

Chapitre I

Exploitation des données et modelage du système marin

par

Jacques C.J. NIHOUL
Georges PICHOT et Jean-Henri HECQ
Daniel M. DUBOIS

1.- Introduction

Le modèle mathématique se veut être un outil préictionnel permettant de prévoir, à partir d'une situation soigneusement répertoriée, l'*évolution* du système marin dans une région test choisie. Cette préiction est particulièrement indispensable lorsque des modifications subites du système (nouveaux rejets côtiers, déversements accidentels ou organisés au large, conditions météorologiques exceptionnelles, ...) mettent en question sa stabilité, son pouvoir récupérateur, la survie de certains de ses éléments et posent les questions préoccupantes de la pollution et de ses incidences sur la nourriture et la santé de l'homme.

L'inventaire du système marin étudié par les nombreuses équipes du Programme est clairement présenté et discuté dans les chapitres qui suivent. Il a paru intéressant d'examiner pour commencer la façon dont cet inventaire peut être intégré au modèle, de faire un premier bilan des techniques de prédition dont ce modèle dispose et des résultats qu'il peut atteindre dès à présent.

2.- Variables d'état et paramètres

Pour décrire l'état du système marin à un instant quelconque, il est nécessaire de pouvoir chiffrer un certain nombre de grandeurs caractéristiques ou *variables d'état*. Pour chacune de ces variables une équation d'évolution est écrite qui permet à l'ordinateur de prédire son évolution temporelle et spatiale connaissant sa valeur initiale aux différents points du réseau et sa valeur en chaque instant aux frontières de celui-ci.

Dans ces équations interviennent inéluctablement des paramètres déterminés empiriquement ou par des modèles auxiliaires. Certains de ces paramètres appartiennent à la physico-chimie même du problème (diffusivités turbulentes, vitesses de réaction, ...), d'autres surgissent de la *fermeture* du système d'équations d'évolution, c'est-à-dire de la limitation imposée pour des raisons pratiques au nombre des variables d'état. (Ainsi la transparence de l'eau peut, dans un modèle simple, intervenir comme un paramètre ajusté empiriquement et, dans un modèle plus élaboré, être ajoutée aux variables d'état et être calculée notamment en fonction des concentrations de suspensions, de plancton, etc.)

Il est convenant de diviser les variables d'état en deux grandes catégories : les variables *entraînantes* et les variables *entraînées*.

Les variables entraînantes sont celles qui interviennent dans l'hydrodynamique du problème, ce sont :

- (i) les trois composantes de la vitesse : v_1 , v_2 , v_3 ;
- (ii) la pression p ;
- (iii) la masse spécifique de l'eau de mer ρ .

Les variables entraînées sont essentiellement de deux types :

- (i) des masses spécifiques telles que la salinité, les masses spécifiques de tel ou tel polluant en solution, en suspension, ..., les biomasses, etc.
- (ii) des variables de nature *thermodynamique* – le mot étant utilisé dans un sens élargi – telles que la température, les indices de diversité, etc.

Il est à noter que les variables entraînées peuvent ne l'être que partiellement et posséder un mouvement propre tel que la migration ver-

ticale des particules qui sédimentent ou la migration horizontale de certaines espèces vivantes.

Les variables entraînées sont notées ρ_a étant entendu que, s'additionnant aux cinq variables entraînantes, l'indice a prendra les valeurs 6, 7, 8, ... jusqu'à épuisement du nombre total de variables d'état.

3.- Équations d'évolution

Les équations d'évolution ont été exposées dans le premier rapport de synthèse et il n'est pas nécessaire de les reproduire ici. Il est utile cependant de rappeler que ce sont des équations aux dérivées partielles de la forme générale (\mathcal{D} désignant un opérateur aux dérivées partielles, faisant intervenir éventuellement certains paramètres) :

$$(1) \quad \mathcal{D}_1(v_1, v_2, v_3, p, \rho) = 0$$

$$(2) \quad \mathcal{D}_2(v_1, v_2, v_3, p, \rho) = 0$$

$$(3) \quad \mathcal{D}_3(v_1, v_2, v_3, p, \rho) = 0$$

$$(4) \quad \mathcal{D}_4(v_1, v_2, v_3, p, \rho) = 0$$

$$(5) \quad \mathcal{D}_a(v_1, v_2, v_3, p, \rho, \rho_a) = Q_a - \nabla \cdot M_a + \sum_{\beta=1}^5 P_{a\beta} \quad (\alpha=6, 7, 8, \dots)$$

Dans les membres de droite des équations (5) pour les variables entraînées apparaissent trois types de termes :

- (i) les termes de source Q_a supposés connus (ces termes pourraient correspondre par exemple aux déversements en mer);
- (ii) les termes de migrations $\nabla \cdot M_a$ (ces termes pourraient correspondre par exemple à la sédimentation);
- (iii) les termes d'interactions chimiques, biologiques et biochimiques $P_{a\beta}$ par lesquels la variable ρ_a peut être produite ou détruite par interaction avec la variable ρ_β .

On note immédiatement que, si on dispose d'une équation d'évolution pour chaque variable entraînée, on n'a que quatre équations pour les cinq variables entraînantes. Le système doit être complété d'une équation d'état qui lie la masse spécifique ρ à la pression et à certaines des

variables entraînées, température, salinité, etc.

$$(6) \quad \rho = \rho(p, \rho_a) .$$

C'est par l'addition de cette équation d'état qu'intervient *en général* un *couplage* entre variables entraînantes et entraînées.

Des études expérimentales et théoriques [Nihoul (1972b)] ont montré cependant que dans le Southern Bight, choisi comme région test pour le programme, le mélange des eaux est suffisamment efficient pour qu'on puisse admettre, avec une très bonne approximation, que dans les équations (1) à (4), ρ est une constante [$\rho = C^{\text{te}}$ est la forme la plus simple de (6)]. Cette approximation semble satisfaisante également pour les problèmes de grande circulation en mer du Nord, en tout cas pendant la période d'hiver. Elle n'est pas valable pour l'étude d'estuaires stratifiés comme celui de l'Escaut et Jäger (1972) a examiné les conséquences de cette situation sur les équations d'évolution et les possibilités d'élaboration d'un modèle d'estuaire applicable à l'Escaut. Cette étude a fait apparaître plusieurs questions importantes qui vont être examinées en collaboration avec les équipes du Ministère des Travaux Publics et il a paru prématûr d'inclure cette contribution dans le présent rapport de synthèse. Les modèles qui seront mentionnés dans le chapitre *Estuaires* seront par conséquent volontairement limités à des modèles parfois très simplifiés servant plutôt à l'interprétation des résultats chimiques et biologiques qu'à la prédition mathématique.

Une importante simplification apparaît dans le système (1-5) lorsque ρ peut être pris constant. En effet, les équations (1) à (4) sont alors découplées des autres et peuvent être résolues séparément. Il est alors possible de déterminer d'abord le champ de vitesse et, ensuite, substituant les valeurs trouvées dans (5), de calculer son influence sur l'entraînement (advection, dispersion, ...) des variables ρ_a .

Le cas le plus simple se rencontre lorsque le membre de droite de (5) est nul ou se réduit à des termes de sources connus. Ce cas est en particulier celui d'un polluant dont les migrations et les interactions sont négligeables et dont le comportement en mer est gouverné par la seule hydrodynamique. C'est le cas des traceurs tels que la rhodamine B. Outre son importance pratique, il est instructif du point de vue fonda-

mental car il permet de comprendre les mécanismes d'entraînement et de mettre au point les techniques numériques appropriées aux équations de diffusion du type (5) en comparant notamment les résultats des simulations numériques aux expériences de dispersion réalisées *in situ* avec les traceurs.

Les modèles consacrés à la détermination du champ de vitesse et à l'étude de la dispersion d'un polluant passif ou traceur portent en général le nom de modèles hydrodynamiques.

Dans une première étape, les modèles hydrodynamiques sont généralement simplifiés en intégrant les équations (1) à (4) sur la profondeur. Ces équations sont alors remplacées par trois nouvelles équations aux dérivées partielles

$$(7) \quad \Delta_1(V_1, V_2, \zeta) = 0$$

$$(8) \quad \Delta_2(V_1, V_2, \zeta) = 0$$

$$(9) \quad \Delta_3(V_1, V_2, \zeta) = 0$$

où V_1 et V_2 sont les composantes horizontales du débit et ζ l élévation de la surface libre.

Le chapitre II est consacré aux résultats des modèles hydrodynamiques pour les marées, les courants résiduels et la dispersion des traceurs. Une attention particulière est prêtée à cet endroit à l'effet de diffusion par cisaillement que l'intégration sur la profondeur met bien en évidence [Nihoul (1971b)]¹.

Dans le cas général, la résolution des équations (5) requiert la connaissance des termes d'interactions.

Cette connaissance n'est encore aujourd'hui que fragmentaire et en particulier de nombreux paramètres interviennent dont les valeurs *in situ* ne sont pas déterminées avec une précision suffisante. La troisième phase du programme mettra l'accent sur la formulation des termes d'interaction.

1. Les travaux du chapitre II ont été menés en étroite collaboration avec les équipes canadiennes dans le cadre d'un programme d'échanges bilatéraux dont les excellents résultats réclament qu'il se poursuive et s'amplifie. Il est agréable d'exprimer ici la gratitude des chercheurs belges au Gouvernement et aux spécialistes canadiens.

A cet effet, on pourra attendre énormément d'une *approche fondamentale* au problème : relevé de ce qui est rapporté dans la littérature, expériences de laboratoire, réflexions théoriques, ...

La nécessité de cette approche a été très bien comprise par les différentes équipes du programme et l'on voit dans les chapitres qui suivent les résultats chimiques et biologiques rapportés avec un souci constant de faire apparaître des corrélations intuitives, des relations interdisciplinaires, des réflexions sur la représentativité et l'importance des variables étudiées et des valeurs mesurées, l'énoncé déjà de premières lois, ...

L'étude fondamentale, en cherchant à comprendre et à formuler les mécanismes d'interactions, contribue également à l'amélioration constante du modèle en faisant apparaître la nécessité d'ajouter de nouvelles variables d'état ou l'intérêt de redéfinir certaines d'entre elles pour mieux les enracer dans la réalité ou mieux les intégrer au modèle.

C'est le cas, en particulier, de variables comme l'indice de diversité qui peut être un élément essentiel d'un modèle préditionnel mais qu'il importe de définir avec soin afin d'en assurer le *réalisme biologique et mathématique*, la qualité d'une représentation adéquate de la vie et d'une donnée adaptée au traitement sur ordinateur.

On trouvera, en appendice de ce chapitre, quelques réflexions sur l'indice de diversité dont l'intention est d'abord d'ouvrir le débat en montrant sous quel angle on peut repenser le problème et quels résultats on peut espérer.

L'étude fondamentale n'est cependant pas suffisante pour le modèle mathématique qui doit interpréter et prédire l'océan réel. Les données de la littérature sont trop fragmentaires et les expériences de laboratoire ont le défaut de leur qualité, en cela qu'elles isolent (pour les mieux étudier) un nombre limité d'interactions, séparées le plus souvent du contexte complexe de la mer dont l'éprouvette ou l'aquarium sont souvent de mauvaises images. Le modèle mathématique doit s'assurer qu'une interaction, dont la réalité est peut-être indiscutable, est bien formulée en termes des meilleures variables et tient compte des paramètres essentiels *in situ*, qu'elle n'est pas inhibée dans l'océan par des facteurs physiques, chimiques et biologiques imprévus, que d'autres interactions

plus importantes en mer ne la supplantent pas dans la hiérarchie des priorités du modèle.

A cet effet, une méthode automatique d'étude des corrélations a été mise au point, capable de discerner, à partir de séries de mesures, et à des seuils de confiance que l'on peut choisir aussi sévères que l'on souhaite, des relations entre variables différentes témoignant d'interactions de ces variables entre elles ou avec une même troisième.

Cette méthode a la puissance de l'ordinateur qui peut examiner très rapidement toutes les possibilités que des études fondamentales mettraient des années à envisager. Elle a, de l'ordinateur, le défaut d'irréflexion et le spécialiste est indispensable pour interpréter, découvrir ou rejeter les corrélations que la machine discerne.

L'étude des interactions se fera par conséquent par l'étroite collaboration de la réflexion fondamentale et du traitement automatique des données.

Un exemple d'application est donné à la section suivante.

4.- Exploitation des données par corrélations croisées

Dans le cadre des campagnes de mesures, une station fixe a été effectuée pendant 72 heures, du six au neuf décembre 1971, au point M6 de coordonnées $51^{\circ}28'25''$ N - $03^{\circ}09'15''$ E situé dans la zone d'influence de l'estuaire de l'Escaut.

Des prélèvements ayant été réalisés toutes les trois heures, on possède — du moins pour les valeurs en surface — de bonnes séries temporelles des variables reprises au tableau 1. Un traitement statistique a été entrepris pour déceler les corrélations qui les lient et établir le graphe des interactions possibles.

La méthode utilisée est celle des fonctions de corrélation croisée. Soient deux séries temporelles discrètes

$$X_i = X_i(t) \quad i = 1, 2$$

dont les moyennes sont

$$E[X_i(t)] = \mu_i .$$

Tableau 1

Liste des variables traitées

$x_1(t)$	concentration des nitrites en solution	$\mu\text{g N/l}$
$x_2(t)$	concentration des nitrates en solution	$\mu\text{g N/l}$
$x_3(t)$	concentration des phosphates en solution	$\mu\text{g P/l}$
$x_4(t)$	concentration des silices en solution	ppm
$x_5(t)$	concentration de chlorophylle A	mg/m^3
$x_6(t)$	concentration de phéopigments a	mg/m^3
$x_7(t)$	concentration de chlorophylle a	mg/m^3
$x_8(t)$	température de l'eau	$^\circ\text{C}$
$x_9(t)$	nombre de bactéries (germes totaux)	col/ml
$x_{10}(t)$	nombre de cnidaires	
$x_{11}(t)$	nombre de copépodes	
$x_{12}(t)$	nombre de chaetognathes	
$x_{13}(t)$	nombre de mysidacés	
$x_{14}(t)$	nombre de schistompsis spiritus (espèce A)	
$x_{15}(t)$	nombre de schistompsis kervillei (espèce B)	
$x_{16}(t)$	nombre de mesopodopsis slabberi (espèce C)	
$x_{17}(t)$	nombre de gastrosaccus sanctus (espèce D)	
$x_{18}(t)$	nombre de gastrosaccus spinifer (espèce E)	
$x_{19}(t)$	indice de diversité de $x_{10}, x_{11}, x_{12}, x_{13}$	
$x_{20}(t)$	indice de diversité de $x_{14}, x_{15}, x_{16}, x_{17}, x_{18}$	

Les variables $x_{10}, x_{11}, \dots, x_{18}$ sont en nombre d'individus rapportés à 2000 tours de flow-meter.

On calcule successivement :

- les fonctions d'autocovariance

$$(1) \quad \gamma_{ii}(u) = E\{[X_i(t) - \mu_i][X_i(t+u) - \mu_i]\},$$

- les fonctions de covariance croisée

$$(2) \quad \gamma_{ij}(u) = E\{[X_i(t) - \mu_i][X_j(t+u) - \mu_j]\},$$

- les fonctions de corrélation croisée

$$(3) \quad \rho_{ij}(u) = \frac{\gamma_{ij}(u)}{[\gamma_{ii}(0) \gamma_{jj}(0)]^{\frac{1}{2}}}$$

avec $|\rho_{ij}(u)| \leq 1$.

Si $\rho_{ij}(u)$ est nul pour tout u , les deux processus i et j sont complètement découpés; s'il est non nul pour u nul seulement, ils ne sont corrélés qu'en simultanéité. Si, pour un u donné, ρ_{ij} est non nul, le comportement de j à un instant t est lié à celui de i à l'instant $t - u$; si, de plus, il est supérieur à un ρ_{ij} de référence, fonction de u et du niveau de confiance choisi, cette liaison est statistiquement significative. Enfin ρ_{ij} peut être positif ou négatif suivant qu'il y a liaison *de pic à pic* ou *de pic à creux*.

Une amélioration a été apportée à cette méthode classique. En effet, si les deux séries (ou même l'une des deux) sont fortement autocorrélées, ces corrélations internes peuvent induire des corrélations croisées parasites. Lorsque ce cas s'est présenté, on a filtré les séries pour ne correler entre eux que leurs reliquats désautocorrélés.

La méthode a été appliquée aux vingt variables disponibles. Les quatre cents résultats sont résumés au tableau 2. En suivant la $i^{\text{ème}}$ ligne du tableau, on lit immédiatement l'influence des j variables sur la $i^{\text{ème}}$: le signe plus ou moins indique celui de la corrélation, le chiffre avant le point, le retard d'influence (soit $u = k \Delta t$ avec $k = 1, 2, 3, \dots$ et $\Delta t = 3$ heures) et les chiffres après le point, le niveau de confiance au-dessus duquel la corrélation a été décelée.

Pour interpréter ce tableau, il faut se rappeler qu'une corrélation décelée entre deux processus signifie que le comportement de l'un à l'instant t présente des similitudes avec celui de l'autre à $t + u$ et malheureusement pas qu'il y a corrélation *de cause à effet* entre le premier et le second.

Par exemple, ils peuvent être soumis à l'influence *réelle* d'un même troisième. C'est probablement ce qui se passe dans le cas traité, où les variables sont influencées, en fait, par la marée responsable d'une émission d'eaux d'estuaire, chargées de nutriments et d'espèces planctoniques caractéristiques. En effet, sur les seize corrélations croisées avec décalage de douze heures, quatorze sont positives.

Tableau 2

	1	2	3	4	5	6	7	8	9	10
1	-	+ 0.001	+ 0.02 + 4.02	-	+ 2.05	-	-	-	-	-
2	+ 0.001	-	-	-	-	-	-	-	-	-
3	+ 0.02	-	+ 4.02	-	-	-	-	-	-	- 0.01
4	- 2.01	- 2.05	-	-	-	-	-	-	-	-
5	-	-	-	-	- 1.02	+ 0.01 - 1.05	-	-	-	-
6	-	-	-	-	+ 0.01 - 1.05	-	- 0.02	-	-	-
7	-	-	-	-	-	- 0.02	-	-	-	-
8	- 1.05 - 4.02	-	- 1.05	-	-	-	-	-	-	-
9	-	-	-	-	-	-	- 2.01	-	-	-
10	-	-	- 0.01	-	-	-	-	+ 3.05	-	-
11	- 4.05	-	-	-	-	-	-	- 1.01	-	-
12	-	-	-	-	-	-	-	- 1.02	-	+ 4.05
13	- 3.05	-	-	-	-	-	-	+ 2.001	+ 4.05	-
14	-	-	-	-	-	-	-	+ 2.01	-	-
15	- 3.05	-	-	-	-	-	-	+ 2.01	+ 4.05	-
16	-	- 1.05	-	-	-	-	-	+ 3.001	+ 5.02	+ 0.05
17	-	- 1.05	+ 0.05	-	-	-	-	-	-	-
18	+ 0.05	+ 0.05	+ 0.01	-	-	+ 2.05	-	-	-	-
19	-	-	-	-	-	-	-	+ 1.05	-	0.05
20	- 1.05	-	-	-	-	-	-	- 2.02	-	-

Tableau 2

11	12	13	14	15	16	17	18	19	20	
-	-	-	-	-	-	+ 1.05	+ 0.05	-	-	1
-	-	+ 3.05	-	+ 3.05	+ 2.05	+ 1.01	+ 0.05	-	-	2
-	-	-	-	-	-	+ 0.05	+ 0.01	-	-	3
-	-	-	-	-	-	-	-	-	-	4
-	-	-	-	-	-	-	-	-	-	5
-	-	-	-	-	-	-	-	-	-	6
-	-	-	-	-	-	-	-	-	-	7
-	-	-	-	-	-	-	-	-	+ 2.05	8
-	-	-	-	-	-	-	-	-	-	9
-	-	+ 1.05	-	+ 1.05	+ 0.05	-	-	+ 0.05	-	10
-	+ 0.001	+ 0.01	+ 1.001	+ 1.01	+ 3.05	+ 2.05	-	- 0.001	- 1.05 + 3.02	11
+ 0.001 + 1.05	-	-	+ 5.05	-	+ 4.001	-	-	-	-	12
-	-	+ 4.02	+ 0.001	+ 0.001 + 4.02	-	-	+ 4.05	-	+ 0.01	13
-	-	+ 0.001 + 4.05	-	+ 0.001 + 4.05	-	+ 0.02	+ 0.05	-	- 0.05	14
-	-	+ 0.001 + 4.02	+ 0.001	+ 4.02	-	-	+ 4.05	-	- 0.01	15
-	-	+ 1.001	+ 1.01	+ 1.001	-	-	-	+ 2.05	-	16
-	-	-	+ 0.02	-	-	-	-	-	-	17
-	-	-	+ 0.05	-	-	+ 1.01	+ 1.01	-	-	18
- 0.001	-	-	- 1.02 - 2.05	-	-	- 2.01	-	-	-	19
-	-	+ 0.01	- 0.05	- 0.01	-	-	-	-	-	20

L'examen global du tableau 2 peut conduire aux remarques suivantes.

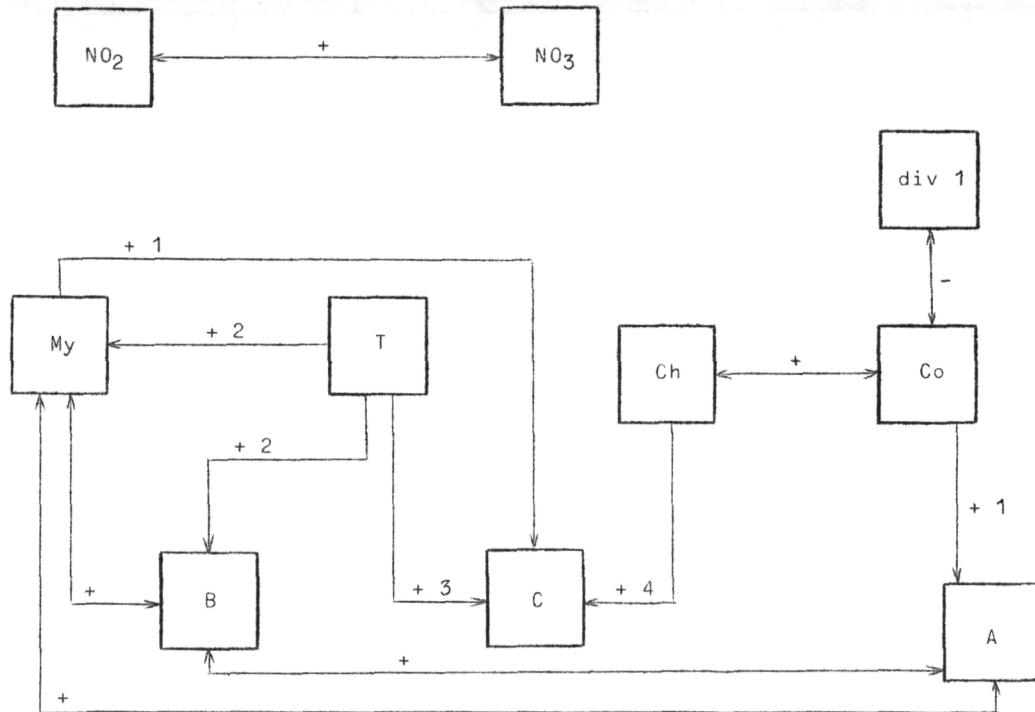
(i) La silice dissoute n'influence aucune autre variable; si cette propriété devait être confirmée aux autres points du modèle, on pourrait songer à l'utiliser en mer comme traceur aussi inerte et même plus sensible que la salinité.

(ii) Il est étonnant de constater que la température de l'eau influence si fortement les variables zooplanctoniques. L'évolution de la température joue-t-elle le rôle d'horloge pour mettre en évidence le cycle nyctéméral du plancton ou est-elle l'indice de cette légère différence de température entre l'eau de l'Escaut et celle de la mer ?

(iii) Il n'y a pratiquement aucune corrélation entre les trois valeurs des concentrations pigmentaires, mesures de la biomasse phytoplanctonique et les onze valeurs relatives au zooplancton. De fait, en période hivernale, les rapports à l'intérieur de ce biotope sont réduits au minimum.

Au niveau de confiance 10^{-3} , les corrélations suivantes sont mises en évidence (tableau 3).

Tableau 3
Corrélations au seuil 10^{-3}



(i) Corrélation (NO₂, NO₃)

Dans une eau chargée de nutriments, il y a présence simultanée de nitrites et de nitrates.

(ii) Corrélations (My,A), (My,B), (My,C), (A,B)

Ces corrélations sont évidentes puisque les espèces A, B et C font partie du groupe des Mysidacés et qu'elles y sont prédominantes.

(iii) Corrélation (A,Co)

Cette corrélation avec décalage de trois heures est un indice de la succession de masses d'eau différentes : l'eau de l'estuaire dont l'espèce A des Mysidacés est l'indicateur privilégié est suivie par une eau du large caractérisée par les copépodes.

(iv) La présence des copépodes est liée à celle des chaetognathes et à une abondance de ces copépodes correspond une diminution de l'indice de diversité de l'ensemble (Co, Cn, Ch, My). Notons ici que cet indice semble être commandé par le groupe le plus nombreux et ne joue pas à cet égard son rôle de tampon.

(v) La corrélation (Ch,C) avec décalage de 12 heures ne trouve pas d'explication satisfaisante.

Au niveau de confiance 10^{-2} , on note *en plus* les corrélations suivantes (tableau 4).

(i) De nombreuses corrélations déjà signalées au niveau 10^{-3} sont confirmées. Ainsi les corrélations (My,Co), (My,B), (PO₄,Cn) et (PO₄,E) sont les indices du même phénomène qu'en [(iii), 10^{-3}] et (My, div 2), (B,div 2) qu'en [(iv), 10^{-3}].

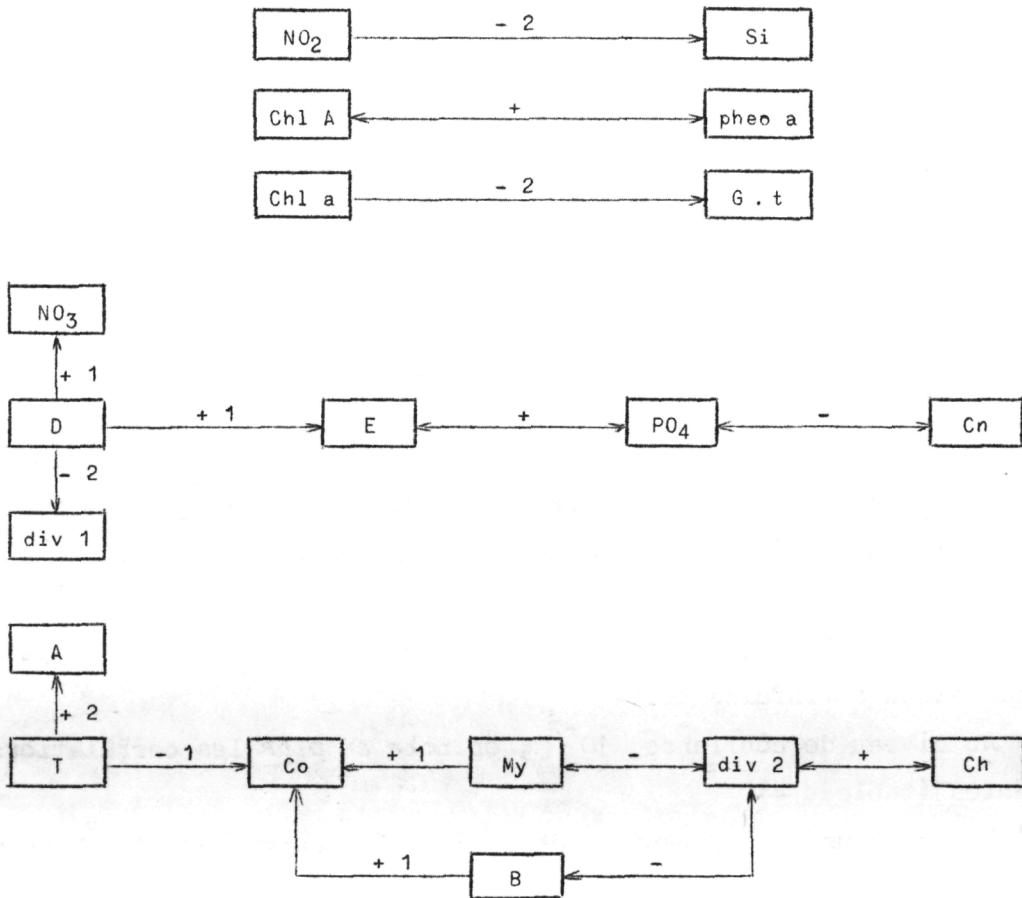
(ii) Corrélation (NO₂, Si)

Cette corrélation négative, avec décalage de 6 heures, montre que les nitrites et les silices sont indicateurs d'eau d'estuaire. Si les seconds en avaient été une caractéristique aussi nette que les premiers, la corrélation aurait été positive et simultanée.

(iii) Corrélation (chl A, phéo a)

Cette corrélation est bien claire, puisque les deux variables sont des mesures de la même biomasse phytoplanctonique, la seconde étant cependant plus générale et englobant le phytoplancton photosynthétiquement inactif.

Tableau 4
Corrélations au seuil 10^{-2}



(iv) On n'a pas de justification immédiate pour les corrélations (chl a, G.t.), (D, NO_3) et (D, div 1).

En conclusion, étudié dans des conditions hivernales, l'écosystème test semble particulièrement inactif. L'évolution des variables et leurs interactions semblent commandées principalement par les marées, le mélange des eaux, la turbulence, c'est-à-dire par l'hydrodynamique.

L'analyse des corrélations croisées est une première étape dans l'établissement du modèle mathématique d'un écosystème. Les suivantes sont la précision des relations d'interactions entre variables d'état. Le passage d'une étape à l'autre est une démarche délicate qui ne pourra se réaliser qu'étayée d'une réflexion écologique approfondie.

Références

- BARBETTE (J.), (1971). Etude des variations dans les populations micro-biennes au point 0006 du 6 décembre 1971 au 9 décembre 1971, *Technical Report/Point fixe : Bact. 01.*
- ELSKENS (I.) et JANSSEN (D.), (1971). Dosages phosphates, *Technical Report/Point fixe : Chim. 02.*
- ELSKENS (I.) et JANSSEN (D.), (1971). Dosages nitrate et nitrite, *Technical Report/Point fixe : Chim. 02.*
- ELSKENS (I.) et JANSSEN (D.), (1971). Dosages silice dissoute, *Technical Report/Point fixe : Chim. 03.*
- GOLDMAN (C.R.), GERLETTI (M.), JAVORNICKY (P.), MELCHIORI-SANTOLINI (U.) et DE AMEZAGA (E.), (1968). Primary Productivity, Bacteria, Phyto- and Zooplankton in Lake Maggiore : Correlations and Relationships with ecological factors, *Mem. Ist. Idrob.*, 23, 49-127.
- GRANGER (C.W.J.) et HATANAKA (M.), (1964). *Spectral Analysis of Economic Time Series*, Princeton University Press, Princeton.
- HECQ (J.H.) et HEYDEN (D.), (1971). Etude du Zooplancton de surface au point M06 pendant 72 heures, *Technical Report/Point fixe : Biol. 01.*
- JAËGER (Ph.), (1972). *Modèle mathématique de l'estuaire de l'Escaut*, Programme National Belge sur l'Environnement Physique et Biologique, N19.
- JENKINS (G.M.) et WATTS (D.G.), (1968). *Spectral Analysis and its applications*, Holden-Day.
- NIHOUL (J.C.J.), (1971a). *Ecosystemology applied to sea pollution*, Programme National Belge sur l'Environnement Physique et Biologique, N6.
- NIHOUL (J.C.J.), (1971b). *Shear effect in the Southern Bight*, Programme National Belge sur l'Environnement Physique et Biologique, N 5, North Sea Conference, Aviemore,
- NIHOUL (J.C.J.), (1972a). *Hydrodynamical and biochemical state variables and evolution equations for the mathematical modelling of sea pollution* in *Proc. Liège Third Coll. on Ocean Hydrodynamics*, Liège, May 3-8, 1971.
- NIHOUL (J.C.J.), (1972b). *Mathematical Models in North Sea Science*, M.I.T. University Press.
- STEVAERT (J.) et VAN BEVEREN (Ch.), (1971). Etude des variations pigmentaires à la station M06, *Technical Report/Point fixe : Biol. 02.*

Appendice : A statistical approach to ecosystems and the definition of a diversity index

Ecosystems are subject to the law of conservation of matter and are open systems with respect to energy, usually light energy. Materials must cycle throughout the ecosystem following the general laws of energy transformation. High quality chemical energy degrades to lower quality energy under the influence of catalysts which control the rates of the various processes. We want to review briefly the response of an ecosystem to different variations of materials in the environment, of the conditions affecting the rates of processes and of the transport of energy. Physiological adaptability allows individual organisms to cope with variable conditions affecting the use of energy.

Variations in the activity or size of individuals imply variations in energy transformations and in energy storage. Due to time lag, variations in the number of individuals in the various species cause oscillations in the size of populations; the regrowth of a population may lag behind the reappearance of the energy supply. The ecosystem self-organizes with compensating variations when a decrease in the abundance of some chemical energy may be accommodated by an increase in the abundance of others.

In general, an ecosystem self-organizes so as to reduce environmental variations (like Lenz's law in electricity).

A necessary condition for the stability of an ecosystem is the balance condition. An unbalanced system will vary until it becomes balanced; the ecosystem must periodically re-establish its own cycle structure in changing its distribution of materials. In certain conditions, the ecosystem can be highly nonlinear and an ecosystem can be very stable even with great fluctuations. Reduction in the oscillation of one species must be compensated by an increase in the fluctuations of some other species. The change in number arising from population oscillations will be reflected in corresponding change of energy.

However, this change of energy can be minimized if the organisms of these oscillating populations are simple and can follow the variable energy supply without too much time delay.

Another stabilizing mechanism in an ecosystem is the development of multifunctionality with resultant increase in the number of possible pathways of energy flow in the food chain. The efficiency of an ecosystem

increases with multifunctionality (parallel processing) and in these conditions the stability of all the system is better because the distribution of the flow of resources is made over a number of different lines.

1.- The food web

Consider the pathway of a food web. Let S_i be the i^{th} species and let F_i be the energy flow to the species S_i , per individual, the number of individuals of which is N_i .

We assume that the population of each species varies directly with the food energy available to that species; the amount of energy entering the community at the lowest trophic level does not vary with time (we take an average value of the energy over a sufficiently long time). Moreover, the energy retained by a species is assumed time independent.

Three macroscopic properties which appear frequently in the literature, the total number of individuals N , the total biomass B and the total energy flow F , are defined at equilibrium by

$$(1) \quad N = \sum_i N_i$$

$$(2) \quad B = \sum_i B_i N_i$$

$$(3) \quad F = \sum_i F_i N_i$$

2.- Diversity-stability of an ecosystem

It is of great interest to know how food webs might be designed so as to increase their stability and how stability based on these design features is related to the diversity of species in the ecosystem.

First, we want to examine the way in which species diversity can be calculated. Two methods are in general use. On one hand, when all the members of a collection can be identified and counted, the diversity of species may be given by Brillouin's formula :

$$(4) \quad H_B = \log \left[\frac{N!}{\prod_i (N_i !)} \right]$$

On the other hand, if P_i is the proportion of the i^{th} species in the population, i.e. $P_i = \frac{N_i}{N}$, the diversity can be written using the Shannon and Weaver formula :

$$(5) \quad H = - \sum_i P_i \log P_i$$

Second, the problem of stability and its relation to diversity has been discussed by a number of authors. In particular, Mac Arthur has proposed an index of stability I_S based on the fraction of energy reaching the top of the web along each distinct pathway; this is given by the quantity :

$$(6) \quad I_S = - \sum_k q_k \log q_k$$

where q_k is the fraction of total energy reaching the top of the web along each distinct pathway. Notice that I_S is highest when q_k is the reciprocal of the number of possible pathways and it may be the same for food webs involving fewer species with many connections as for those with many species with few connections.

Odum has suggested that stability increases with the amount of choice which the energy has in following the paths up through the food web.

3.- A new index of diversity-stability

We propose a new index of diversity-stability. At each species node, S_i , we define a quantity

$$(7) \quad T_i = \frac{(N_i + L_i)!}{(N_i!)(L_i!)}$$

where L_i is proportional to the number of links reaching the i^{th} species. A more precise definition of L_i will be given below. The diversity at each node is then written in the logarithmic form :

$$(8) \quad D_i = \log T_i$$

The function T_i depends upon two variables N_i and L_i ; its maximum is reached when $N_i = L_i$ and two minima are obtained when N_i or L_i goes to zero.

The total index of diversity-stability is given by

$$(9) \quad D = \sum_i \log T_i = \log \prod_i \frac{(N_i + L_i)!}{(N_i!)(L_i!)}$$

since the logarithm of a product is equal to the sum of the logarithms of the factors in the product.

We assume that the length of time that energy is retained by one species before being passed on to the next is time independent.

Our purpose is to calculate the most diversified distribution of N_i in maximizing D subject to the two constraints, fixed B and N (this last condition is not essential), or equations (2) and (1). The theory will also be made with the third constraint, fixed F or equation (3).

We want to account for the detailed interaction between stability at any one trophic level and the stability of the total trophic web. For example, Watt has held that stability at one particular trophic level (*e.g.* competitive stability among predators) may promote overall instability (*e.g.* an herbivore species escapes control). Paine has conjectured that although one particular trophic level may, in isolation, be unstable due to competition, the effects of other levels (*e.g.* predation) can lead to a total ecosystem that is stable.

So, we have chosen an index of diversity-stability which is large at a node of the pathway, when $N_i \sim L_i$ in order to see if the maximum diversity D , *i.e.* a total stable system, can co-exist with unstable subsystems, *e.g.* some parts of the ecosystem with $N_i \gg L_i$ and others with $N_i \ll L_i$.

Equation (9) may be written :

$$(10) \quad D = \sum_i [\log (N_i + L_i)! - \log N_i! - \log L_i!] .$$

In using Stirling's formula to eliminate the factorials :

$$(11) \quad \log N! \approx N \log N - N$$

we obtain :

$$(12) \quad D = \sum_i [(N_i + L_i) \log (N_i + L_i) - N_i \log N_i - L_i \log L_i] .$$

This new index takes into account three terms : (i) the first term is the diversity of the number of individuals of each species and of links reaching each species, (ii) the second and third terms are the diversity of species and the diversity of links respectively.

The second term corresponds to the formulae of Brillouin and Shannon-Weaver [equations (4) and (5)] : diversity increases with the number of species. The third term corresponds to the diversity of the number of links of species and is in agreement with the stability discussed by Mac Arthur and Odum : stability increases with the number of links.

This index of diversity can be interpreted as a measure of the interaction between the number of individuals of each species and the number of links reaching each species.

4.- The most diversified ecosystem

To find the most diversified ecosystem, subject to only two constraints (N and B), the maximum of D , we differentiate equation (11) with respect to N_i and set δD equal to zero. Since all the L_i 's are constant, it follows that :

$$(13) \quad \delta D = \sum_i \left[\log \frac{N_i + L_i}{N_i} \right] \delta N_i = 0 .$$

For a fixed N :

$$(14) \quad \delta N = \sum_i \delta N_i = 0 .$$

Furthermore, for a given biomass in the ecosystem,

$$(15) \quad \delta B = \sum_i B_i \delta N_i = 0 .$$

The most diversified ecosystem is given by equation (13) subject to the two constraints, constant N and B , that is equations (14) and (15). Consequently, the expression giving the maximum of D can be found by the method of Lagrange multipliers. Let us multiply equation (14) by some arbitrary value $\frac{1}{N_0}$ and equation (15) by some other constant $\frac{1}{B_0}$. Subtracting (for convenience) these results from equation (13), we obtain :

$$(16) \quad \delta D - \frac{\delta N}{N_0} - \frac{\delta B}{B_0} = 0$$

or

$$(17) \quad \sum_i \left[\log \frac{N_i + L_i}{N_i} - \frac{1}{N_0} - \frac{B_i}{B_0} \right] \delta N_i = 0 .$$

Two of the N_i 's are not independent, but their coefficients can be made identically zero by the proper choice of the multipliers $\frac{1}{N_0}$ and $\frac{1}{B_0}$. Since all the remaining δN_i 's are arbitrary, the coefficient of each must be zero in order that equation (17) be satisfied. Therefore,

$$(18) \quad \log \frac{N_i + L_i}{N_i} - \frac{1}{N_0} - \frac{B_i}{B_0} = 0$$

for all values of i , or

$$(19) \quad N_i = L_i \exp \left(-\frac{B_i}{B_0} \right) \left[\exp \left(\frac{1}{N_0} \right) - \exp \left(-\frac{B_i}{B_0} \right) \right]^{-1}$$

For the most diversified ecosystem, we find that the diversity is :

$$(20) \quad D_{\max} = - \sum_i L_i \log \left[1 - \exp \left(-\frac{1}{N_0} \right) \exp \left(-\frac{B_i}{B_0} \right) \right] + \frac{N}{N_0} + \frac{B}{B_0}$$

in substituting equation (19) into equation (12).

Moreover, two important relations are the following :

$$(21) \quad \left[\frac{\partial D_{\max}}{\partial N} \right]_B = \frac{1}{N_0}$$

and

$$(22) \quad \left[\frac{\partial D_{\max}}{\partial B} \right]_N = \frac{1}{B_0} .$$

5.- Fluctuations in ecosystems

From the general formula for diversity we deduce :

$$(23) \quad \frac{\partial D}{\partial N_i} = \log (L_i + N_i) - \log N_i$$

$$(24) \quad \frac{\partial^2 D}{\partial N_i^2} = - \frac{L_i}{N_i (L_i + N_i)}$$

The second derivative of D is negative; the distribution D_{\max} obtained by equating to zero the first variation is thus seen to be a maximum of D .

Equations (23) and (24) show that the diversity is independent of the number of individuals in a species if there is no link between this i^{th} species and the others. Then L_i can be interpreted as a parameter representing the degree of interaction of one species with the whole system. The fluctuations of species will depend strongly on the value of L_i .

The mean square of the number of individuals in the i^{th} species is obtained in the following way. In the neighbourhood of the most diversified distribution (*i.e.* D_{max}), the first variation of D_{max} being null for all variations of the numbers N_i which do not modify N and B , we have to second order :

$$(25) \quad D = D_{max} + \frac{1}{2} \sum_i \left(\frac{\partial^2 D}{\partial N_i^2} \right) \delta N_i^2$$

because D is a sum of terms depending each on only one variable N_i . From this expression and from the expression for the second derivative of D_{max} we obtain easily

$$(26) \quad T = T_{max} \exp \left[- \frac{1}{2} \sum_i \frac{\frac{\delta N_i^2}{N_i^2}}{\frac{1}{N_i} + \frac{1}{L_i}} \right]$$

in defining T by $D = \log T$.

Since the variations δN_i are independent, we have, for the mean square of the fluctuation of N_i from its mean value \bar{N}_i ,

$$(27) \quad \frac{\overline{\delta N_i^2}}{\bar{N}_i^2} = \frac{1}{\bar{N}_i} + \frac{1}{L_i} .$$

The mean population of a species depends strongly on the mean quantity of energy flow F_i , variations in which can be very important. The stability of the population of a species depends strongly on the variation in received energy flow $\overline{\delta F_i^2}$. Moreover, this quantity $\overline{\delta F_i^2}$ will be small when the number of links reaching this species is large.

If \bar{N}_i is large, equation (27) becomes

$$(28) \quad \frac{\overline{\delta N_i^2}}{\overline{N_i^2}} \sim \frac{1}{L_i} .$$

If we assume that the relative mean square deviation of the number of individuals in the i^{th} species is proportional to the relative mean square deviation of energy flow, we obtain :

$$(29) \quad \frac{\overline{\delta N_i^2}}{\overline{N_i^2}} \sim \frac{\overline{\delta F_i^2}}{\overline{F_i^2}} .$$

From equations (28) and (29), we obtain a practical measure of the coefficient L_i :

$$(30) \quad L_i \sim \frac{\overline{F_i^2}}{\overline{\delta F_i^2}} .$$

6.- A more complete description of ecosystems

In the precedent section we did not take into account the energy flow F at equilibrium. In differentiating equation (4) with respect to N_i and with δF equal to zero, it follows that

$$(31) \quad \delta F = \sum_i F_i \delta N_i = 0$$

The neglected term, $\sum_i N_i \delta F_i$, is taken into account by the parameter

$$L_i \sim \frac{\overline{F_i^2}}{\overline{\delta F_i^2}} .$$

With this new contribution, equation (16) becomes

$$(32) \quad \delta D - \frac{\delta N}{N_0} - \frac{\delta B}{B_0} + \frac{\delta F}{F_0} = 0$$

where $\frac{1}{F_0}$ is a new Lagrange multiplier.

Equation (18) is written :

$$(33) \quad \log \frac{N_i + L_i}{N_i} - \frac{1}{N_0} - \frac{B_i}{B_0} + \frac{F_i}{F_0} = 0$$

for all values of i , or

$$(34) \quad N_i = L_i \exp\left(\frac{F_i}{F_0}\right) \left[\exp\left(\frac{1}{N_0}\right) \exp\left(\frac{B_i}{B_0}\right) - \exp\left(\frac{F_i}{F_0}\right) \right]^{-1}.$$

One more relation, like equations (21) and (22), is then obtained :

$$(35) \quad \left[\frac{\partial D_{\max}}{\partial F} \right]_{N,B} = \frac{1}{F_0} .$$

When the number of links in an ecosystem is very large ($L_i \gg N_i$) , equation (12) may be written :

$$(36) \quad D \sim \sum_i [N_i \log \frac{L_i}{N_i} + N_i]$$

and we get

$$(37) \quad N_i \sim L_i \exp\left(-\frac{1}{N_0}\right) \exp\left(-\frac{B_i}{B_0}\right) \exp\left(\frac{F_i}{F_0}\right)$$

If we define

$$(38) \quad N_{i1} = L_i \exp\left(-\frac{1}{N_0}\right) ,$$

equation (37) is written :

$$(39) \quad N_i = N_{i1} \exp\left(-\frac{B_i}{B_0}\right) \exp\left(\frac{F_i}{F_0}\right) .$$

So, for an ecosystem composed of species with many links between them, the distribution of individuals is given by equation (39).

Equation (39) shows that the number of individuals increases with the quantity of energy flow received, F_i , and decreases with the biomass, B_i .

Finally, it will be noted that use of the Brillouin formula [equation (4)] leads to the same result [equation (39)] when N_{i1} is a constant (independent of i). Indeed, taking the first variation of H_B :

$$(40) \quad \delta H_B = \sum_i \log N_i \delta N_i = 0$$

and, using the Lagrange multipliers,

$$(41) \quad \delta H_B - \frac{\delta N}{N_2} - \frac{\delta B}{B_0} + \frac{\delta F}{F_0} = 0$$

we obtain

$$(42) \quad N_i = N_1 \exp\left(-\frac{B_i}{B_0}\right) \exp\left(\frac{F_i}{F_0}\right)$$

with

$$(43) \quad N_1 = \exp\left(-\frac{1}{N_2}\right).$$

7.- Conclusion

In our description of ecosystems, from three macroscopic properties defined at equilibrium, *i.e.* the total number of individuals N , the total biomass B and the total energy flow F , we get a statistics of the distribution of individuals in species in maximizing the diversity subject to the three constraints N , B and F .

On one hand, for species which received energy flow with large variations, the following formula would represent a rather good description

$$(42) \quad N_i = N_{i1} \exp\left(-\frac{B_i}{B_0}\right) \exp\left(\frac{F_i}{F_0}\right) [1 - \exp\left(-\frac{1}{N_0}\right) \exp\left(-\frac{B_i}{B_0}\right) \exp\left(\frac{F_i}{F_0}\right)]^{-1}$$

On the other hand, for species which receive stable energy flow, *i.e.* without variations, or for species with great multifunctionality, this description would be more appropriate :

$$(43) \quad N_i \sim N_{i1} \exp\left(-\frac{B_i}{B_0}\right) \exp\left(\frac{F_i}{F_0}\right)$$

References

- BRILLOUIN (L.), (1960). *Science and Information Theory*, 2nd ed., Academic Press, New York.
- Mc ARTHUR (R.), (1955). *Ecology*, 36, 533.
- MARGALEF (R.), (1958). *General Systems*, 3, 36.
- ODUM (E.P.), (1953). *Fundamentals of Ecology*, Saunders, Philadelphia.
- PAINE (R.T.), (1966). *Am. Nat.*, 100, 65.
- PIELOU (E.C.), (1966). *J. Theoret. Biol.*, 13, 131-144.
- SHANNON (C.E.) and WEAVER (W.), (1963). *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana.