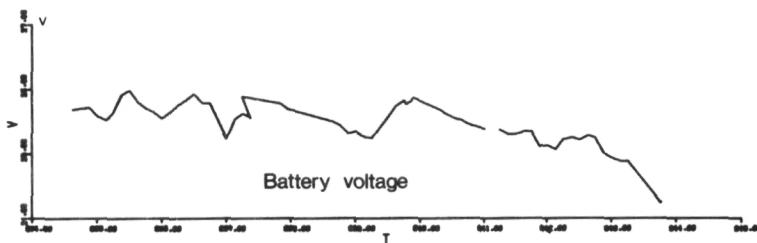
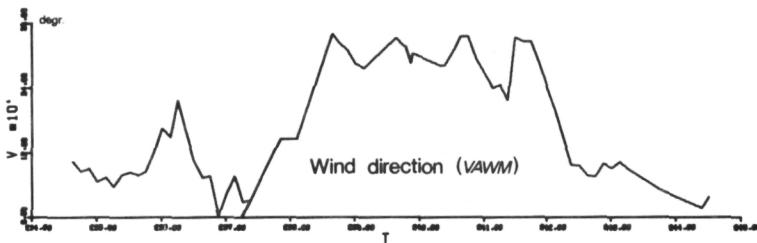
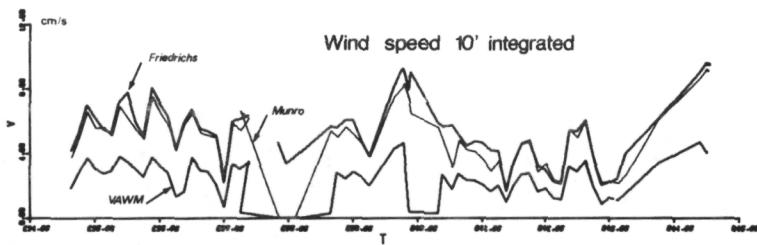
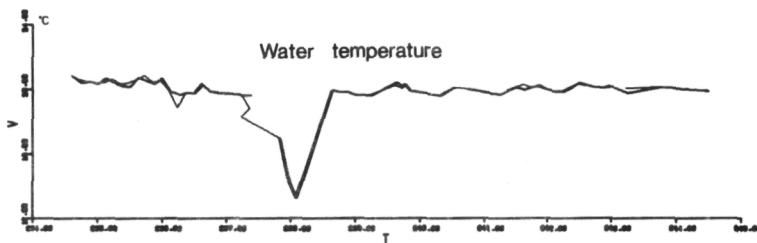
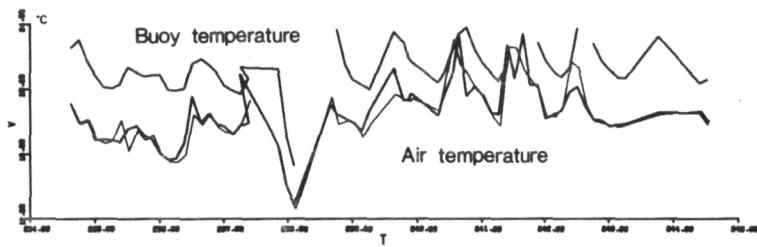
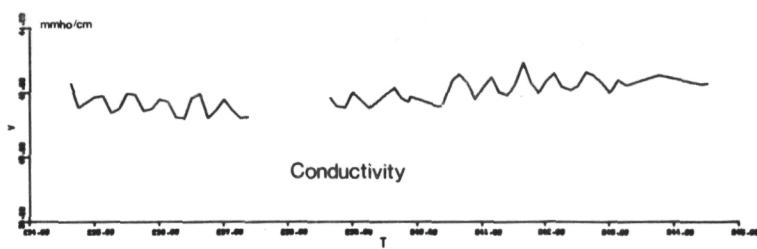


fig. 22.

Sysna II Mooring 24-2-1975 18-3-1975  
Position 51°21'30" N 2055'40" E



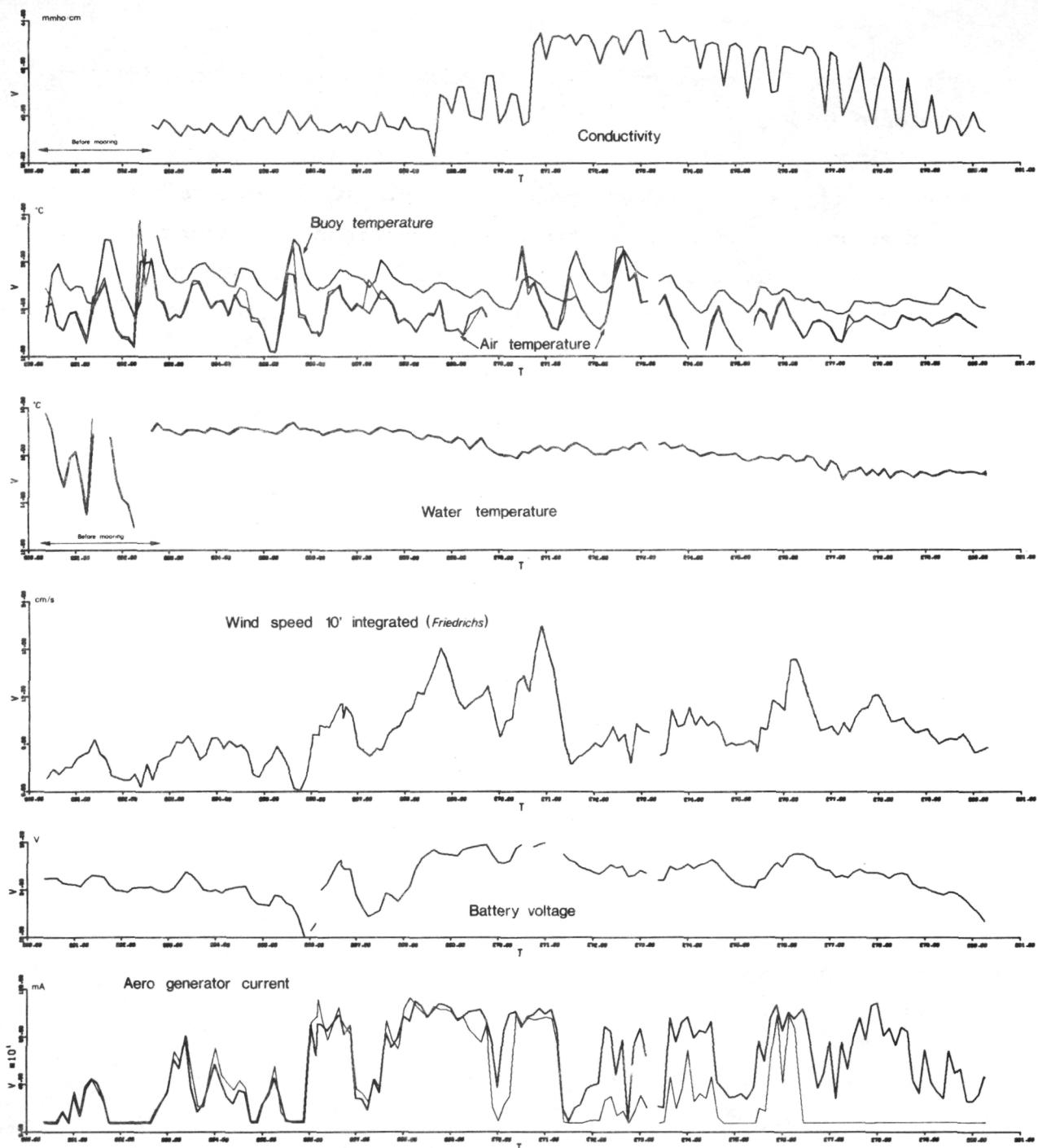


fig. 23b.  
Sysna III Mooring 19-9-1975 10-10-1975

fig. 23a.

Sysna III Mooring 22-8-1975 2-9-1975



après une inspection générale du D.A.S. Elle fonctionna ensuite jusqu'au 10 octobre 1975.

L'important à souligner dans ce mouillage réparti en deux périodes, est que les données transmises (fig. 23a,b) furent décodées à la station de réception et introduites directement dans l'ordinateur. *Tout le système fonctionna donc en temps réel.* Toutefois quelques pannes survenues tant dans le *hardware* que dans le logiciel justifièrent pleinement le stockage de sécurité sur ruban perforé.

## Chapitre III

### La station fixe

Les senseurs, actuellement opérationnels sur bouées, ont principalement trait à la météorologie marine et à l'océanographie physique. Cependant, un modèle mathématique décrivant et prédisant l'évolution de l'environnement marin requiert en outre l'acquisition automatique de grandeurs chimiques et biologiques comme la concentration des nutriments dissous, la biomasse du plancton, la production primaire, la teneur en métaux lourds, etc.

Des moyens élégants pour mesurer ces grandeurs ont été développés dans le cadre du Programme, soit sous forme de senseurs intelligents capables de fournir un signal électrique soit comme échantilleurs programmés (par exemple, en fonction des conditions météorologiques).

Toutefois ces prototypes ne sont pas encore suffisamment achevés pour pouvoir être montés sur bouées. C'est principalement pour les développer sans limitation de place ou d'énergie que l'implantation d'une station fixe a été envisagée (fig. 24 et 25).

Celle-ci, prête en juin 1975, n'a pu encore être mise en place à  $51^{\circ} 21' 25''$  N  $3^{\circ} 03'$  E à cause de conditions météorologiques défavorables, mais devrait l'être dans les premiers mois de 1976.

Un tuyau Mannesman de 35 mètres de long, de 1,42 mètres de diamètre, de 16,5 mm d'épaisseur et pesant 20 tonnes est surmonté par une cabine octogonale de 3 mètres de diamètre, dont les parois, de sandwich polyester-polymétre armé ont 50 mm d'épaisseur.

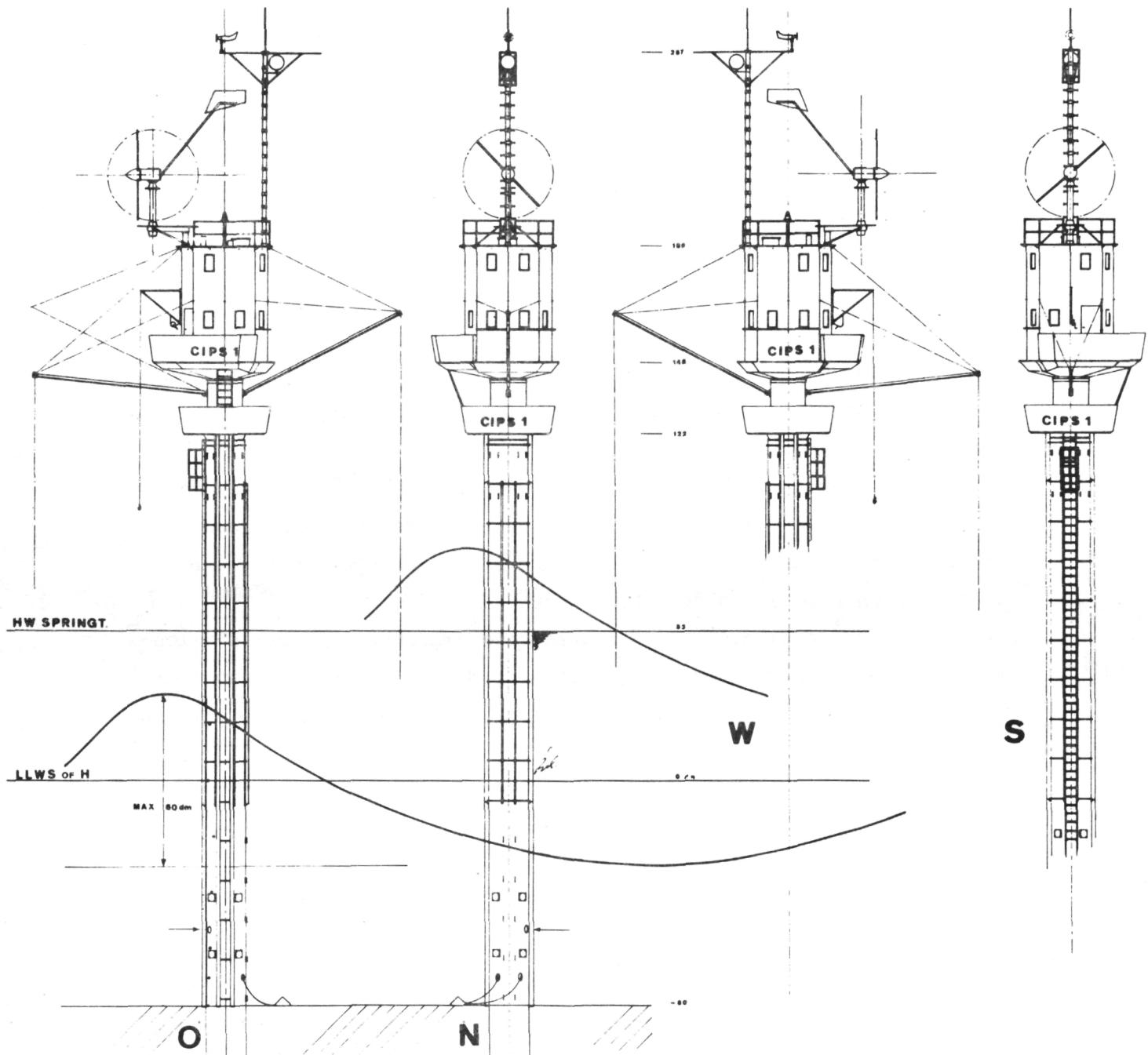


fig. 24.  
La station fixe

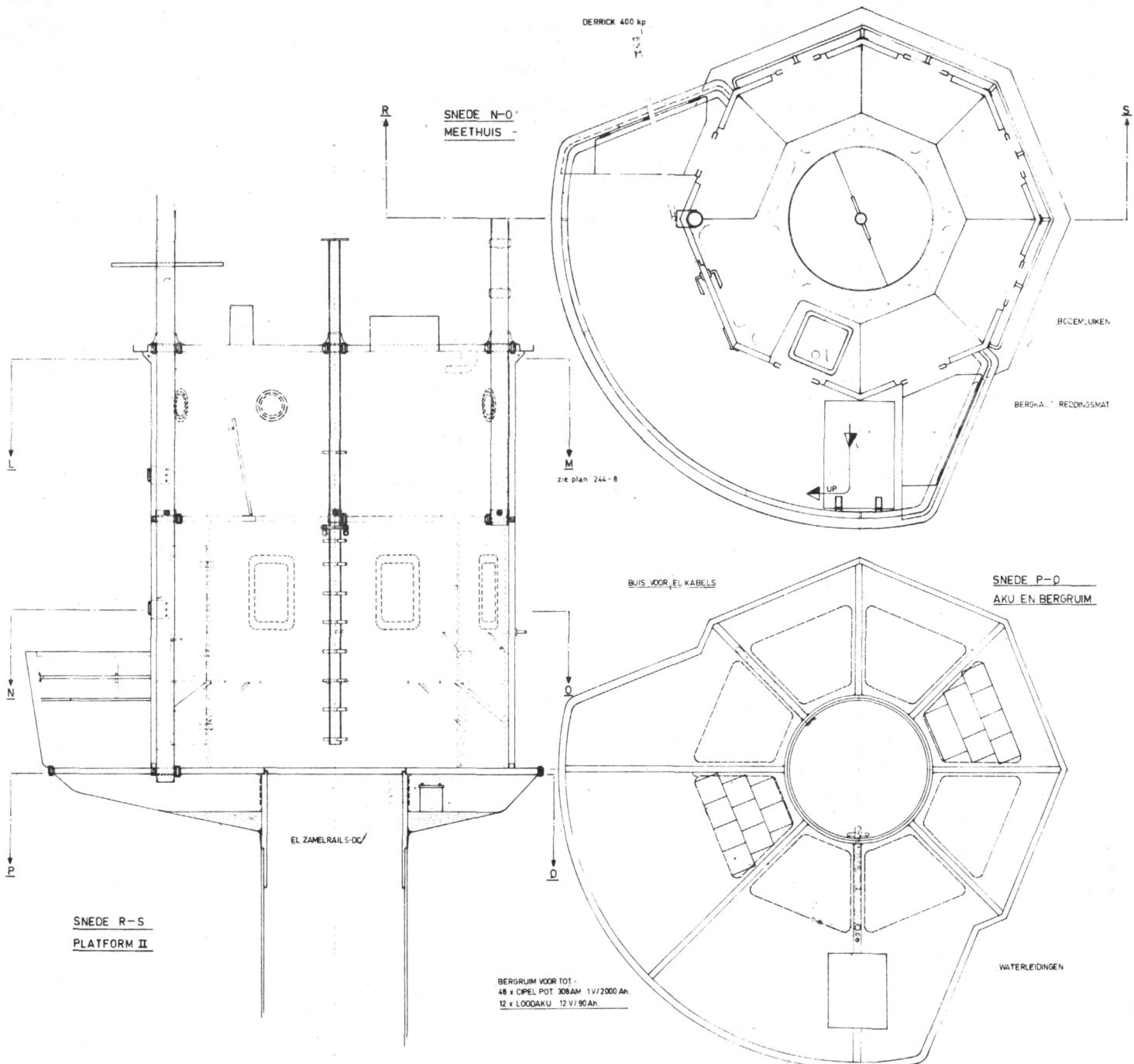


fig. 25.  
La station fixe - Plan du laboratoire

Le hauteur d'eau à l'endroit prévu est de 8 m , aux marées basses de vives eaux tandis qu'il y a une distance minimum de 10 m entre la surface de l'eau et le laboratoire.

Un premier balcon servira principalement d'entrepôt pour les batteries, tandis qu'un second, 2 mètres plus bas, permettra la manipulation des treuils et l'exécution des mesures.

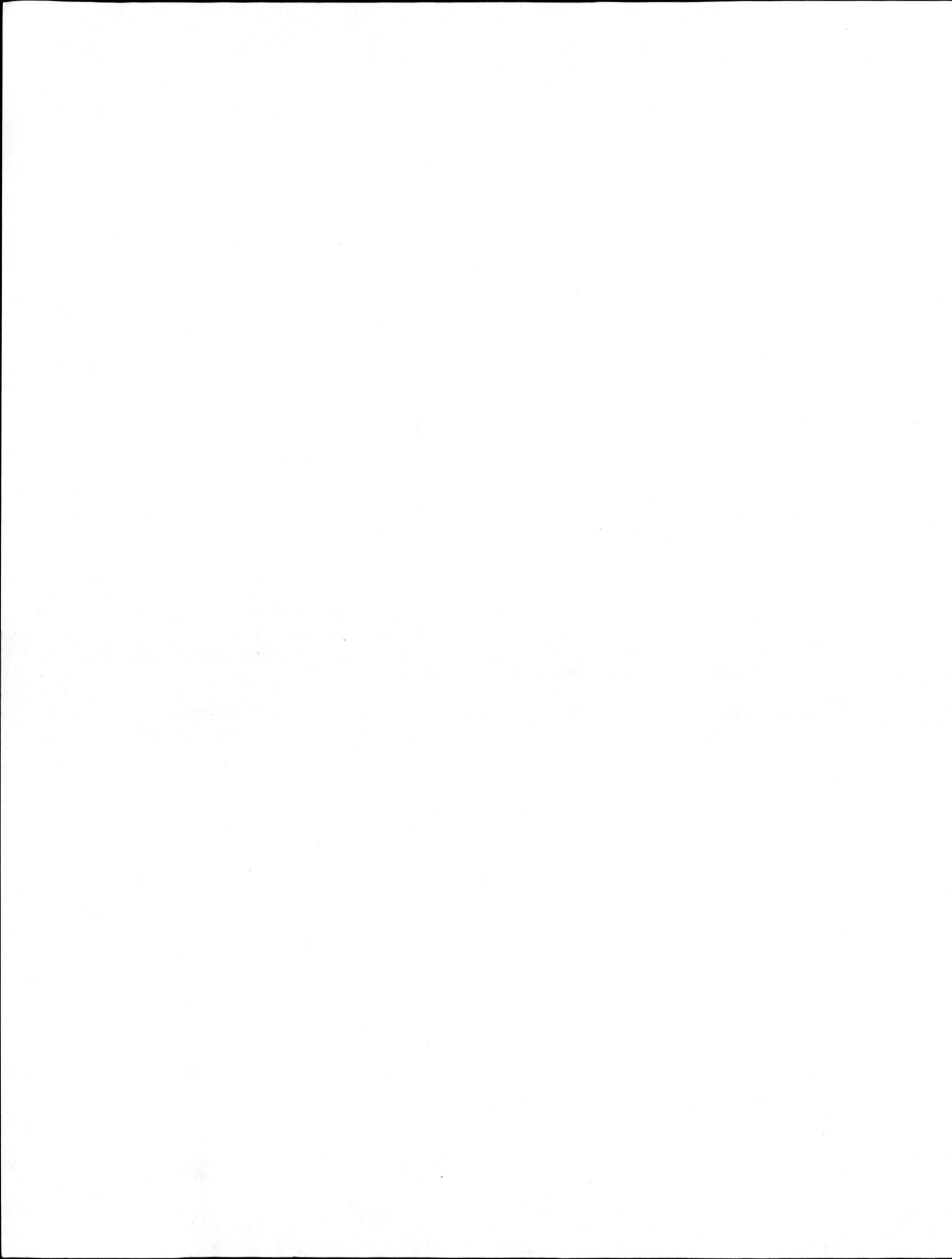
Cette station sera équipée de quatre supports de senseurs mobiles (deux à l'intérieur, deux à l'extérieur du tuyau), de tuyaux d'adduction d'eau et de protection de câbles, d'échelles intérieure et extérieure, d'un groupe alternateur diesel de 33 kVA, d'une installation basse tension, de sanitaires, d'équipements de sécurité, etc.

Elle ne causera aucun danger à la navigation car elle portera une signalisation optique (activée par une cellule photoélectrique) et acoustique (mise en action par un signal donné par la station côtière).

Les chercheurs pourront y séjourner pendant les quelques heures nécessaires à la mise en place, l'entretien, la calibration ou la récupération des senseurs qu'ils développent.

## Références

- PICARD, H., (1974). Calibration des compas des courantomètres, Technical Report 1974/Instrumentation 01, Programme National sur l'Environnement Physique et Biologique, Projet Mer.
- PICHOT, G., DE HAEN, A. and NIHOUL, J.C.J., (1974). The Belgian automatic oceanographic and meteorological data station, *in Proc. of the First European Symposium on Offshore Data Acquisition Systems*, Southampton (U.K.), Sept. 1974.
- PICHOT, G., (1975). Automatic acquisition of meteorological and oceanographic data : further developments and first results, Rapport de Synthèse IV, 159-172, Programme National sur l'Environnement Physique et Biologique, Projet Mer.
- POLLENTIER, A., VANDENBOSSCHE, M. en RIGOLE, F., (1974). Vektorieel integrerende windmeter voor meteo-oceanografische boeien, Technical Report 1974/Instrumental 02, Programme National sur l'Environnement Physique et Biologique, Projet Mer.
- POLLENTIER, A., (1976). Project Stroommetingen : September 1973 - December 1975, Report N, Programme National sur l'Environnement Physique et Biologique, Projet Mer.
- PRANGSMA, G.J., (1975). Scientific report on the automatic handling of data obtained from recording current meters, K.N.M.I., De Bilt.
- RAMSTER, J.W. and HILL, H.W., (1974). A first assessment of two long-term pilot current meter networks in the North Sea, *in Proc. of the First European Symposium on Offshore Data Acquisition Systems*, Southampton (U.K.), Sept. 1974.
- TALBOT, J.W. and BAXTER, J.C., (1975). The direction calibration of Plessey current recording meters, Fisheries Laboratory Technical Report Series M.A.F.F. n° 12.



## LA BANQUE DE DONNEES

par

Yves ADAM, Huguette LAVAL, Pol CLOSSET, Jean-Pierre FOGUENNE, Walter KEUTGEN



## Chapitre I

### La gestion des données océanographiques

La tendance actuelle en océanographie est à la construction de modèles mathématiques pour représenter les phénomènes physiques, chimiques et biologiques. Pour être validés et initialisés, ces modèles demandent la connaissance d'une foule de données expérimentales, soit prises *in situ* soit déterminées en laboratoires, qui doivent être traitées, analysées et injectées ensuite dans les programmes de simulation.

A cause de l'énorme quantité de données nécessaires, la modélisation des systèmes marins implique deux obligations :

- a) la réalisation d'expériences de grande envergure, étalées dans le temps et dans l'espace, qui fournissent des données cohérentes entre elles;
- b) le traitement automatique des données.

Ce dernier fait l'objet du présent travail de recherche.

La seule manière d'établir un compromis entre le grand nombre de données disponibles et l'accessibilité à ces informations, est d'établir une base de données structurée selon le modèle mathématique et de développer un ensemble de programmes de gestion et d'analyse qui en rendent l'accès aisés. Les variables à étudier, à mesurer, à analyser sont de nature très diverse; ce peuvent être en effet :

- a) des séries temporelles, régulières ou non, de paramètres physiques ou chimiques;
- b) de séries de paramètres variant en fonction de la profondeur;

c) des séries de paramètres dont les valeurs sont mesurées aux points d'une grille régulière couvrant la zone à étudier.

Les paramètres sont eux-mêmes très différents les uns des autres, tant dans leur signification physique que dans leur méthode de mesure. Ils peuvent être engendrés par des instruments automatiques ou non, venir de différents laboratoires; une série peut n'avoir de signification que comparée à une autre.

Il est nécessaire, d'abord, de définir un *format de stockage* des données suffisamment souple pour faire face au problème de la diversité des données; ensuite, de développer des programmes généraux pour mettre en forme, écrire, retrouver et sélectionner les différents types de données, chaque programme étant construit de façon à être facilement exploitable par l'utilisateur.

La plupart du temps, et plus particulièrement lors d'expériences regroupant de nombreuses institutions, ces données doivent être échangées entre différents laboratoires, sur des supports digitaux susceptibles d'être traités automatiquement par les systèmes informatiques dont disposent les centres de recherche. Etant donné l'énorme variété des données à traiter, ainsi que la grande diversité de conception des différents systèmes de traitement et d'analyse de l'information, il est exclu de développer des logiciels particuliers à chaque type de données, et compatibles seulement avec quelques systèmes. Il faut donc définir un *format d'échange* de données sur support digital (et de préférence sur bande magnétique), autodescriptif (c'est-à-dire qui contient toutes les informations nécessaires pour définir les données) et susceptible d'accepter n'importe quel type de données. Il faut de plus que ce format d'échange soit compatible avec la structure de la majorité des ordinateurs en fonction dans les centres océanographiques. Il faut enfin développer un logiciel qui lise et écrive ces structures et qui soit portable d'ordinateur à ordinateur. On va s'attacher, dans les chapitres suivants, à expliquer les raisons qui ont conduit à l'élaboration de tels formats et à décrire sommairement leur structure.

## Chapitre II

### Construction de la base de données

#### 1.- Etat présent de la base de données

• Ce chapitre est consacré à la description et au développement de la base de données. On va mettre l'accent sur les difficultés qui ont été rencontrées lors de la construction de ce système d'information. Les scientifiques impliqués dans des problèmes de traitement et d'exploitation de données ont été, jusqu'ici, bien trop occupés par d'autres recherches fondamentales (développement de modèles mathématiques, création d'algorithmes numériques) ou par des questions de technologie (réalisation de systèmes d'acquisition de données) pour pouvoir consacrer leur temps à l'élaboration d'une banque de données.

Toutefois, il existe déjà plusieurs fichiers de données aisément accessibles sur des supports lisibles par ordinateurs (le plus souvent des bandes magnétiques) et un logiciel élémentaire capable de stocker, de lire, de classer et de visualiser ces données. Le logiciel n'a pas d'unité et consiste en diverses chaînes de plusieurs programmes, chacune étant spécialisée dans le traitement d'un type particulier de données. Jusqu'à présent, les outils mathématiques développés pour un type de données ne peuvent être immédiatement adaptés à un autre type. De plus, les programmes ont été réalisés pour des ordinateurs de types particuliers, mais des travaux sont actuellement en cours pour assurer la portabilité du logiciel et des données (pour rappel, deux types d'ordinateurs sont disponibles dans le cadre du projet pour l'exploitation des données : un HP 2100A et un IBM 370/158).

Les données proviennent essentiellement de deux types de sources :

- les stations d'acquisition automatiques opérant dans le cadre du Programme National sur l'Environnement Physique et Biologique;
- les échanges avec d'autres institutions coopérant avec les chercheurs du Programme dans le cadre d'expériences communes.

## 2.- Données issues du Programme

Les chaînes d'acquisition automatique des données concernent :

- a) les données des bouées,
- b) les données des courantomètres,
- c) les données météorologiques.

Toutes sont des séries temporelles (données toutes les 15 minutes pour les courantomètres, toutes les heures pour la bouée, toutes les 3 heures pour la météorologie).

La première chaîne a été installée à Ostende, les deux autres à Liège.

### 2.1.- Données météorologiques

La chaîne traitant les données météorologiques est la plus sophistiquée de toutes, à cause de la structure complexe des séries temporelles (28 échantillons et 14 variables à chaque pas de temps).

Cette chaîne traite les données météorologiques qui proviennent en temps réel de la Régie des Voies Aériennes (RDVA) grâce à un récepteur telex. Les informations météorologiques codées (exemple fig. 1) sont enregistrées telles quelles sur ruban perforé par la perforatrice automatique associée au récepteur. Un certain nombre de stations météorologiques dont les émissions sont régulières ont été sélectionnées; elles forment un réseau relativement dense couvrant le Southern Bight et ses côtes (fig. 2); elles comptent notamment des stations terrestres dont la validité des informations en ce qui concerne la représentation de la situation en mer n'est pas toujours certaine, notamment pour les données relatives au vent.

vents geostrophiques

du 26 / 1 / 1976 a 0900 z

a = 327 / 26 nds n = 4  
b = 270 / 19 nds n = 10  
c = 335 / 25 nds n = 11  
d = 291 / 8 nds n = 16  
e = 316 / 17 nds n = 11

sauk22 egrr 260900  
03898 63221 98022 40607 13253 145//=

san140 ehwx 260929  
ehvb 20004 4500 1cu025 3ci180 m01/m03 1012  
grn wht tempo grn=  
sin123 ehdb 260900  
06220 83006 98261 11104 7937/ 01713 30201 35502 4068//  
sin121 ehdb 260900  
06235 51906 65151 10701 39462 52810 81708 83916=

sa ehts not available  
sin123 ehdb 260900  
06300 53115 98272 13304 49360 01801 30402 00/00 4073//  
sin122 ehdb 260900  
06310 12109 57050 13300 18400 51303 81815=

sa ehgo not available

san1 eham 260925=  
ehrd19005 5000 1cu022 3ci250 m01/m01 tempo 2000 9//005=

sabx50 ebwm 260925  
ebfn 26002 9999 3cu020 01/m01 1014  
black wht wht inter amb=

sa ebwh not available  
mmxx 2609  
06407 13507 70010 13902 12500 52105 81825=  
mmxx 2609  
06408 30204 98027 14603 32400 51205 91838=  
sifr22 lfpw 260900  
07002 83412 66022 15000 885// 51312 91827=

sa lfac not available

sa lfaj not available

+  
41746 enviro b  
25757 enviro b

fig. 1.  
Exemple de messages codés produits par la R.D.V.A.

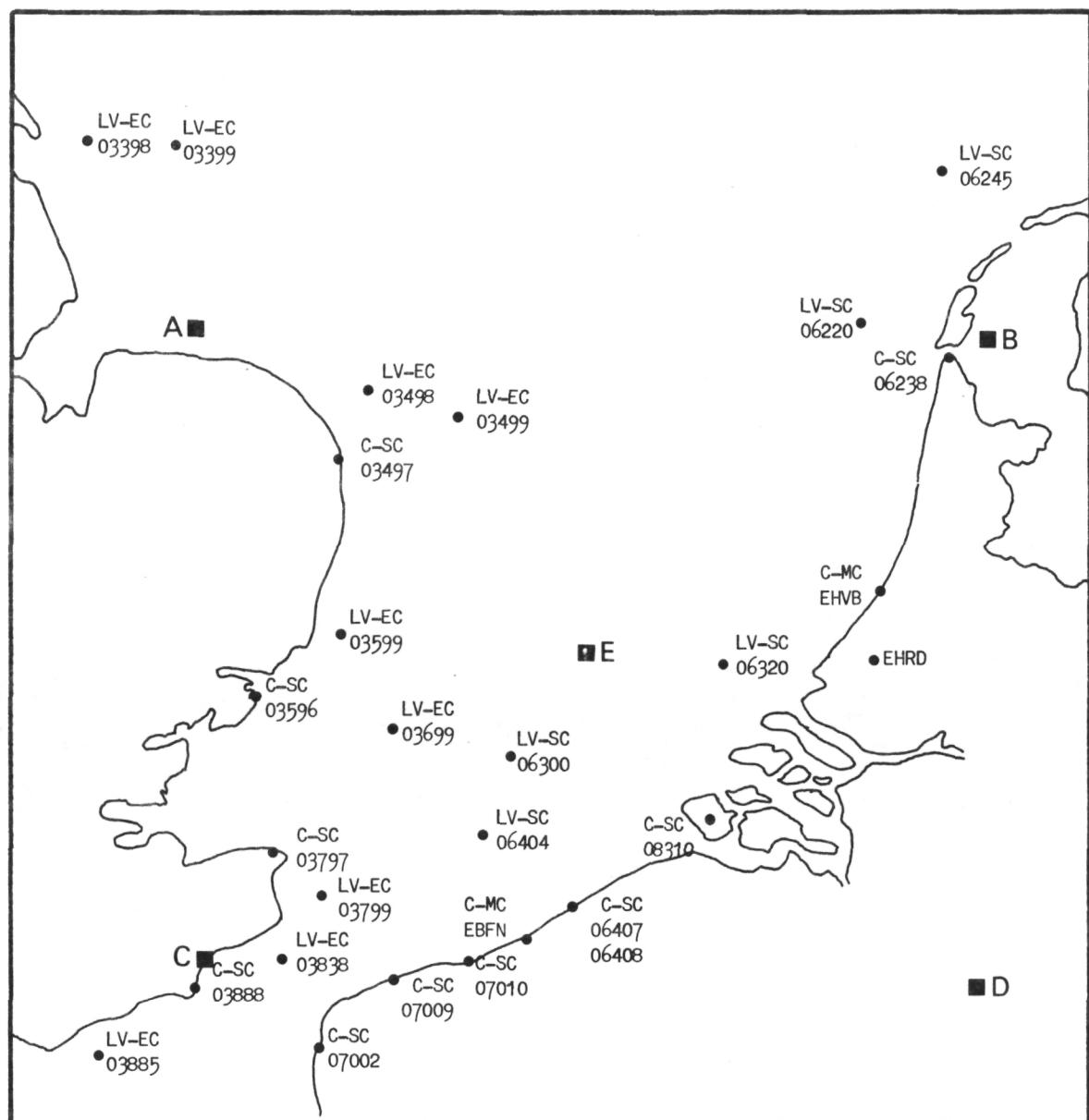


fig. 2.  
Position des stations météorologiques dans le sud de la mer du Nord

La plupart des stations émettent toutes les 3 heures, mais certaines sont plus irrégulières. La liste des paramètres mesurés en ces stations et retenus dans les fichiers est donnée au tableau 1. En plus

Tableau 1

Liste des données synoptiques stockées dans la base de données

Ce tableau montre les divers types de données stockées par le second programme de la chaîne de traitement. Quelques stations fournissent les informations en code "anglais", d'autres en code "Synop" ou "Metar". L'ensemble des informations varie d'après le type de code. Dans le but d'enregistrer les données sous un format commun, les informations inexistantes dans certains codes sont remplacées par une valeur fictive 9999. Quelques stations, bien qu'utilisant un code donné, n'envoient pas un ensemble complet d'informations : les astérisques dans le tableau indiquent les informations qui manquent dans les messages de certaines stations. Une croix indique une information toujours présente.

Type de données	Unités	Code anglais	Code Synop	Code Metar
Nom de la station	-	x	x	x
Visibilité horizontale	mètres	x	x	x
Hauteur	demi-mètres	*	9999	9999
Période } des vagues	secondes	*	9999	9999
Direction de propagation	dizaines de degrés	x	9999	9999
Température } de l'air	°C	x	x	*
Température } de la mer	°C	x	9999	9999
Température } de rosée	°C	9999	x	*
Pression barométrique	dixièmes de mb	*	x	*
Vitesse } du vent	noeuds	x	x	x
Direction	dizaines de degrés	x	x	x
Temps } présent	code	x	x	9999
Temps } passé	code	x	x	9999
Tendance barométrique	code	9999	x	9999
Caractéristique de la tendance barométrique	code	9999	x	9999

de mesures, la R.D.V.A. fournit les vents géostrophiques calculés et prédits (toutes les 6 heures) en 5 points : A,B,C,D,E sur la figure 2. Le contenu des rubans perforés (support d'information primaire) est transféré sur une bande magnétique; chaque fois que le volume d'informations est suffisant (tous les quinze jours) ou que le besoin en est exprimé (en cas d'urgence), un premier programme de la chaîne

- a) traduit le code SIEMENS du ruban en code EBCDIC et transfère le résultat de ce décodage sur une autre bande magnétique;
- b) optionnellement produit une liste des codes transmis.

La bande magnétique créée par ce premier programme sert d'entrée à un second programme, qui interprète les codes synoptiques en valeurs numériques, utilisable par les programmes de traitement subséquents. Le second programme produit :

- a) une bande magnétique comprenant 3 types de fichiers :
  - fichiers de résultats d'observations,
  - fichiers de vents géostrophiques calculés,
  - fichiers de vents géostrophiques prédictifs;
- b) optionnellement une liste de messages en clair.

Ces deux programmes doivent absolument traiter les données avant que celles-ci soient effectivement utilisables. Aucun filtrage, aucune épuration, aucune interpolation ou extrapolation n'est réalisé à ce stade : les valeurs numériques fournies sont l'exacte et stricte traduction des messages codés (à condition toutefois que le sens de ceux-ci n'ait pas été altéré par des erreurs de transmission); de la sorte on peut après examens répétés, déterminer les éventuelles erreurs systématiques, repérer les stations dont l'information est peu sûre et prendre les dispositions nécessaires dans les programmes d'exploitation des données. Plusieurs de ceux-ci ont été développés, notamment :

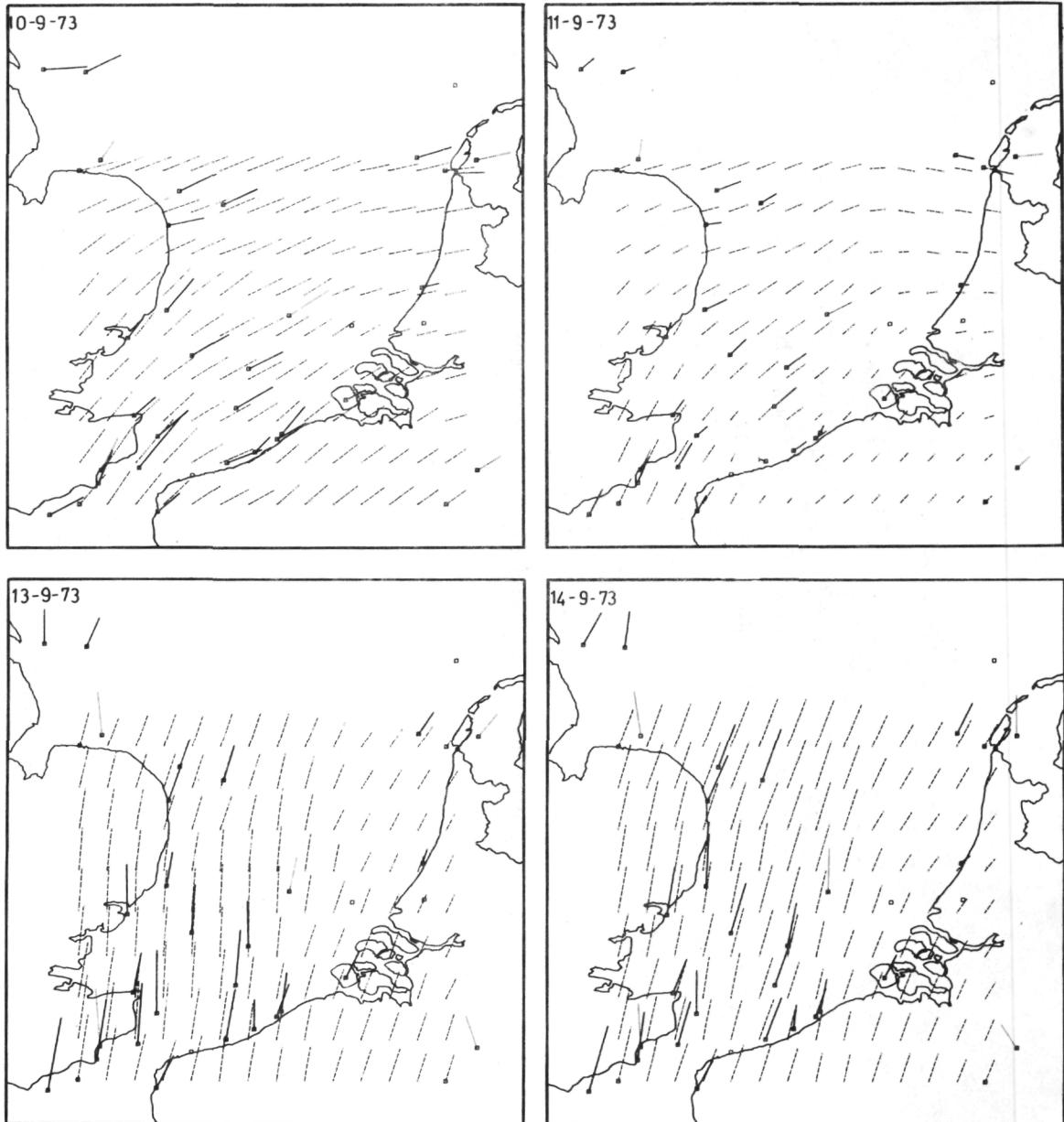
- a) un programme fournissant, par interpolation sur un réseau régulier couvrant le Southern Bight, la valeur du champ de vent (vent observé et vent géostrophique calculé). La figure 3 montre plusieurs répartitions du vecteur-vitesse du vent (moyenne sur 24 h) pendant quelques jours de la campagne JONSDAP 73. Les mêmes valeurs interpolées sont utilisées pour le calcul de la tension due au vent dans des programmes de simulation hydrodynamique. Les mêmes calculs seront effectués tout au long de la campagne JONSDAP 76;
- b) un programme permettant de faire le bilan thermique du Southern Bight en tenant compte de la température de la mer et de l'air, du taux d'humidité et de la couverture nuageuse.

Les données stockées sous le format actuel peuvent facilement être mises sous le format standard décrit plus loin.

fig. 3.

Champ de vent dans le Southern Bight durant JONSDAP 73

- Vent observé
- ..... Vent géostrophique observé
- - - - Vent interpolé
- 10 m/s



## 2.2.- Données des bouées automatiques

Le logiciel développé pour le traitement des données transmises par les bouées automatiques comprend 2 chaînes de programmes; toutes deux se limitent à construire des séries temporelles de données pré-traitées, c'est-à-dire d'où l'on a éliminé les erreurs dues à la transmission et au décodage. Les données résultantes sont prêtes à être analysées et exploitées.

a) La première chaîne est l'ensemble des programmes qui effectuent le prétraitement en temps différé (*off-line data acquisition system : DAS*) des informations fournies par le système d'acquisition de données sur ruban perforé. La bouée émet ses informations (canal par canal) qui sont reçues à Ostende par le système d'acquisition; celui-ci traduit les signaux radio en code digital et envoie les résultats vers une perforatrice automatique qui fournit un ruban de papier portant les données brutes écrites en code-caractère ASCII. Le ruban est lu par un premier programme qui stocke les données sur une bande magnétique, sous un format fixe; cette bande est alors traitée par une série de programmes qui

- produisent une liste des données;
- corrigent les erreurs de décodage;
- éliminent les cycles de données trop mal transmis.

La plupart des corrections (caractères incorrects, transmissions tronquées, ...) se font à l'aide de programmes standard du système d'exploitation. La bande purifiée est interprétée par un programme final qui, suivant les instructions de l'opérateur, exécute différents traitements des informations correspondant aux divers canaux de la bouée et qui construit les séries temporelles définitives. On ne décrira pas davantage cette chaîne étant donnée qu'elle ne sert plus que de système de sécurité en cas de défaillance de la seconde chaîne, beaucoup plus automatisée et facile à manipuler.

b) La seconde chaîne est un ensemble de programmes d'acquisition de données en temps réel (*on-line data acquisition system : OLDAS*). Grâce à elle, le prétraitement des données se fait de manière quasi-automatique, ne nécessitant l'intervention d'un opérateur que pour la

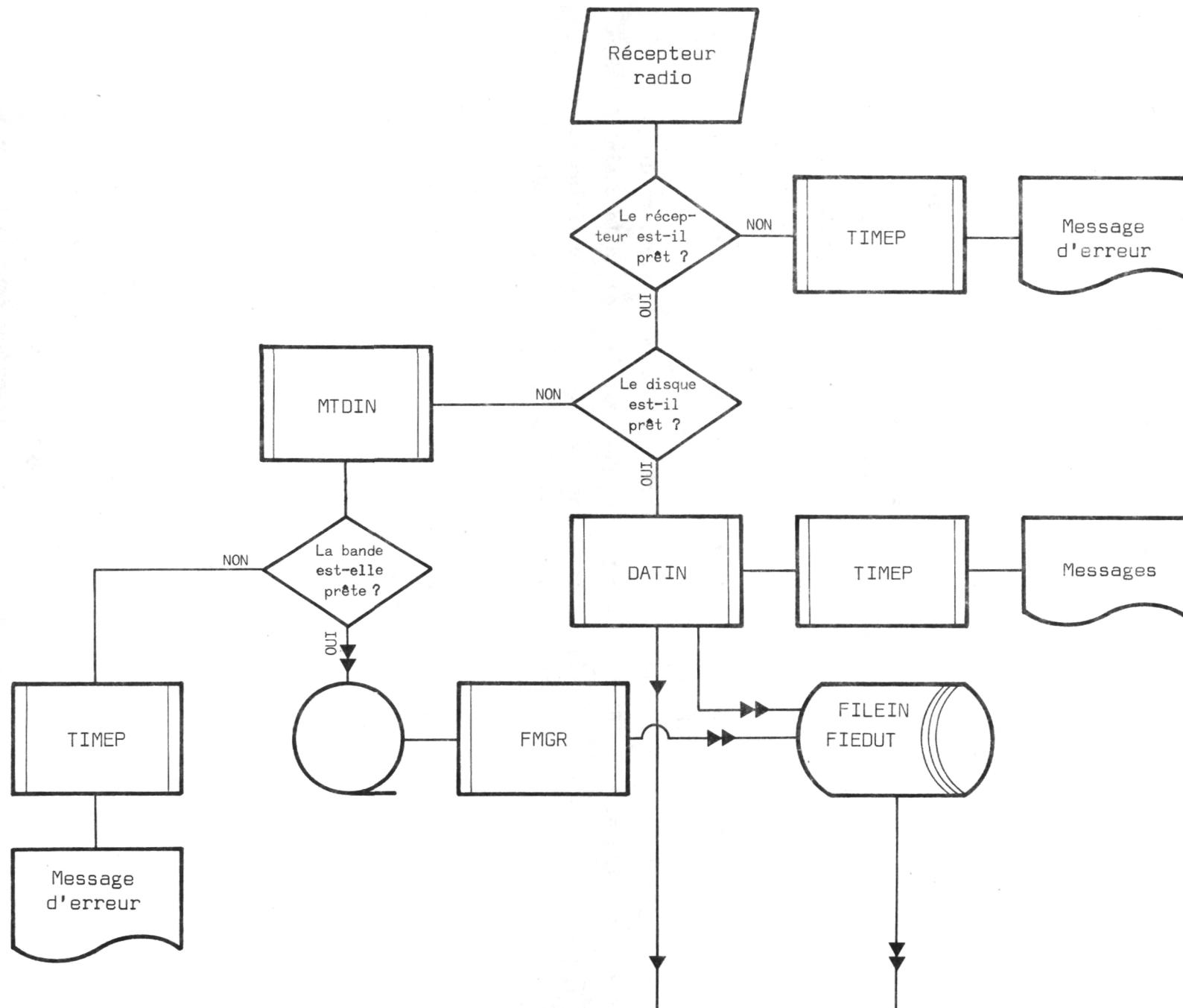
définition exacte des traitements à opérer (qui peuvent éventuellement différer d'une bouée à l'autre) et en cas d'erreur grave; l'acquisition elle-même peut se faire sans intervention humaine pendant des périodes fort longues. Pour réaliser ce logiciel très élaboré, il a fallu faire appel à toutes les ressources du système d'exploitation en temps réel (*Real Time Executive System*) qui gère actuellement l'ordinateur d'acquisition de données.

Les différents programmes sont activés automatiquement et selon les besoins du traitement par leur prédecesseur dans la chaîne.

La première phase est l'acquisition des données elle-même : dès que le récepteur capte une émission, un premier programme (DATIN) est activé automatiquement par le système-récepteur; selon que le disque est capable de stocker les données ou non, ce programme "lit" les données une à une et les sauve sur un fichier ou transfère le contrôle à un autre (MTDIN) qui lit les données et les sauve sur bande magnétique; cette bande peut soit être transformée en fichiers disques par la suite et être traitée par les phases suivantes de la chaîne OLDAS, soit être analysée par la chaîne DAS. Dès que l'acquisition est terminée, le contrôle passe à la phase de prétraitement. Pendant toute la durée de l'acquisition, le récepteur est logiquement déconnecté de sorte qu'il ne puisse y avoir d'interférences entre les émissions de plusieurs bouées. Bien entendu, des sécurités sont prévues pour éliminer les transmissions erronées (par exemple les messages sur la même fréquence envoyés par des émetteurs autres que la bouée). Les programmes DATIN et MTDIN restent en permanence dans la mémoire centrale de l'ordinateur, parce que :

- i) ils doivent pouvoir être exécutés même en cas de panne ou d'arrêt du disque,
- ii) ils doivent pouvoir être activés avec la plus grande rapidité possible.

La phase de prétraitement comprend les programmes PLINK et PRICES; le premier peut être activé par la première phase (DATIN) et dans certains cas spéciaux par PRICES; le second est activable par l'opérateur ou par PLINK. PLINK est un programme interface dont le but principal



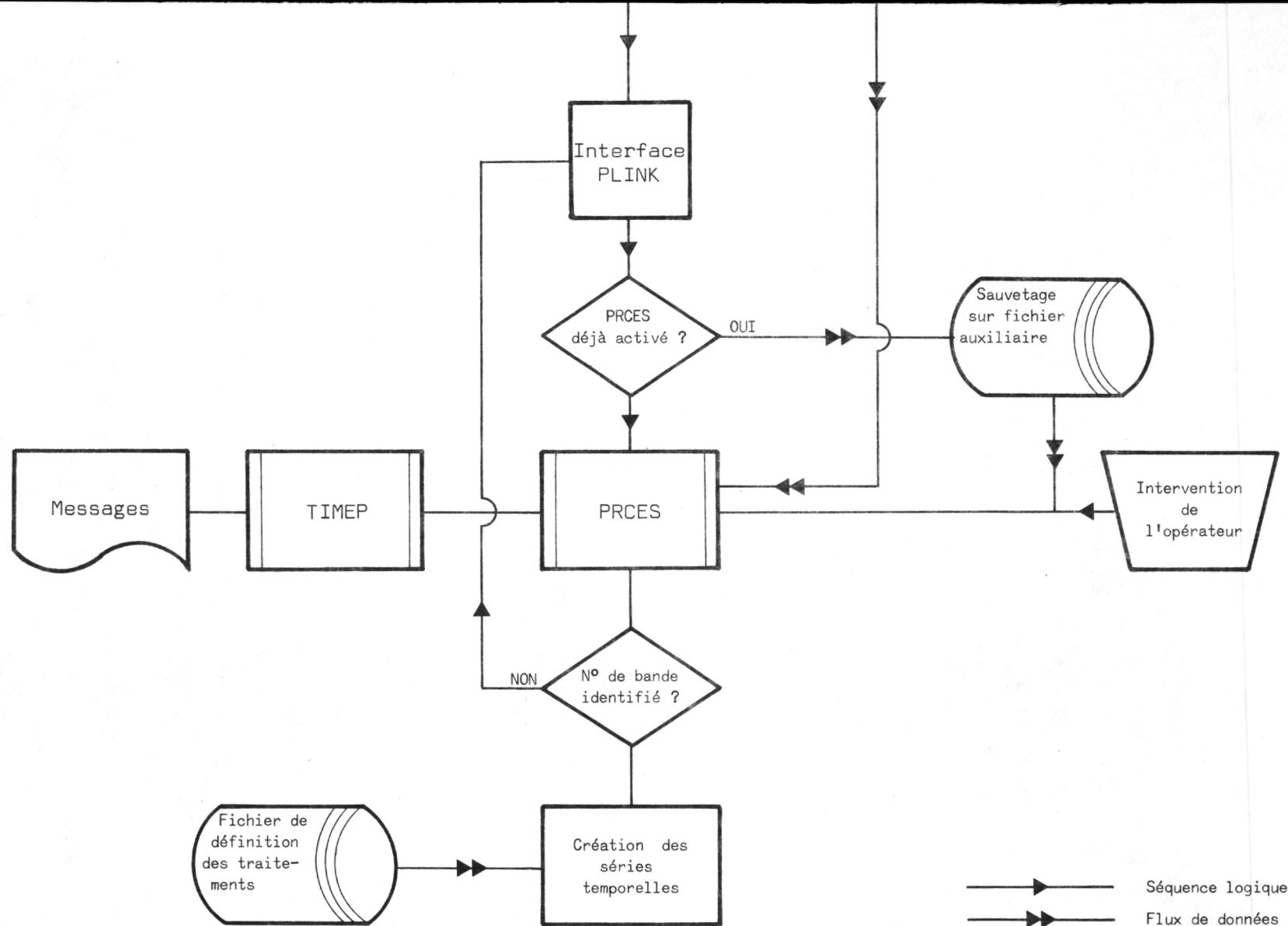


fig. 4.

Organigramme de la chaîne OLDAS

est de mettre provisoirement à l'abri des données en attendant que PRICES les manipule (ce dernier programme peut ne pas être disponible lorsqu'il est déjà en train de traiter les données d'une autre émission). PRICES identifie la bouée qui a émis les données contenues dans le fichier qu'il est en train d'examiner, va chercher dans un fichier auxiliaire des informations concernant les traitements qu'il a à effectuer (interprétation des informations des canaux) et opère ceux-ci. Les résultats sont stockés sur des fichiers-disque puis sur bande magnétique sous forme de séries temporelles dont la structure est semblable à celle des fichiers en format standard de traitement décrits plus loin. Les fichiers descriptifs mentionnés plus haut doivent être élaborés avant tout traitement effectif et ne contiennent que des informations concernant les bouées actives. Le diagramme de la figure 4 résume la logique des opérations effectuées par la chaîne OLDAS.

### 2.3.- Données courantométriques

Cette chaîne d'acquisition est la moins sophistiquée, car le décodage des cassettes provenant des instruments se fait chez le fabricant. La forme sous laquelle ses données sont reçues dépend du type de l'instrument.

Les données des VACM arrivent sous une forme prétraitées de la WHOI, où elles ont été décodées, nettoyées et mises en forme (avec notamment, l'interpolation effectuée pour des valeurs manquantes. Les bandes magnétiques fournies sont lisibles par les programmes de traitement directement, mais les données sont malgré tout mise sous un format intermédiaire, qui est le même pour tous les types de courantomètres, de manière à unifier les programmes subséquents.

Les données brutes NBA arrivent sur rubans perforés en caractères ASCII. Aucune édition ni prétraitement n'a été fait. Elles doivent être filtrées avant d'être mise sous le format intermédiaire.

Les données Plessey et Anderaa arrivent encodées selon un code binaire sur rubans perforés ou sur bandes magnétiques. Elles doivent être traduites en ASCII ou EBCDIC, avant d'être disponibles pour les

programmes d'exploitation. Les données doivent également être filtrées préalablement à leur insertion dans la base de données.

Pour chaque type d'instrument, il existe un programme qui établit les séries temporelles, les édite et les stocke sur des fichiers d'entrée pour les programmes d'analyse et de tracé.

### 3.- Données reçues par échange

La base de données comprend, en plus des données qui proviennent d'instruments de mesure automatiques, des informations qui nous sont parvenues dans le cadre d'expériences communes.

#### 3.1.- Données de l'expérience JONSDAP 73

##### 3.1.1.- Historique

JONSDAP 73 est une expérience entreprise en collaboration par 4 centres océanographiques riverains de la mer du Nord. Son but était l'étude approfondie de l'hydrologie et l'hydrodynamique du Southern Bight pendant une période relativement étendue, en l'occurrence les mois d'automne. Les institutions qui prirent part à cette campagne de mesures étaient :

- le Fisheries Laboratory de Lowestoft,
- l'Institute of Oceanographic Sciences (I.O.S.) de Bidston,
- le Koninglijk Nederlands Meteorologisch Instituut (K.N.M.I.) De Bilt,
- les chercheurs du Projet Mer, et plus particulièrement le groupe Math. Modelsea.

Les données acquises au cours de la campagne étaient essentiellement de nature physique, à savoir :

- des relevés de bathysondes,
- des élévations de marées (côtières et pélagiques),
- des mesures courantométriques,
- des données météorologiques diverses, dont on a déjà parlé précédemment.

L'expérience avait été soigneusement préparée en ce qui concerne la phase d'acquisition de données et de fait, la moisson fut abondante.

Il avait été convenu que les données acquises par chaque institution devaient être échangées avec les autres participants, de sorte que chaque groupe de recherche puisse appliquer à l'ensemble des mesures ses méthodes d'analyse propres et nourrir les modèles mathématiques qu'il développe avec toutes les données disponibles relatives à la région qu'il étudie.

L'échange des informations n'avait pas été préparé, chacun ayant estimé que le transfert des données d'un centre à l'autre ne poserait guère de problèmes.

### 3.1.2.- Difficultés de l'échange

Or la quantité de données à échanger (notamment les séries temporelles issues de courantomètres à enregistrement automatique) est telle que seuls peuvent être utilisés des supports de données susceptibles d'être traités automatiquement par des systèmes informatiques (en effet, les risques d'erreurs dues à d'éventuelles transcriptions manuelles deviennent rapidement prohibitifs).

Les supports utilisables sont :

- les bandes magnétiques,
- les bandes de papier perforé,
- les cartes perforées

(ce dernier support est peu recommandé quand le volume des informations est très grand car les risques de mélange et le poids des cartes fixent rapidement une limite à la quantité d'informations qu'il peut transférer). Seules quelques données, comme les élévations horaires du niveau de la mer ont été échangées par ce moyen, sans problème majeur heureusement. Par contre, il est exclu de l'utiliser pour l'échange des données courantométriques, beaucoup plus volumineuses. Les 2 autres supports se sont avérés indispensables mais ont posé de graves problèmes de compatibilité entre systèmes de traitement de l'information. En effet, la situation est la suivante :

- le Fisheries Laboratory, équipé de matériel ICL, fournit des données mises en forme sur bande magnétique;
- l'IOS-Bidston, équipé de matériel IBM, fournit également des données sur bande magnétique;
- le KNMI ne peut fournir des données que sur ruban perforé;
- Math. Modelsea, ayant accès au Centre de Calcul de l'Université de Liège, (matériel IBM) fournit des données sur bande magnétique.

Il se fait que si les bandes produites à Liège et à Bidston sont parfaitement lisibles par les ordinateurs réciproques (c'est le même matériel), Bidston et Lowestoft avaient jusqu'alors tenté en vain d'échanger leurs données, leurs systèmes respectifs pour l'écriture/lecture des bandes magnétiques étant incompatibles pour des raisons qui seront détaillées plus loin; d'autre part, aucun de ces deux derniers centres n'est équipé pour traiter des rubans perforés. Après de multiples essais infructueux, il a été décidé d'adopter la procédure suivante : toutes les données de quelqu'origine qu'elles soient, sont envoyées à Liège, où le Groupe Math. Modelsea tente de les décoder et de les mettre à la disposition des autres centres. La base de données courantométriques issues de JONSDAP 73 est donc centralisée à Liège; encore faut-il déchiffrer les données issues des divers centres, en faire une base de données exploitable, et les redistribuer aux institutions coopérantes sous une forme telle que disparaissent les incompatibilités qui avaient rendu impossible l'échange direct. Le choix de Liège comme centre d'échange est fondé sur deux raisons :

- i) le système d'ordinateur y est le plus complet et accepte tous les types de support;
- ii) l'équipe scientifique qui s'y trouve compte notamment des scientifiques expérimentés dans les problèmes de traitement de l'information et est plus apte que toute autre à surmonter les difficiles problèmes rencontrés.

### 3.1.3.- Mise au point des techniques d'échange

En principe, il suffit de résoudre trois problèmes, étant donné que les systèmes de Liège et de Bidston sont parfaitement compatibles :

- décoder les données fournies par le K.N.M.I. sur ruban perforé et les faire entrer dans la base de données centralisée à Liège; en effet, à l'heure actuelle, l'ordinateur du K.N.M.I. est capable de lire sans problème des bandes magnétiques créées à Liège;

- décoder les données fournies par le Fisheries Laboratory sur bandes magnétiques produites par un système ICL;

- encoder les données de la base centrale de manière à ce qu'elles puissent être utilisées directement par le système de traitement de Lowestoft.

Le problème du décodage des données courantométriques originaires du K.N.M.I. est similaire, dans les difficultés qu'il a fallu contourner, au problème posé par les données météorologiques qui est exposé dans le paragraphe suivant; on ne le détaillera donc pas ici; il faut cependant rappeler que le système d'exploitation du Centre de Calcul de l'Université de Liège (OS/360-VS fonctionnant sous ASP) n'est absolument pas conçu pour faciliter le traitement des bandes perforées; il a donc été nécessaire de créer une procédure complète de décodage, comprenant notamment :

- un programme de décodage proprement dit (passage du code-ruban perforé au code EBCDIC);

- un programme de translation (traduction des chaînes de caractères EBCDIC en valeurs significatives à introduire dans la base de données) pour "assimiler" ces données.

Les difficultés rencontrées lors de l'échange entre Lowestoft et Bidston ou Liège sont d'une nature nature : elles proviennent d'une incompatibilité de structure fondamentale entre bandes magnétiques produites par des ordinateurs de conception différente. En effet,

- les ordinateurs IBM utilisent des caractères de 8 bits, qui groupés par 4, peuvent former des mots de 32 bits tandis que les ordinateurs ICL utilisent des mots de 24 bits (unités fondamentales de la mémoire centrale) sur lesquelles sont éventuellement définis des groupes de 4 caractères de 6 bits; dans le premier cas, l'unité de base de la mémoire centrale est le caractère (ou byte) de 8 bits, tandis que dans le second cas, c'est le mot de 24 bits;

- le système d'exploitation d'IBM permet d'écrire des bandes magnétiques sans étiquette d'identification spéciale; le code-caractère (permettant un échange relativement aisé avec d'autres systèmes) d'IBM est le code EBCDIC largement utilisé; le système d'exploitation d'ICL exige la présence en début de bande, d'une étiquette d'identification de format déterminé; de même, le début et la fin des enregistrements de données sont délimités par des étiquettes (sentinelles) de format spécial; toutes ces étiquettes doivent être écrites en code interne (octal); le code-caractère d'ICL est un code spécial, basé sur des caractères de 6 bits (code ICL);

- chez IBM, les enregistrements logiques sont définis par le système d'exploitation sur des enregistrements physiques; chez ICL, les enregistrements logiques compris dans les enregistrements physiques sont séparés, sur la bande magnétique elle-même, par des identificateurs qui en définissent la longueur.

On conçoit donc la difficulté de transférer des données d'un système à l'autre; le passage ne peut se faire qu'en analysant, au niveau des chaînes de bits écrites sur les bandes, les structures des enregistrements physiques.

Pour mettre des données provenant d'une bande ICL sous une forme directement lisible par le système d'exploitation OS/360, il faut :

- ignorer les diverses étiquettes et sentinelles,
- découper chaque chaîne de bits correspondant à 1 enregistrement physique en groupes de 6 , et faire correspondre à chacun de ces groupes un caractère EBCDIC;
- éliminer les identificateurs de longueur, en tenant compte du format d'écriture des données et reformer des chaînes de caractères continues; ces chaînes sont lisibles alors par des ordres de lecture standard du système IBM.

Pour mettre des données provenant d'une bande IBM convenablement structurée, sous une forme directement lisible par le système d'exploitation ICL, il faut :

- recréer les étiquettes et sentinelles sous la forme imposée;

- déterminer les identificateurs de longueur et les insérer (sous forme de chaînes de caractères EBCDIC) entre les enregistrements logiques;

- faire correspondre à chaque groupe de 8 bits un caractère du code ICL et reformer des chaînes continues; ces chaînes sont alors lisibles par des ordres de lecture standard du système ICL.

Il a fallu mettre au point toute une panoplie de programmes pour réaliser le transfert dans les deux sens, en passant par l'écriture des données sous un format intermédiaire.

Le format intermédiaire est celui sous lequel sont écrits les fichiers composant la base de données courantométriques; c'est un format transitoire entre le format d'échange et le format standard de la banque de données générale, qui sera décrit plus loin; le format intermédiaire est d'ailleurs provisoire et les données sont écrites sous format standard au fur et à mesure que les programmes d'exploitation de la banque de données sont mis au point. Quant au format d'échange, il a actuellement la structure suivante : chaque fichier correspondant à une série de mesures issues d'un même instrument comporte :

a) un enregistrement (en-tête) définissant les données, il définit :  
- la zone marine où les mesures furent prises (ici, la mer du Nord),

- l'époque approximative des mesures (septembre-octobre),
- le nom de l'expérience en cours (ici, JONSDAP 73),
- la position du mouillage (mot code définissant une station de prise de mesures),
- la profondeur à laquelle se trouve l'instrument (fond, milieu, surface),
- le numéro de série de l'instrument,
- l'heure du début des mesures (jour, mois, année, heure, minute),
- le nombre de cycles de données,
- la variation magnétique,
- l'indication du traitement déjà effectué (par exemple, si une moyenne a été calculée),
- l'intervalle de temps entre cycles de données;

b) des enregistrements contenant les cycles de données;  
c) des enregistrements contenant les moyennes horaires des mesures;  
ceci a été introduit pour permettre des comparaisons directes entre mesures fournies par des instruments dont l'intervalle d'échantillonnage n'est pas le même.

A l'heure actuelle, le problème de l'échange des données de JONSDAP 73 peut être considéré comme résolu. Il faut rappeler que le format actuel d'échange est exclusivement conçu pour la manipulation de données courantométriques.

### 3.2.- Données de l'expérience JONSMOD

L'expérience JONSMOD est encore actuellement en cours. Elle consiste essentiellement à comparer différents modèles pour la prédiction des marées et des ondes de tempête en mer du Nord, nourris par un ensemble de données communes, en l'occurrence des élévations connues du niveau de la mer et les conditions météorologiques qui existaient pendant la période simulée (automne 73); ces données communes, ainsi que des résultats permettant la validation des modèles, constituent l'ensemble à échanger. Les mêmes raisons qui imposèrent le choix de Liège comme centre d'échange pour JONSDAP 73 ont conduit les membres du groupe JONSMOD à y centraliser l'ensemble commun de données. Les problèmes posés sont en effet fort similaires, puisque le groupe JONSMOD comprend, outre les quatre institutions qui coopérèrent pour JONSDAP 73, des chercheurs des centres suivants :

- Institut Royal Météorologique Danois (Copenhague),
- Institut für Meereskunde (Hambourg),
- Service Hydrographique et Océanographique de la Marine (Paris).

Les données sont de nature moins sophistiquée que dans le cas précédent, mais le nombre de partenaires (donc de systèmes) est plus élevé. La complexité des échanges est donc fort similaire. On va dans les sous-paragraphes qui suivent, détailler les problèmes qui ont surgi à chaque étape de l'échange.

### 3.2.1.- Détermination des restrictions sur les supports de données

Le premier souci est d'essayer d'obtenir de chacun une liste exhaustive des possibilités et contraintes concernant leur matériel informatique. Après de nombreux rappels et recherches, on est en mesure de connaître au moins un type de données standard pour chaque centre concerné :

- Lowestoft	bandes magnétiques	code ICL
- Bidston	bandes magnétiques	code EBCDIC
- K.N.M.I.	rubans perforés	MC flexowriter code
- Paris	bandes magnétiques	code EBCDIC
- Hambourg	rubans perforés	code ISO (ASCII)
- Copenhague	rubans perforés	code ISO (ASCII) .

### 3.2.2.- Les formats

Si la gamme des machines et des codes de représentation est très étendue, les formats eux aussi sont très variables, chaque centre de recherche construisant le format propre à ses besoins et ses applications sans tenir compte d'un éventuel échange de données. De plus, certaines limitations du matériel informatique ne laissent aux utilisateurs que très peu de liberté quant au choix du support et de la disposition des données sur ce support. C'est ainsi que l'on rencontre des enregistrements sur bandes magnétiques et rubans perforés avec ou sans format, de longueurs fixe ou variable. Dans de semblables conditions, on peut s'attendre à recevoir n'importe quoi, sous n'importe quelle forme.

### 3.2.3.- Traitement du support de données

#### 1) Lecture

Pour effectuer le traitement d'un support quelconque de données, il est indispensable de connaître un certain nombre de renseignements techniques. Lors des premiers travaux, il fut souvent nécessaire de réclamer un complément d'informations ou d'effectuer soi-même certains tests pour vérifier la validité des renseignements donnés. Parfois,

il est indispensable de déterminer soi-même des détails manquants. Les informations nécessaires pour lire les fichiers concernent la méthode d'enregistrement, le code de représentation accompagné éventuellement d'une table d'équivalence et enfin quelques détails techniques tels que : longueur de l'enregistrement physique, densité d'enregistrement, parité, etc. Toutes ces caractéristiques permettent d'effectuer une première lecture du fichier déterminant son contenu. Chaque institution employant des formats différents, il faut concevoir à l'entrée et à la sortie, des chaînes de programmes particulières à chaque application.

## 2) Traitemet de rubans perforés

Les données reçues sur rubans perforés et restaurées sous cette même forme nécessitent un travail supplémentaire. En effet, le Centre de Calcul de l'université de Liège ne dispose pas d'un lecteur de rubans perforés directement relié à l'ordinateur. Il faut donc au préalable passer par un système d'encodage sur bande magnétique. Un encodeur lit l'information et la recopie sur une bande magnétique selon un format fixe comprenant 100 caractères. A la sortie, la procédure s'effectue en sens inverse : il faut créer un fichier sur bande magnétique avec des enregistrements de 100 caractères afin de perforer le ruban par la suite.

Cette opération d'encodage est très délicate. Elle nécessite beaucoup de manipulations de la part des opérateurs et du programmeur qui prépare le ruban perforé. En effet, pour être traité, ce ruban doit répondre à certaines exigences :

- il doit comporter des amorces au début et à la fin,
- il doit être enroulé dans un certain sens avec indication de début et fin. Il est de plus conseillé de ne pas utiliser des rubans d'une longueur supérieure à 300 mètres.

La plupart des rubans reçus ne correspondent pas à ces normes. Des vérifications minutieuses s'imposent afin de remédier aux imperfections. Les erreurs demandent de longues corrections manuelles et alourdissent considérablement le travail d'échange de données.

### 3) Décodage de l'information

L'information transférée des rubans encodés sur bande magnétique ou reçue directement sur bande magnétique est maintenant traitable par l'ordinateur. Les données sont prises en charge par un programme de décodage. Ce programme crée un fichier reprenant en EBCDIC tous les caractères bons ou mauvais. A chaque appel du sous-programme décodeur un enregistrement est transcodé. On obtient une chaîne de caractères EBCDIC correspondant aux caractères perforés sur le ruban. L'enregistrement peut avoir une longueur variable limitée toutefois à 100 caractères. Ce transcodage est effectué au moyen d'une table d'équivalence à deux entrées. Chaque caractère y est situé à une place correspondant à sa valeur numérique dans le premier code. On calcule donc automatiquement la valeur décimale des perforations de tous les caractères à décoder, et on obtient chaque fois un numéro de case de la table. On lit le contenu de cette case de la table, c'est le caractère EBCDIC correspondant. Lorsque toute l'information est traitée de cette façon, on obtient un fichier brut décodé.

### 4) Toilette des informations

Si les fichiers reçus sur bandes magnétiques correspondent généralement à la description fournie, il n'en est pas de même pour les fichiers sur rubans perforés. Ils contiennent parfois des caractères totalement étrangers aux données. Par exemple, il arrive de constater la présence d'une multitude de signes négatifs devant une donnée ou encore un décalage de l'information par rapport au format annoncé. Il s'est avéré nécessaire de mettre au point des programmes de nettoyage de données créant un fichier correct. Les enregistrements de celui-ci correspondent alors tous aux caractéristiques annoncées par l'expéditeur. Ce travail de nettoyage des données est très fastidieux car chaque cas doit être traité séparément, les erreurs étant distribuées de façon tout à fait aléatoire. Il consiste principalement à :

- éliminer tous les enregistrements composés de caractères spéciaux ne concernant pas les données;

- éliminer tous les caractères erronés se situant à l'intérieur d'enregistrements normaux;
- insérer ou supprimer des blancs en cas de décalage;
- corriger des inversions de chiffres.

Le fichier correct peut alors être traité et traduit dans les formats des divers organismes auxquels il est destiné.

#### 5) Fourniture des données

Cette recodification s'effectue suivant un processus inverse. On constitue une table de l'alphabet avec les équivalents binaires dans le code désiré. On crée alors un fichier où chaque caractère est remplacé par son équivalent binaire dans le code de sortie, tout en respectant les normes de présentation de chacun des organismes. La fourniture des données peut, elle aussi, réclamer la copie du fichier sur ruban perforé.

#### 6) Les aléas de l'échange des données

Le transcodage peut se compliquer légèrement lorsque la longueur des mots-machine varie d'un ordinateur à un autre. C'est le cas lorsqu'on reçoit des données provenant d'un ordinateur ICL, ou que l'on doit fournir des informations à un centre équipé de ce matériel. Cette difficulté a déjà été traitée précédemment.

Certaines difficultés supplémentaires peuvent surgir lorsque, d'un envoi à un autre, la structure de l'information d'un même organisme est modifiée. L'ensemble des programmes constitués pour transcoder les données de ce centre doit être alors totalement repensé. Le cas se présente lors du passage d'une information structurée à une suite de données séparées les unes des autres par un ou plusieurs blancs. Aucune spécification de longueur des blocs d'information n'est indiquée. A ce moment, de nouveaux programmes doivent isoler les données et reconstituer un fichier structuré.

D'autres problèmes peuvent encore se présenter si certaines données sont imprimées ou fournies sur graphique. Le traitement de telles informations entraîne toujours un surcroît de travail fastidieux.

### 3.2.4.- Acquis actuel de l'échange des données JONSMOD

Les premiers échanges de données effectués au cours de la campagne JONSMOD ont déjà permis à plusieurs centres d'utiliser, dans le cadre de leurs recherches, une masse de données jusqu'alors inaccessibles. Cette distribution de données est destinée à s'accentuer et à sortir du cadre de JONSMOD. Toutefois, cette opération JONSMOD a permis de tirer quelques enseignements précieux.

Le support le plus maniable reste la bande magnétique pour autant que la description de l'information enregistrée soit suffisante pour en permettre la lecture. Presque tous les autres supports sont acceptables mais nécessitent trop souvent des traitements supplémentaires, longs et coûteux.

De plus, devant la grande variété des formats employés par tous les centres concernés, il est nécessaire de discipliner et standardiser le format des données échangées. Cette standardisation permet à chacun de pouvoir utiliser directement les données reçues dans son logiciel de traitement. Plusieurs structures de format autodescriptif sont proposées. L'une de celles-ci prend de plus en plus d'extension et est déjà développée dans certains grands centres océanologiques. Il s'agit du format GF2 (anciennement GATE) dont il sera question plus loin.

Il est en tous cas indispensable de poursuivre un échange de données plus intensif afin de fournir aux chercheurs une banque de données plus complète, et de continuer le développement d'un format généralisé afin d'accélérer et faciliter cet échange de données.

## Chapitre III

### Exploitation des données

#### 1.- Construction de la banque de données et du logiciel d'exploitation

Une base de données devient une banque de données à partir du moment où les données sont structurées de manière telle qu'on puisse écrire un logiciel capable de les rechercher, les analyser, les traiter; ce logiciel doit être de maniement aisément et commun à tous les types de données. La structure proposée ici n'a pas la prétention d'être accessible de manière transparente pour les utilisateurs, tel que certains systèmes utilisés dans le domaine commercial ou bancaire. Le logiciel ne sera pas étendu aussi loin pour deux raisons essentielles :

1) les données scientifiques (telles que les données océanographiques et météorologiques) sont de types et de structures beaucoup plus variées que les données commerciales, et nécessitent une exploitation beaucoup plus complexe. Il est absurde d'établir un logiciel très sophistiqué, qui augmenterait considérablement le temps de calcul alors que l'on est limité en moyens matériels et humains.

2) la banque de données et le logiciel d'exploitation sont créés dans le but d'être utilisés par des programmeurs qualifiés et par des scientifiques qui ont un minimum d'expérience en programmation et qui sont supposés connaître le type d'information qui leur est nécessaire et la manière dont ils veulent traiter les données. Une petite équipe de spécialistes du logiciel est cependant nécessaire pour aider les scientifiques à assembler les chaînes de programmes et pour les conseiller sur le choix des techniques de calcul.

Cette solution a déjà été expérimentée dans plusieurs centres de recherche océanographique, comme la WHOI, l'université de Southampton et le BNDO au Centre Océanologique de Bretagne.

Le but de ce projet est de construire un système suffisamment flexible pour permettre le stockage, l'archivage, la sélection et l'analyse des jeux de données présents et futurs, quelles que soient leur origine et leur structure. Le système est destiné :

- à faciliter le travail des scientifiques en leur donnant des outils standards
  - 1) pour l'analyse de leur matériel,
  - 2) pour la validation et la comparaison des modèles mathématiques;
- à être étendu facilement par des programmes aux fonctions spécifiques.

Finalement, le stockage des données et le logiciel doivent être portables (c'est-à-dire capables d'être utilisables avec peu de modifications sur la plupart des ordinateurs) afin que le plus grand nombre de centres de recherche bénéficie du travail d'élaboration. Il est donc essentiel que la structure des programmes et des enregistrements de données soit compatible avec la grande majorité des ordinateurs utilisés dans les centres.

## 2.- Conditions de portabilité du format de stockage et du logiciel

Les spécialistes en programmation savent parfaitement que toute information écrite par un ordinateur peut être lue par un autre, pourvu que le logiciel de translation existe.

La portabilité du logiciel est moins évidente. Tous les langages ne sont pas utilisables sur tous les ordinateurs, et la translation d'un langage à un autre demande souvent beaucoup de temps et d'argent. En général, les centres de recherche océanographique ne disposent pas des moyens nécessaires pour une telle conversion. C'est pourquoi le logiciel développé dans ce projet de recherche est destiné à être utilisé avec seulement quelques simples modifications relatives aux

opérations d'entrée/sortie, sur la plupart des ordinateurs scientifiques.

En raison de la structure interne des programmes, de la flexibilité du format de stockage et de la totale impossibilité d'écrire des programmes d'exploitation de données absolument indépendants du matériel (*hardware*), il est nécessaire que le matériel et le logiciel présentent les caractéristiques suivantes :

- 1) un accès direct entre la mémoire centrale et les bandes magnétiques et disques est possible,
- 2) un mot réel en simple précision est stocké en deux mots entiers,
- 3) un compilateur pour ANSI FORTRAN IV existe,
- 4) des enregistrements de différentes longueurs peuvent être lus et écrits sur bande magnétique (éventuellement, au moyen d'un ensemble restreint de programmes en langage spécial d'assemblage).

ANSI FORTRAN IV est le langage de programmation le plus répandu pour des applications scientifiques et techniques. C'est donc le meilleur langage pour écrire une bibliothèque scientifique.

La seconde condition ci-dessus, signifie qu'un nombre en virgule flottante en simple précision est représenté en mémoire sous la forme de deux mots entiers. Ceci se fait automatiquement sur des ordinateurs dont la structure est basée sur l'utilisation des mots de 16 bits ou 24 bits et ceci peut être obtenu au moyen de déclarations (instructions) spéciales sur des ordinateurs dont la structure est basée sur l'utilisation de caractères de 8 bits ou de mots de 60 bits.

Le format de la bande magnétique est portable en ce sens que la structure des enregistrements sur bande magnétique est la même pour tous les ordinateurs. (Les principes énoncés ici sont valables pour le traitement des données et le format de stockage et non pour le format d'échange. Ce dernier sera développé plus loin.) Ainsi, on utilise le mode d'enregistrement binaire pour la concision et la performance des entrées/sorties, c'est-à-dire que les données sur bande magnétique sont une copie exacte de leur représentation en mémoire centrale. Un format d'échange de données sera écrit dans un code caractère international, comme EBCDIC ou ASCII. Un tel code ne convient pas pour une analyse

rapide des données et est beaucoup moins compact que le code binaire. Mais en dépit du mode d'enregistrement binaire, les bandes écrites par un ordinateur peuvent être lues facilement par un autre si les conditions suivantes sont satisfaites :

- 1) les spécifications des bandes magnétiques sont compatibles sur les deux ordinateurs (c'est-à-dire, même nombre de pistes, même densité, même technique d'encodage, même marque de fin d'enregistrement et de fin de fichier).
- 2) la structure des deux ordinateurs est basée sur l'utilisation de mots de 16 bits ou de caractères de 8 bits; la structure des deux ordinateurs est basée sur l'utilisation de mots de 24 bits; la structure des deux ordinateurs est basée sur l'utilisation de mots de 60 bits (les ordinateurs sont alors appelés ordinateurs de structure compatible au niveau du mot ou du caractère).
- 3) des programmes existent pour la translation des codes-caractère et des codes internes.

Des dispositions sont prises au niveau du format de stockage pour définir le code interne utilisé dans l'écriture des nombres réels et des caractères, de sorte que les programmes appropriés (condition 3) puissent être sélectionnés. Les clefs définissant ces codes peuvent être décodées par tout ordinateur de structure compatible avec l'ordinateur source, parce que les nombres entiers positifs ont le même format dans tous les systèmes basés sur la même longueur du mot entier.

### 3.- Description générale du logiciel

Le logiciel peut être défini comme une extension d'un sous-ensemble de "WHOI Standard buoy format package" [Maltais (1969)] plus simple sous divers aspects. Le logiciel original est modifié dans le but d'être plus général (pour traiter une grande variété de données) et plus portable. Il est étendu pour répondre aux exigences du modèle mathématique. Le logiciel ainsi défini consiste en :

- 1) Un ensemble de programmes généraux pour lire et écrire les enregistrements de données, pour étiqueter et retrouver les fichiers de

données (sur bandes magnétiques et sur disques) et pour transférer les données entre les mémoires de masse et la mémoire centrale (dans des tableaux accessibles aux utilisateurs) et vice-versa.

2) Un ensemble de programmes de traitement de données (utilisant les programmes généraux) pour sélectionner, éditer, imprimer et tracer les données de façons différentes selon leur structure.

3) Un ensemble de programmes mathématiques :

- pour analyser les données,
- pour comparer et adapter les modèles aux résultats expérimentaux.

#### 4.- Stockage des données en format standard

Le format se décrit par lui-même : chaque fichier de données contient toutes les informations nécessaires pour être lu correctement quelque soit le nombre de variables, le nombre d'échantillons, leur nature, leur structure (fonction continue d'une autre variable ou non), leur mode d'enregistrement, et le nombre d'enregistrements sur le fichier. Les bandes magnétiques écrites en format standard contiennent également, au début des données, un fichier d'identification et, après le dernier fichier de données, un fichier de fin de données. Chaque fichier de données consiste en :

- 1) deux enregistrements d'étiquette,
- 2) des enregistrements de données (en nombre quelconque).

L'enregistrement de la première étiquette a une longueur fixée et définit :

- les clefs pour la conversion des données binaires (tableau 2),
- la date de création du fichier,
- le nom du fichier,
- le nombre de variables,
- le nombre d'échantillons,
- la date des mesures,
- le type de la seconde étiquette.

Tableau 2  
Etiquette 1

Zone	Contenu de la zone
1	mot clé (entier = 1)
2	Interchange code; convention : 1 → EBCDIC 2 → ASCII (entier) 3 → BCD
3	Internal code; convention : 1 → IBM (entier) 2 → HP
4-11	nom de données (16 bytes)   4 b. (code mnémonique)   4 b. # croisière ICES   2 b. # code ICES pays   2 b. # code ICES bateau ou bouée   4 b. Histoire des données
12 - 16	époque de la création du fichier : année - jour - heure - minute - seconde (5 entiers)
17 - 18	source des données (4 b.)
19	# variables (entier)
20	# échantillons (entier)
21	Format de la 2e étiquette (entier)
22 - 26	époque initiale des données : année - jour - heure - minute - seconde (5 entiers)
27 - 98	144 b. de commentaires

L'enregistrement de la seconde étiquette définit (tableau 3) :

- la position géographique où les mesures ont été effectuées (latitude, longitude ou autre information);
- les paramètres d'échantillonnage;
- la valeur initiale de la variable continue;
- pour chaque variable :
  - le nom,
  - les unités,
  - le code de l'appareil de mesure,
  - le code de l'instrument,
  - le numéro de série,
  - le type des données à l'enregistrement (entier, réel, synoptique),
  - la profondeur ou la pression,
  - trois attributs.

Tableau 3

Etiquette 2

Zone	Contenu de la zone
1	mot clé (entier = 2)
2	# échantillons/intervalle (entier)
3 - 4	intervalle d'échantillonage (réel)
5 - 6	intervalle entre échantillons (réel)
7 - 8	jour julien } ou { valeur initiale
9 - 10	secondes } - 1 } (2 réels)
11 - 14	8 b. : latitude degrés - minutes - secondes - N, S (4 x 2 b.)
15 - 18	8 b. : longitude idem
19 - 22	8 b. : variation magnétique
23 - 28	12 b. : nom de variable
29 - 34	12 b. : unités dimensionnelles
35 - 36	4 b. : code mesure
37 - 38	4 b. : code instrument
39 - 40	4 b. : série
41	2 b. : type des données { 'I' entier { 'R' réel (T,V,R) { 'S' synoptique
42	profondeur (en m) ou pression (mb) ou 0 (entier)
43 - 44	1er attribut (réel) : valeur de biais
45 - 46	2ème attribut (réel) } non attribué
47 - 48	3ème attribut (réel)

} 24 bytes  
pour indiquer  
un autre format

jusqu'à 64 fois cette zone

Tableau 4  
Enregistrement des données

Zone	
1	mot clé (entier = 3)
2 - 9	nom des données (16 b.)
10	# échantillons précédents (entier)
11	# échantillons dans l'enregistrement (entier)
12	pas utilisé
13 - 14	jour julien ou valeur initiale
15 - 1166	données proprement dites

Un enregistrement de données consiste en (tableau 4) :

- le nom du fichier de données (pour contrôler si les données lues sont celles décrites dans l'enregistrement de la première étiquette);
- le nombre d'échantillons dans les enregistrements précédents;
- le nombre d'échantillons dans l'enregistrement courant;
- la valeur initiale de la variable continue;
- les données.

La structure détaillée de chaque type d'enregistrement est décrite dans les tableaux 1 à 3, uniquement pour les ordinateurs dont la structure est basée sur l'utilisation de mots de 16 bits ou 32 bits.

Tous les fichiers existant actuellement dans la banque de données sont en voie d'être mis en forme suivant les spécifications décrites ci-dessus.

## 5.- Etat actuel du logiciel d'exploitation

### 5.1.- Mémorisation d'ensembles de données sous une forme standardisée

Vu la nécessité de manipuler de grands ensembles de données, il s'est avéré utile de déterminer une méthode d'archivage. Un groupe de programmes qui permettent de conserver n'importe quel ensemble de données sous une forme standardisée ont été mis au point. Ces programmes rangent les données sur bande magnétique car c'est ce support qui offre le plus d'avantages lorsque l'on traite de grands volumes d'informations. Les données sont répertoriées en une série de fichiers, chaque fichier étant constitué par des données qui ont un caractère commun. La structure d'un fichier est élaborée de manière à être autodescriptive. La définition des données ainsi que tous les renseignements signalétiques propres à chaque variable sont eux-mêmes enregistrés sur le fichier. Ce procédé permet de conserver dans un même fichier un nombre quelconque de variables.

### 5.2.- Structure des fichiers

Dans le format que l'on a adopté, trois types de fichiers existent.

#### 1) Fichier d'identification de la bande

Il est constitué d'une suite d'enregistrements identiques qui identifient la bande par un numéro de série. Ce fichier est unique sur une bande.

#### 2) Fichier de données

Il contient toutes les informations qui décrivent les données à enregistrer, plus les données proprement dites.

#### 3) Fichier de fin des données

Ce fichier est unique sur la bande et contient un enregistrement spécial qui détermine la fin de la bande.

### 5.3.- Description synthétique des programmes

#### 1) Programme d'écriture du fichier d'identification de la bande (TAPID)

Ce programme initialise la bande avec des paramètres tels que le numéro d'immatriculation, le jour et l'année de création. A la suite du fichier d'initialisation, il crée un fichier de fin de données.

#### 2) Programme d'écriture du fichier de fin de données (EØD)

Le fichier de fin de données contient la juxtaposition de 24 fois la chaîne de caractères 'EØD'.

#### 3) Programme d'acquisition des paramètres décrivant les données (GLABE)

Au début d'un fichier contenant les données, se trouvent deux enregistrements décrivant les données qui suivent. Le premier enregistrement contient des informations qui concernent le fichier globalement (par exemple le nom du fichier, le nombre de variables, ...). Cet enregistrement est de longueur fixe. Le second enregistrement contient les informations qui définissent les variables (par exemple, le nom de la variable, le type, le code mesure, ...). Cet enregistrement est de longueur variable.

Le programme GLABE initialise les deux enregistrements au moyen de paramètres lus sur des cartes perforées. A l'issue de l'exécution du programme, les deux enregistrements ne sont pas encore écrits sur la bande, mais leurs images sont conservées en mémoire de l'ordinateur dans deux vecteurs.

4) Programme d'écriture des deux enregistrements de tête d'un fichier de données (PLABE)

L'image des deux enregistrements initialisés par GLABE sont effectivement copiés sur la bande magnétique.

5) Programme de recherche d'un fichier déterminé sur la bande (RLABE)

Un nom de fichier et un nombre signifiant la quantité de fichiers à explorer sont fournis au sous-programme. Celui-ci va passer en revue une suite de fichiers jusqu'à ce qu'il ait découvert celui qui est recherché, ou jusqu'à ce que le nombre de fichiers examinés soit égal à celui qui lui a été spécifié. Un code retourné par le programme indique s'il a trouvé le fichier ou si une anomalie existe.

6) Programme d'écriture des données (OUTDA)

L'utilisateur doit ranger ses données par lots de taille arbitraire dans un tableau qui sera pris en charge par le programme. Le programme OUTDA puise les données dans ce tableau et les transfère dans une mémoire tampon interne. Le nombre de cycles ainsi transférés à chaque appel est fourni à OUTDA par un paramètre. Pour mémoriser une variable temps, il faut avoir initialisé l'une des séries avec les secondes dans la journée et le paramètre 'JULIAN' avec le jour julien. Le contenu de la mémoire tampon est écrit sur la bande lorsqu'elle est remplie à son maximum ou lorsque le paramètre 'JULIAN' est modifié.

En effet, avec la structure adoptée, des données réparties chronologiquement sur deux jours ne peuvent figurer dans le même enregistrement.

7) Programme de lecture des données sur la bande (INDA)

L'appel de ce sous-programme provoque la lecture d'un enregistrement de la bande et le transfert des informations qui y sont contenues

dans un vecteur. Les informations y sont structurées d'une manière qui permet leur exploitation directe par l'utilisateur ou par les programmes de sélection.

8) Programme évolué permettant de sélectionner sur la bande en format standard des variables particulières (SHELL)

Il a pour fonctions :

- a) de retrouver un fichier déterminé sur une bande,
- b) de sélectionner des variables particulières hors du fichier,
- c) de n'extraire que les données comprises entre deux limites (limites incluses),
- d) de modifier les données à l'aide d'opérations algébriques élémentaires,
- e) de transférer les données sélectionnées sur un autre support.

Chaque variable est transférée dans un fichier séparé sur disque.

A chaque appel de SHELL, un groupe de données est déplacé dans chaque fichier disque. Un paramètre détermine la manière dont il faut échantillonner les données :

- HOW=0 : Toutes les données sont transférées par un seul appel à SHELL.
- HOW=1 : Les données sont transférées par une suite de blocs correspondant à un intervalle de temps d'échantillonnage. Il faut que les données soient équidistantes dans le temps. L'intervalle d'échantillonnage est enregistré dans l'étiquette # 2.
- HOW=2 : Les données sont transmises par blocs de longueur arbitraire. Si la variable continue est le temps, le paramètre SAMP fixe la durée d'échantillonnage sur la bande. Si la variable continue n'est pas le temps, SAMP fixe le nombre de données qu'il faut regrouper à chaque appel.
- HOW=3 : Les données sont transmises en une suite de blocs qui correspondent à la taille du bloc sur la bande.
- HOW=4 : Les données sont transmises par blocs correspondant à l'intervalle de temps d'échantillonnage enregistré dans l'étiquette # 2 du

fichier. Dans ce cas-ci, les données enregistrées ne sont pas équidistantes dans le temps.

### 9) Programme de lecture des paramètres sur la carte standard (SICC)

Le programme SHELL a besoin pour son fonctionnement d'un ensemble de paramètres. Certains sont obligatoires, d'autres sont facultatifs. Les paramètres obligatoires sont obtenus par le programme SICC. Ces paramètres sont :

- le nom du fichier à traiter,
- le numéro de la variable continue; si la variable continue est le temps, il faut mettre 0,
- la période de temps exprimée en jours juliens et secondes qui détermine l'intervalle sur lequel les données seront échantillonnées sur la bande; ou bien le nombre de données à regrouper pour former un bloc de données sur les nouveaux fichiers disques,
- l'endroit du fichier à partir duquel il faut extraire les données, exprimé soit sous la forme de jours juliens et secondes, soit sous la forme d'un numéro de séquence,
- la limite supérieure des données à extraire, au-delà de laquelle il ne faut plus échantillonner. Le programme signale qu'elle est atteinte au moyen d'un code.

### 5.4.- Travaux en cours

On procède actuellement à la transformation de la base de données (données météorologiques, données de la bouée, données des courantomètres, données des expériences communes) en une banque de données exploitable à l'aide du logiciel qui vient d'être décrit. Ce logiciel est par ailleurs en développement constant.

### 6.- Référence

MALTAIS, J. A nine channel digital magnetic tape format for storing oceanographic data, Woods Hole Oceanographic Institution, Ref. 69-55. Unpublished manuscript.

## Chapitre IV

### Echange des données

#### 1.- Introduction

Une autre fonction du centre d'exploitation des données est de fournir ces données à des institutions extérieures qui coopèrent avec les chercheurs du Programme dans le cadre d'expériences communes. Ces expériences sont destinées à nourrir les modèles mathématiques développés par différents groupes de recherche; le volume de données qu'elles fournissent impose que celles-ci soient échangées à l'aide de supports d'information directement traitables par ordinateur et portables d'ordinateur à ordinateur.

L'expérience de plusieurs années a montré les difficultés inhérentes à l'échange d'informations entre systèmes différents (voir le chapitre II). Pour éviter de semblables problèmes lors d'expériences futures, il est nécessaire de définir un format flexible et standardisé pour l'échange international de données entre différents systèmes de traitement. Ce format n'est pas destiné à traiter ou analyser des informations, mais à les transférer facilement d'un ordinateur à l'autre. Le but principal est donc de développer *un format standard pour l'échange international des données* d'un centre à un autre, tel qu'il soit utilisable par chacun des chercheurs dans tous les centres traitant les données.

L'apparition du format de données GATE pour l'échange international adopté par la COI, avec quelques modifications, sous le nom GF2 (*Generalized format 2*) a résolu une partie du problème. Cependant, peu

de centres océanographiques disposent du logiciel nécessaire pour échanger des données en GF2. Il apparaît donc du plus haut intérêt d'étudier un logiciel et ses possibilités d'adaptation aux différents systèmes existants. Le travail est considérable car le GF2 n'a jamais été utilisé pour l'échange de données océanographiques de nature chimique ou biologique, et la possibilité de l'adaptation de telles données au GF2 doit être étudiée de manière approfondie. Cependant une étude préliminaire a montré que la souplesse des structures préconisées est telle qu'il ne semble pas y avoir de difficultés de principe à l'incorporation de données biologiques, chimiques ou géologiques à un système d'échange basé sur GF2.

## 2.- Description générale du GF2

Les données peuvent être écrites ou retrouvées à partir du format au moyen de programmes très simples et assez rapides. Par ailleurs, la structure du format (enregistrements, fichiers, code et format d'écriture, ...) est destinée à être entièrement auto-descriptive et peut être entièrement automatisée si on le désire.

Les analogies entre le GF2 et le format de stockage précédemment décrit sont nombreuses. En fait, tous deux sont issus du même format, [Maltais (1969)] mais ont évolués différemment car ils ont été développés dans des buts différents. En effet, le format de stockage n'est utilisé que sur un seul ordinateur à la fois et est conçu pour être lu et écrit rapidement; ceci impose l'utilisation de code interne et l'écriture en enregistrements de longueur variable; le GF2 doit servir de lien entre ordinateurs différents; l'utilisation d'une représentation en code caractère et d'enregistrements de longueur fixe s'impose au prix d'une manipulation plus lente et plus lourde; tous deux ont en commun leur grande souplesse d'utilisation, leur capacité de s'adapter à n'importe quel type de données et leur caractère totalement auto-descriptif.

2.1.-

La structure de la bande consiste en six types d'enregistrements physiques séparés par des marques d'enregistrement et bloqués en fichiers (usuellement séparés par une marque de fin de fichier 'EOF', quelque fois appelée marque de bande), comme suit :

"fichier test"

EOF

### **enregistrement de début de bande**

EOF

## enregistrement de données

•

### **enregistrement de données**

EOF

### **enregistrement de données**

•

autant que nécessaire

#### **enregistrement de données**

EOF

enregistrement de fin de bande

EOF

EOF

le double EOF signifie la fin des données sur cette bande

Deux marques de fin de fichier 'EOF' à la fin des informations sur chaque bande, même dans les fichiers à bandes multiples, sont obligatoires. D'autres usages de 'EOF', excepté pour la séparation des fichiers et après l'enregistrement début de bande, sont interdits.

2.2.-

Tous les enregistrements physiques doivent être de longueur constante avec exactement 1920 caractères (24 groupes de 80 caractères) complétés par des blancs ou des '9' si nécessaire pour obtenir la longueur imposée. Les enregistrements de début de bande et d'en-têtes sont bloqués par groupes de 80 caractères de sorte que l'information nécessaire à la description de la bande et des en-têtes peut être facilement copiée à partir de (ou sur) cartes perforées si l'on le désire.

2.3.-

Le type d'enregistrement de la bande magnétique doit être identifié par son premier caractère :

- fichier test<sup>1</sup>,
- (0) enregistrement de début de bande,
  - (1) enregistrement "en-tête de fichier" type 1,
  - (2) enregistrement "en-tête de fichier" type 2,
  - (3) enregistrement de données,
  - (4) enregistrement de fin de bande.

Le sommaire du contenu de chacun de ces enregistrements est donné ci-dessous.

---

1. Le fichier test se reconnaît directement à son contenu, par son premier caractère, octal 77 (BCD), hexadécimal FF (EBCDIC) ou 11111111 (binaire).

Type	Nom	Contenu
	"Fichier test"	<p>Un enregistrement test consiste en 1920 caractères – test. Un caractère test est défini comme une suite de "1", l'octal "77" en BCD, l'hexadécimal "FF" en EBCDIC. C'est pourquoi le fichier test consiste en un nombre suffisant d'enregistrements test, composé chacun de 1920 caractères, pour remplir environ 20 mètres de la bande (protection contre une mise en place défectueuse).</p>
0	En-tête de la bande	<ul style="list-style-type: none"><li>- Identification de la bande, nom du pays et de l'Institution écrivant la bande et information de la succession des bandes.</li><li>- Date/heure de création de la bande et type d'ordinateur utilisé.</li><li>- Tables de translation (codes des caractères) utilisées par l'ordinateur écrivant la bande.</li><li>- Zone non attribuée, pouvant être utilisée pour des remarques ou explications en langage clair.</li></ul>
1	"En-tête de fichier" type 1	<ul style="list-style-type: none"><li>- Détails de la provenance des données (pays et institution) dans le fichier et nom du fichier.</li><li>- Information descriptive au sujet de la plateforme primaire qui a observé des données se trouvant dans ce fichier.</li><li>- Informations spéciales pour le cas où une plateforme secondaire supporte une plateforme primaire (pour un système de bouée, la bouée est la plateforme primaire tandis que le bateau effectuant le mouillage est la plateforme secondaire).</li><li>- Position, époque et valeur de la variation magnétique au début de la période d'observation ou de la station.</li><li>- Mêmes informations mais à la fin de la période d'observation ou de la station (incluant un code pour résumer la validité des résultats pour tout le fichier).</li></ul>

- Information sur l'environnement (météorologique et océanographique) des stations d'observation (utilisation optionnelle).
- Information sur le nombre de paramètres répétés une seule fois par enregistrement, nombre de paramètres par cycle de données, intervalle entre les échantillons, nombre de cycles de données dans le fichier, et indication de l'utilisation de nombres entiers ou réels (ou les deux) et de caractères alpha-numériques comme valeurs de paramètres.
- Spécifications détaillées du format pour l'enregistrement des 1920 caractères des enregistrements de données. Zone libre mais utilisable pour des remarques ou explications en langage clair.

2

"En-tête de fichier"  
type 2

- Spécifications du paramètre # 1 dans le cycle de données incluant un code résumant la validité des résultats pour le paramètre # 1 (pour tout le fichier) et définition d'un des deux attributs pour le paramètre # 1.
- Définition du second attribut et documentation sur la méthode de mesure et de collection du paramètre # 1.
- Idem pour les paramètres 2-12.

(Pour les cycles de données de 13 → 24 paramètres, on a besoin d'un deuxième enregistrement d'"en-tête de fichier" type 2, de 25 → 36, un troisième, etc.)

3

Enregistrement  
de données

- Nombre de cycles de données dans cet enregistrement, nombre de cycles de données précédant cet enregistrement et compteur d'enregistrement de données.  
(Note : Ces nombres *ne sont pas* les paramètres définis au point 7 de la description de l'en-tête de fichier type 1)  
Le paramètre d'enregistrement appartient aux données observées, et pas aux enregistrements de données eux-mêmes. Les caractères 1-20 de cet

enregistrement contiennent des explications pour la lecture des données sur la bande et pas pour l'identification des observations.

- Cycles de données analysées.

4

Enregistrement fin de  
bande 'EOT'

- Nom ou numéro de la bande qui suit.
- Caractères tous mis à '9' pour que cet enregistrement puisse être lu avec la même instruction de format que l'enregistrement en tête de fichier type 1.
- Zone libre mais utilisable pour des commentaires.

#### 2.4.-

Si un fichier est trop long pour une seule bande magnétique, il peut être continué sur d'autres bandes. L'enregistrement début de bande procure une information indiquant si le fichier est la suite d'un fichier se trouvant sur la bande précédente et l'enregistrement de fin de bande montre si le fichier se prolonge sur une nouvelle bande. L'enregistrement 'EOF' après l'enregistrement en début de bande se trouve sur toutes les bandes, mais les enregistrements "en-tête de fichier" ne sont pas répétés sur des bandes suites.

Chaque fichier de données est entièrement auto-décrit par des enregistrements "en-tête de fichier" placés au début de celui-ci. En particulier, l'enregistrement "en-tête de fichier" type 1 doit contenir des détails de format sur tous les enregistrements de données (en Fortran, ceci serait une instruction du format utilisé pour écrire un enregistrement de données). L'enregistrement en-tête de fichier type 2 contient des informations sur chacun des paramètres du jeu de données. Un fichier peut être défini comme étant toutes les observations individuelles qui constituent un ou des jeux de données.

#### Exemples

- \* Chaque profil STD, si l'information dans les enregistrements de tête change pour chaque profil.

\* Tous les profils STD reçus d'un bateau pour une période de temps donnée si l'information contenue dans les enregistrements de tête est toujours la même, et si les enregistrements de données contiennent toutes les informations qui changent à chaque profil (c'est-à-dire, latitude, longitude, temps, etc.).

\* Un bref aperçu de la température à la surface de la mer en provenance d'observations aériennes.

\* Une série temporelle qui provient d'un courantomètre.

#### 2.5.-

Chaque enregistrement de données consiste en un nombre entier de cycles de données (c'est-à-dire un cycle de données ne peut pas chevaucher deux enregistrements physiques). Par exemple, tous les instruments d'échantillonnage prennent des échantillons selon des séquences régulièrement répétées; la répétition d'une séquence de base constitue un cycle de données.

#### Exemples

\* (profondeur ou pression et température) à partir d'un XBT ou MBT.

\* (temps, pression, température, humidité, élévation, vents) à partir d'une sonde atmosphérique.

\* (temps, direction, direction, direction, vitesse) à partir d'un instrument de mesure pour les courants océaniques (Ceci arrive quand l'instrument de mesure échantillonne un paramètre beaucoup plus souvent qu'un autre; ici, c'est le cas de la direction du courant).

\* (latitude, longitude, temps, profondeur, température, salinité) permettant d'établir une carte horizontale de température et de salinité.

Les paramètres peuvent ou non être tous échantillonés simultanément. Si l'échantillonnage est séquentiel, comme dans l'exemple du courantomètre, à chaque paramètre est associé un déphasage par rapport

au temps spécifié dans le cycle de données et ces déphasages doivent être spécifiés dans l'enregistrement en-tête du fichier type 2. En général, la pression ou quelqu'autre variable peut être la variable indépendante, et cette variable doit être le premier élément du cycle de données. Les déphasages spécifiés dans l'enregistrement en-tête du fichier type 2 ont alors les mêmes unités que la variable indépendante. Si les déphasages sont tous nuls, le premier paramètre dans un cycle de données ne doit pas être nécessairement une variable indépendante.

2.6.-

En Fortran, il est recommandé d'utiliser uniquement les formats A, I, F et X. Les formats E et D ne sont pas toujours compatibles entre les différentes installations. Les facteurs d'échelle des paramètres (voir enregistrement en-tête du fichier type 2) doivent être utilisés pour éviter les formats E et D et pour faciliter la conversion des unités d'enregistrement aux unités communes.

2.7.- Il faut bien veiller à ce que les longueurs de mots des ordinateurs utilisés soient suffisantes pour la résolution que nécessite chaque paramètre. De nouveau, l'usage des facteurs d'échelle peut résoudre ce problème.



## Conclusion

L'équipe chargée de la compilation et du traitement des données a rassemblé une importante base de données, dont la plupart proviennent d'instruments automatiques. Elle développe le logiciel nécessaire pour exploiter les données et les échanger avec d'autres équipes de recherche dans le cadre d'expériences communes. A cet égard, cette équipe a été choisie par les responsables du projet INOUT (partie de JONSDAP 76), sur base de l'expérience et de la compétence qu'elle a acquise dans ce domaine, pour réaliser la circulation des informations de ce projet, dont le centre de données (INOUT Data Center) sera pour lors situé à Liège. Le format GF2 a été proposé pour simplifier les problèmes de l'échange, et le logiciel adéquat pour l'utilisation pratique de ce format sera également réalisé à Liège.

2.5% of the total number of individuals in each group. This information will then be used to calculate the proportion of each group that has been sampled.

Achevé d'imprimer le 20 mai 1976  
sur le duplicateur et la presse offset  
d'é.t.a.b.é.t.y.p.  
72, quai de Longdoz - 4020 LIEGE  
Tél.: 041/42.59.21

