

# LES BASES DE DONNEES OCEANOGRAPHIQUES, CONCEPTION ET EXPLOITATION

## **L'EXEMPLE DE LA BASE DE DONNEES COADS DANS LE PROGRAMME CEOS**

*C. Roy et M. -H. Durand*

L'océanographie physique a longtemps été le champs de l'observation et de la mesure, elle s'est ensuite orientée vers la compréhension et la description de la dynamique des océans. Aujourd'hui, l'océanographie prend une part importante dans les recherches menées sur le climat et s'oriente vers la compréhension de la dynamique couplée du système océan-atmosphère. Depuis le début du siècle, des échanges ont été entretenues avec la biologie marine dans le but de mieux comprendre l'impact des fluctuations de l'environnement marin sur la variabilité des populations marines. Plus récemment, la perception d'un changement climatique global et l'urgence de produire quelques résultats fiables en ce domaine a posé de façon nouvelle et cruciale la question de la disponibilité et de la qualité des données historiques. De longues séries de données sont nécessaires pour que l'évaluation scientifique puisse se faire convenablement et soutenir les décisions à prendre pour répondre à ce problème (Houghton *et al.*, 1990). Dans ce contexte, la constitution de bases de données historiques a pris une importance grandissante. Le coût élevé de collecte et de traitement et le caractère universel et fondamental de ces données a donné lieu à la mise en place de projets de grande envergure, rassemblant de nombreux laboratoires et s'appuyant sur des collaborations internationales.

La base de données COADS (Comprehensive Ocean Atmosphere Dataset - Woodruff *et al.*, 1987) est un exemple, parmi d'autres, des produits issus de ces projets. Nous présentons ici un bref historique de la collecte de ces données et les différentes étapes de l'élaboration de cette base de données. Les motivations nous ayant conduit à sélectionner COADS pour le projet CEOS (Climate and Eastern Ocean Systems) sont ensuite brièvement soulignées (Bakun *et al.*, 1993). Le fonctionnement en réseau de CEOS nous a amené à développer, à

partir de COADS, un produit facilement diffusable et accessible à des laboratoires ne disposant pas de moyens informatiques lourds.

Des exemples illustrant les difficultés rencontrées dans l'utilisation de produits standards dérivés de cette base sont donnés. Pour s'affranchir de ces nombreux biais et erreurs, il s'est avéré indispensable de préserver l'accès aux données originales. Du fait du volume très important que représentent les 100 millions d'enregistrements contenus dans COADS, ce fut une contrainte majeure. Les développements technologiques récents ont cependant permis d'arriver à un compromis raisonnable offrant une souplesse d'utilisation inédite.

En parallèle au développement de cette base de données, une réflexion sur la nature et la caractérisation des changements climatiques à partir de séries temporelles a été menée. Des outils spécifiques, issus de l'économétrie, peuvent être appliqués à notre domaine de recherche.

Pour terminer et afin d'illustrer l'utilisation de COADS qui a été faite dans le réseau CEOS, quelques exemples et résultats sont brièvement résumés.

## I. Un bref historique

1626: Le Capitaine J. Smith publie le premier code pour la codification du vent dans les livres de bord.

1775: B. Franklin utilise un thermomètre comme aide à la navigation et pour la cartographie du Gulf Stream.

1805: F. Beaufort définit un code pour quantifier la vitesse du vent. Des versions révisées de ce code sont toujours employées aujourd'hui.

1831: W. Marsden définit un système universel de codification des positions ("Marsden square") afin de faciliter l'archivage et l'analyse des observations réalisées par les bateaux.

1841: M.F. Maury publie une série de cartes de vent et de courant basées sur une première analyse des livres de bord qu'il distribue en échange d'informations météorologiques.

1853: Conférence Maritime de Bruxelles dont l'objectif est de mettre un place un système international de collecte et d'échange de données météorologiques sur les océans. Cette conférence est le premier pas vers l'établissement de l'Organisation Météorologique Mondiale (OMM).

Tableau 1: Chronologie des principales étapes ayant conduit à la mise en place d'un système mondial d'observation des océans.

Au temps de la marine à voile et jusqu'à l'avènement des outils modernes de navigations (DECCA, Radar, LORAN, TORAN et maintenant le GPS), les observations météorologiques réalisées à bord des bateaux étaient d'une importance cruciale pour entretenir l'estime (i.e. connaître sa position) et donc arriver à bon port. Ces observations étaient scrupuleusement consignées sur les livres de bord. La direction et la vitesse du vent, la pression atmosphérique, la température de l'eau et de l'air, l'état de la mer, la nébulosité, la visibilité, etc. étaient les paramètres les plus couramment archivés. D'autres informations, comme la vitesse et de la direction du courant, étaient non pas mesurées mais évaluées à partir de la différence entre les positions estimées et celles obtenues par la navigation astronomique (sextant). Rapidement, il est apparu que la compilation de ces informations était d'un intérêt considérable pour améliorer notre connaissance de ce vaste univers, les océans, qui représente

près de 70% de la surface totale du globe terrestre. Les principales étapes qui ont conduit à la mise en place, au niveau international, d'un système homogène de collecte de ces données sont résumées dans le tableau 1.

La conférence de Bruxelles en 1853 a jeté les bases de l'organisme qui supervise aujourd'hui la collecte, le traitement et la distribution des données météorologiques: l'Organisation Météorologique Mondiale (OMM) dont le siège est à Genève (Suisse). A l'initiative de M.F. Maury du service hydrographique de la Marine américaine, cette conférence avait pour objectif de discuter de la mise en place d'un système uniforme de collecte de données météorologiques et océanographiques sur les océans et de l'utilisation de ces données comme aide à la navigation. La conférence ayant approuvé les idées avancées par M.F. Maury, différents centres nationaux ont pris en charge la collecte et l'archivage des informations contenues dans les livres de bord. A partir de 1895, la mise au point des premières machines à cartes perforées par Hollerith va révolutionner l'archivage et le traitement de ces données. A partir de 1906, une autre étape importante est franchie par la généralisation progressive de la télégraphie sans fil à bord des bateaux. En 1921, un système de transmission des informations en temps réel à 0000, 0060, 1200, 1800 TU est mis en place. C'est à cette époque que l'archivage et le codage sous forme numérique (carte perforée puis par la suite supports magnétiques) se sont généralisés.

A partir de 1940, afin de fournir des informations à la flottille US opérant dans le Pacifique Ouest et pour préparer le débarquement des Alliés en Normandie, des collections de données ont été rassemblées. A la fin de la Seconde Guerre Mondiale, de gros efforts ont alors été consacrés pour combiner ces jeux de données sous un format unique. Ce travail était justifié par la nécessité de mieux connaître le climat à l'échelle mondiale, par les besoins de plus en plus importants exprimés par les services météorologiques afin d'affiner les prévisions météorologiques et le développement des premiers modèles. Dans les années 1960 aux Etats-Unis, 15 jeux différents de données ont été rassemblés par le National Climate Data Centre (NCDC) sous un format unique: Tape Data Family 11, plus connu sous l'abréviation TDF-11 (NCDC, 1968). La mise au point et la diffusion de ce premier jeu de données uniformisées, associée à la croissance exponentielle des moyens de traitement informatique, a montré l'intérêt que pouvait représenter la mise en valeur de ces bases de données historiques pour la recherche océanographique et pour la prévision météorologique. Cela a donné une impulsion considérable aux tâches de sauvetage et d'archivage des données historiques. Au cours des années 1970, de nombreux projets ont alors vu le jour afin de

compléter ces jeux de données avec des archives encore inexploitées et d'évaluer la validité des informations ainsi obtenues. De nombreuses biais et inconsistances ont ainsi été mis en évidence. Parmi ceux-ci, on peut citer le biais introduit dans les séries temporelles de Température de Surface de la Mer (TSM) suite aux modifications intervenues dans le mode de collecte des données au cours des années 1940. Jusqu'à la seconde Guerre Mondiale, la mesure de la TSM était réalisé à partir d'un échantillon d'eau de mer prélevé à l'aide d'un seau en toile ou en bois. Une fois sur le pont du navire, l'évaporation peut entraîner un refroidissement de l'échantillon et ainsi fausser la mesure. Dans l'Atlantique Nord, des études ont montré qu'en hiver, la correction à apporter aux séries de TSM peut atteindre 1°C (Folland and Hsiung, 1987; Folland and Parker, 1990, Jones and Wigley, 1992). A la fin des années 1940, l'apparition de seaux isolés et la mise en place progressive de thermomètres sur la prise d'eau des machines a supprimé ces problèmes ... mais d'autres sont apparus !

PRODUITS	TYPE	NOMBRE DE BANDES MAGNETIQUES
Long Marine Reports (LMR 5)	R	48
Decadal Summaries, Trimmed or Untrimmed	D	2
Compressed Marine Reports (CMR 5)	R	18
Monthly summaries Untrimmed	M	9
Monthly summaries Trimmed	M	18
NCDC TD-1129	R	115

Tableau 1 : Liste des différents produits de la base de données COADS ("release 1"). Le type de produit est identifié par : R = enregistrements individuels, M = moyennes mensuelles par carré 2°x2° de 1854 à 1979, D = climatologie décennale par mois et par carré 2°x2°. Le nombre de bandes magnétiques est basée sur une densité de 6250 bpi. LMR correspond à l'ensemble des observations individuelles enregistrées en ASCII de 1800 à 1979 (volume = 39,5 Gb). CMR est une sélection de 28 éléments les plus couramment utilisés, de 1854 à 1979, enregistrements individuels compressés en binaire comportants chacun des indicateurs de qualité (volume = 13,7 Gb). TD 1129 est le jeu de données initial à partir duquel COADS a été élaboré (volume 84.6 Gb).

Le projet qui a conduit à la création de la base de données COADS (Comprehensive Ocean-Atmosphere Data-Set) a vu le jour en 1981 à la suite d'une initiative conjointe de plusieurs laboratoires américains : le CIRES (Cooperative Institute for Research in Environmental Sciences), le ERL (Environmental Research Laboratories), le NCAR (National Centre for Atmospheric Research) et le NCDC (National Climatic Data Centre). L'objectif du projet était de mettre en place une base de données, facilement utilisable, relatant l'histoire de la variabilité climatique des océans de 1854 à 1979. Plus de 100 millions d'observations météorologiques couvrant l'ensemble des océans de 1854 à 1979, provenant de sources diverses, ont été rassemblées. Diverses procédures visant à trier les enregistrements, éliminer les doubles, corriger et valider ce jeu de données ont été appliquées. Ce travail a été réalisé sur les super-ordinateurs du NCAR : CDC 7600 et CRAY 1, les seules machines pouvant traiter à cette époque une telle masse de données. L'ensemble des procédures de validation qui ont été mises en oeuvre constitue indéniablement le point fort et l'originalité de COADS. Ces procédures ont conduit à associer, à chaque enregistrement, différents indicateurs de la "qualité" de la mesure (mode de prélèvement, position par rapport à la moyenne climatique, etc.). Ainsi, avec une telle structure, chaque utilisateur peut lui-même définir ses propres critères de qualité pour sélectionner ses données. C'est un atout majeur par rapport à d'autres bases de données. La première version de COADS dénommée "COADS release 1" a été diffusée en 1985, après plus de 4 années de travail intensif. Les principaux paramètres accessibles dans COADS sont : température de l'air, de la mer, vitesse des composantes Nord-Sud et Est-Ouest du vent, pression atmosphérique, humidité, couverture nuageuse (Woodruff *et al.*, 1987).

"COADS release 1" est constitué de 14 produits différents couvrant la période 1854-1979: enregistrements originaux qui ont permis de constituer la base de données, enregistrements traités selon la procédure COADS sous différents formats (ASCII : LMR5, Binaire compressé: CMR 5), moyennes des différents paramètres calculées avec différents pas de temps incluant l'ensemble des données ("untrimmed records") ou une sélection basée sur des tests statistiques ("trimmed records") (Slutz *et al.*, 1985). Les principaux produits sont présentés dans le tableau 1. Pour la période 1981-1990, un jeu de données temporaire sur lequel une validation partielle a été réalisée fut diffusé en 1991. Seuls les formats LMR et CMR permettent l'accès à la donnée individuelle, les autres produits sont constitués de moyennes mensuelles ou de climatologies par carré  $2^{\circ} \times 2^{\circ}$ . Ces derniers produits sont ceux les plus couramment utilisés du fait de leur maniement aisé et de leur volume relativement compact. Depuis la diffusion de COADS, plusieurs groupes de travail ont été organisés afin de discu-

ter de l'utilisation de COADS, des problèmes rencontrés et des améliorations à apporter (Woodruff, 1986; Diaz *et al.*, 1992, Diaz et Isemer, 1994). En 1993, une mise à jour de la période 1980-1992 (release 1a) a été diffusée avec des améliorations notables mais qui ont cependant entraîné une modification du format CMR. Une mise à jour partielle de la période 1947-1992 a été diffusée en 1995 (Release 1b). Le traitement de l'ensemble des données sous ce nouveau format, avec l'ajout de nombreuses données et de nouvelles procédures de contrôle, est prévue en 1996 (Release 2) (Woodruff *et al.*, 1993).

## **II. Motivations pour l'utilisation de COADS dans le programme CEOS.**

Un des objectifs premiers du programme CEOS (Climate and Eastern Ocean Systems) était de rassembler et d'analyser en suivant une approche comparative, les données concernant les quatre grands écosystèmes d'upwelling au cours des quarante dernières années (Bakun *et al.*, 1993). Pour les aspects climatiques, il existe peu de bases de données avec une telle couverture spatiale et temporelle. Une des contraintes majeures était aussi de disposer d'un jeu de données homogènes spatialement et temporellement, afin de réaliser les analyses rétrospectives et comparatives envisagées par CEOS. Rapidement, il est apparu que seule COADS permettait de satisfaire ces besoins.

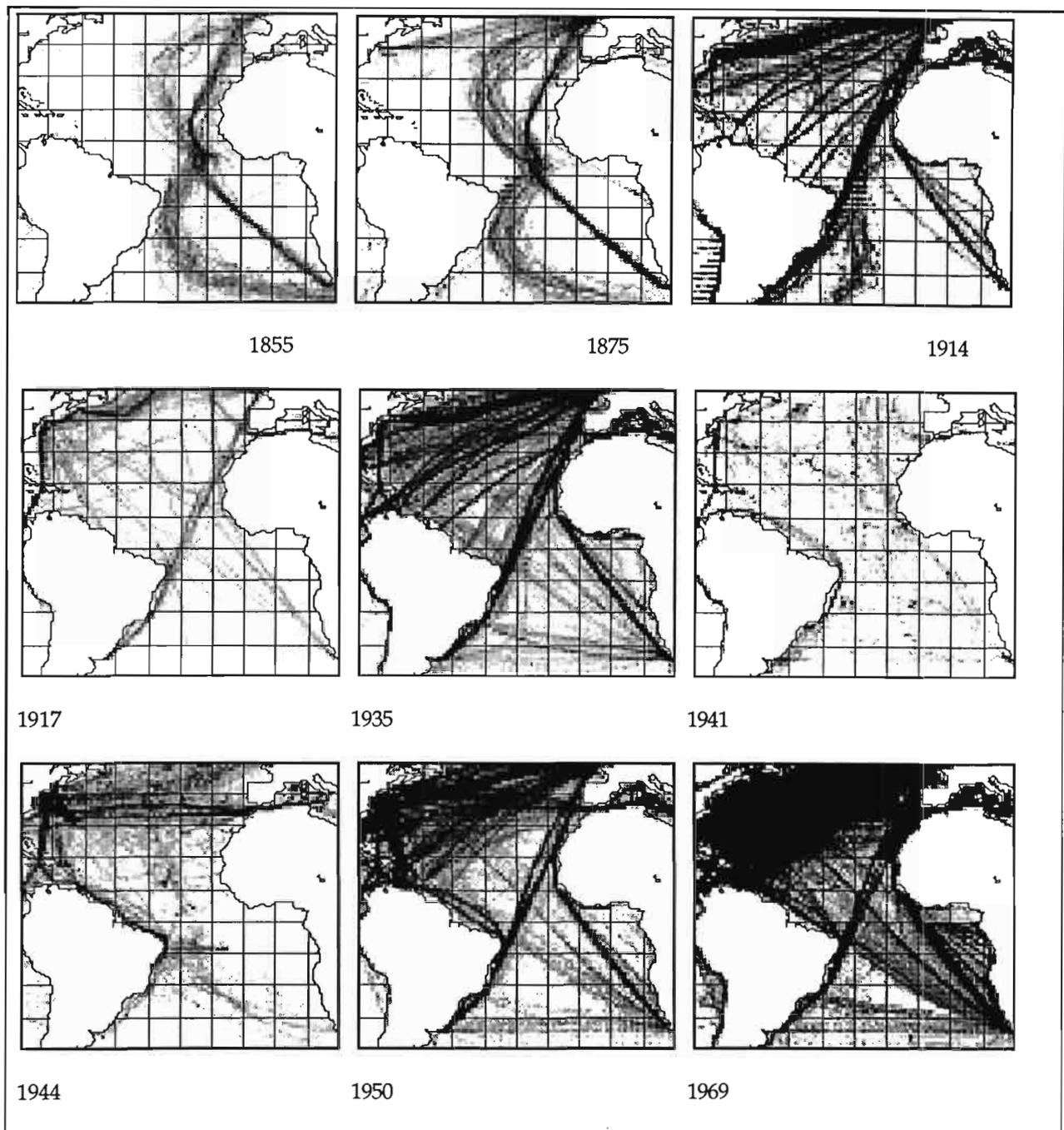


FIGURE 1 : Quelques exemples de cartes illustrant la densité spatiale des données COADS (TSM) dans l'Atlantique en 1855, 1875, 1914, 1917, 1935, 1941, 1944, 1950 et 1969. La densité des données est maximale le long des principaux rails de navigation. Au début du siècle, avec l'avènement de la marine à vapeur et la disparition des grands voiliers, les routes suivies par les navires deviennent rectilignes. À partir de la fin de la Seconde Guerre Mondiale, la densité des données est élevée sur l'ensemble de l'Atlantique Nord et dans l'Atlantique Est au nord de 10°N. Les façades maritimes de l'Afrique de l'Ouest, de l'Afrique du Sud, de la Namibie, du Brésil et de l'Argentine sont relativement bien échantillonnées par la présence de rails de navigation longeant ces côtes. La densité des données est faible dans le Golfe de Guinée et l'Atlantique Sud. Les conséquences, en terme de trafic maritime, de nombreux faits historiques sont repérables sur ces cartes. L'ouverture du Canal de Suez à la fin du siècle dernier, a entraîné une augmentation importante du trafic en Méditerranée au détriment du rail Dakar-Le Cap (Cap de Bonne-Espérance). L'activité des sous-marins allemands au cours des deux guerres mondiales a eu pour conséquence une diminution très forte du trafic maritime sur l'Atlantique. À partir de 1942, on observe cependant une augmentation significative du trafic entre les USA et la Méditerranée : cela marque l'implication des USA en Afrique du Nord et l'ouverture d'un deuxième front par le Général Patton (opération Torch). La fermeture du

Canal de Suez entre 1967 et 1975 a conduit à une forte augmentation du trafic maritime autour du Cap de Bonne Espérance.

COADS ne rassemble que des données de surface, c'est une limitation importante mais la densité des données océanographiques de subsurface disponible est encore faible et il est illusoire de vouloir baser des analyses rétrospectives sur de telles données, notamment dans les zones tropicales.

Pour CEOS, la base de données COADS présente donc de nombreux atouts. L'étendue spatiale échantillonnée par les bateaux marchands couvre la majeure partie des océans : dès la fin du 19ème siècle, il est possible de construire des chroniques de la variabilité climatique le long des rails de navigations qui s'étendent dans l'Atlantique Sud en direction du Cap Horn et du Cap de Bonne Espérance (fig. 1 et 2).

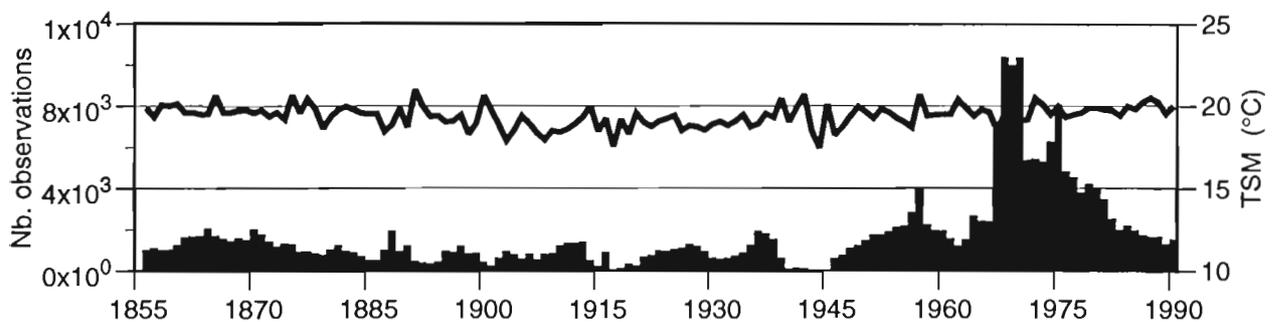


Figure 2: Evolution annuelle du nombre d'observations et de la Température de Surface de la Mer (TSM) dans l'Atlantique Sud, de 1854 à 1990 (zone : 20°S-30°S, 0-10°E).

Du fait de la relative uniformité des plates-formes d'observations et des méthodes de mesures (observations réalisées à partir de la passerelle des bateaux), il est possible de comparer d'une région ou d'un océan à l'autre, des séries temporelles obtenues à partir des observations rassemblées dans COADS. De telles comparaisons sont difficiles à réaliser avec d'autres types de données. Avec des stations côtières par exemple, les principales difficultés résident dans les modes de prélèvement et le pas d'échantillonnage qui diffèrent d'une station à l'autre; souvent, ces données ne reflètent que des conditions très locales qu'il est difficile d'étendre à l'ensemble d'une zone (influence du relief sur le vent, prélèvements réalisés dans les ports situés en général dans des zones abritées, ...). Les données satellites offrent une couverture mondiale mais qui ne remonte guère au delà des années 1980; pour ces données les problèmes liés à la calibration et la dérive des capteurs sont encore loin d'être entièrement résolus; dans de

nombreuses régions (zone équatoriale par exemple) la trop forte couverture nuageuse ne permet pas de disposer d'observations de manière continue (fig. 3). Les données COADS ne sont pas absentes de biais et d'erreurs, cependant les principales limitations sont connues et documentées du fait du travail de recherche réalisé par un grand nombre d'organismes et des échanges constamment entretenus entre les concepteurs et les utilisateurs de COADS. Pour les zones tropicales, COADS possède un atout majeur : ce sont très souvent les seules et uniques données disponibles ! Un autre élément essentiel est la présence de rails de navigation situés près des côtes avec une forte densité de données permettant ainsi un échantillonnage fin et régulier (fig. 1). En Afrique de l'Ouest, du Maroc au Sénégal, la densité des données avec un maillage de  $2^{\circ} \times 2^{\circ}$  est en général supérieure à une centaine d'observations par mois, permettant ainsi de bâtir des séries temporelles continues et relativement fiables depuis la fin de la Seconde Guerre Mondiale.

L'intérêt de ces données pour les recherches halieutiques a été démontré lors d'un programme sur la variabilité et l'instabilité des ressources pélagiques côtières d'Afrique de l'Ouest, développé par l'ORSTOM et les partenaires régionaux (Maroc, Mauritanie, Sénégal, Côte d'Ivoire et Ghana) à partir du CRODT de Dakar (Cury et Roy, 1991) à la fin des années 1980. A cette occasion, nous avons utilisé un sous-ensemble du Fichier TDF-11 (Atlantique 30N-20S, de 1964-1984) pour quantifier la variabilité de l'environnement en Afrique de l'Ouest. Profitant des moyens informatiques disponibles au CRODT, nous avons pu développer un certain nombre de programmes d'interrogations et de traitement, tout d'abord sur le calculateur central du centre et ensuite nous avons assuré le transfert des données et des programmes sur des micro-ordinateurs de type PC. Ce produit a ensuite été diffusé aux partenaires régionaux, leur permettant ainsi d'avoir pour la première fois accès aux données récoltées devant leurs côtes.

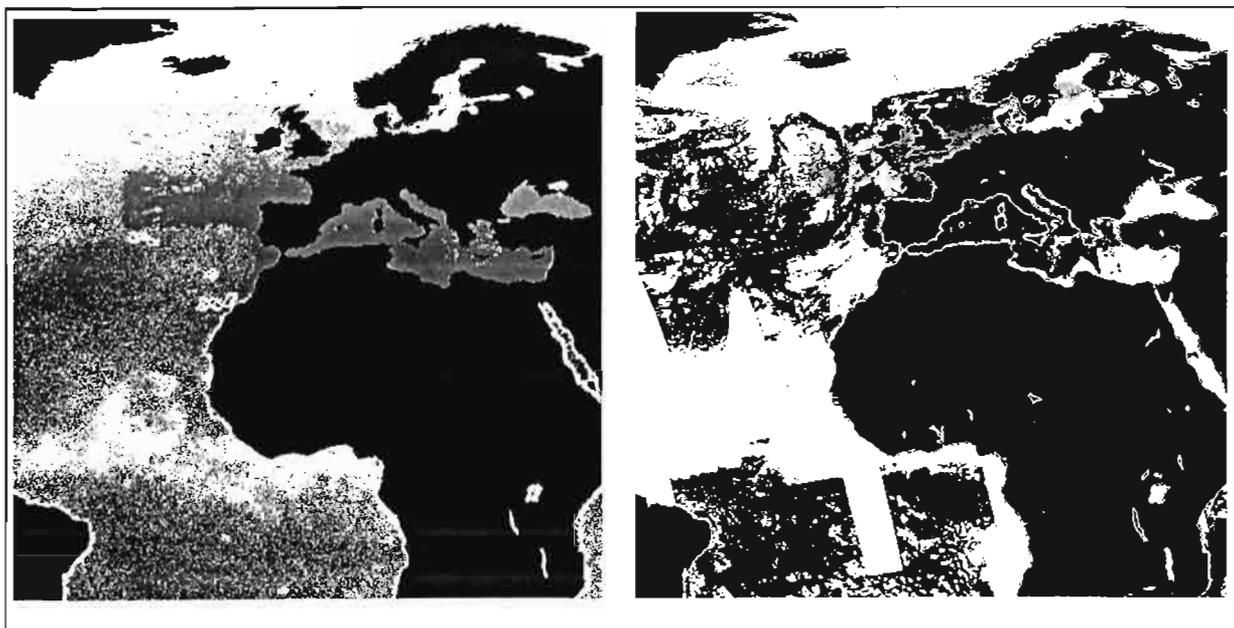
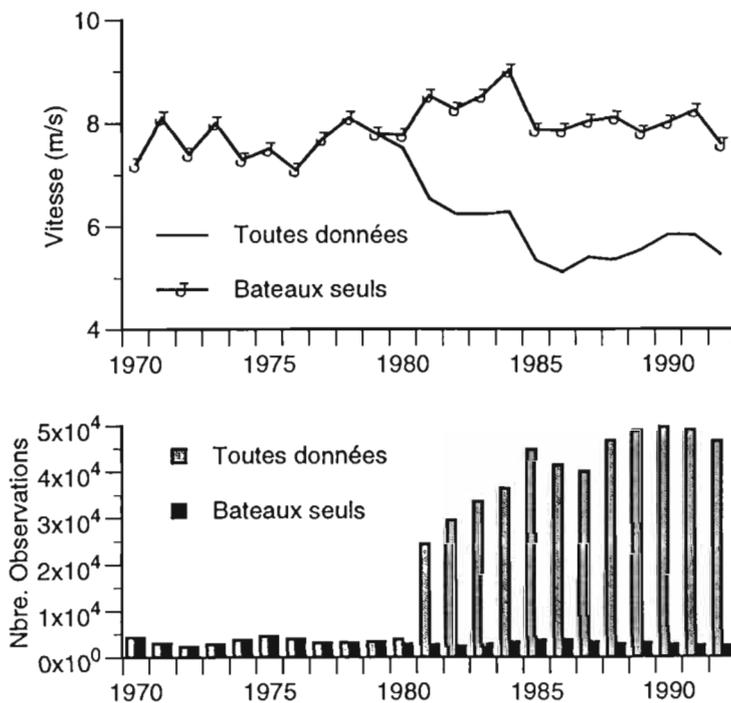


Figure 3: Synthèses mensuelles de la distribution spatiale de la TSM (janvier 1986) et de la chlorophylle (juin 1986) obtenue à partir des satellites NOAA dans l'Océan Atlantique. Malgré une répétitivité élevée, de larges régions ne sont pas échantillonnées, notamment en raison des difficultés rencontrées dans les régions à forte nébulosité. (Source : CD-ROM-NASA-JPL)

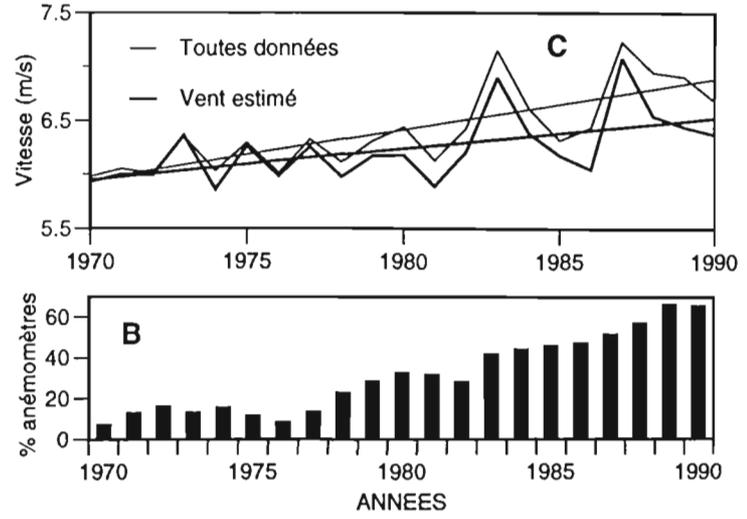
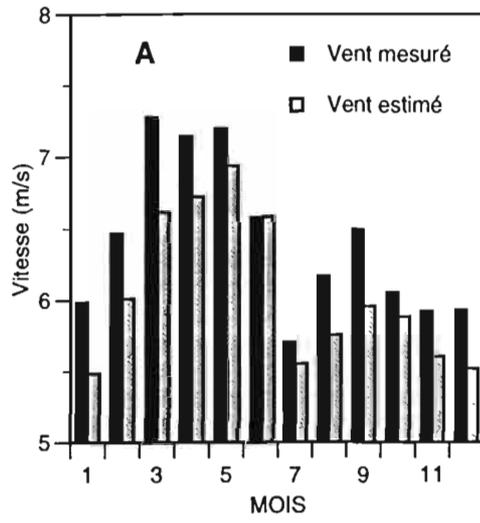
### III. Conception et développement de la base de données

Lors de la mise en place de CEOS, un accord est rapidement intervenu entre les différents partenaires du projet (NOAA, ICLARM, ORSTOM) concernant la priorité à donner au développement d'une base de données climatiques basée sur COADS. Le réseau scientifique CEOS qui associe des pays en développement aux trois instituts leaders de CEOS était une composante majeure du projet. Il s'avérait donc indispensable d'assurer la diffusion de cette base de données à une communauté scientifique ne disposant pas de moyens informatiques lourds. La décision fut prise de développer l'ensemble de la base et les outils d'interrogation sur micro-ordinateur. Parmi les nombreux produits disponibles au NCAR, notre choix s'est évidemment porté sur le jeu de données pré-traités (moyennes mensuelles des paramètres les plus usuels par carré  $2^{\circ} \times 2^{\circ}$ ). Le volume relativement faible de ces fichiers (environ 80 Mb) comparé à celui du jeu de données originales (environ 3,2 Gb sous le format CMR 5) et la facilité de mise en oeuvre ont constitué les éléments déterminants de ce choix.

Un premier prototype a été développé sur PC et Macintosh et les premières analyses ont démarré. Rapidement, nous nous sommes heurtés à de nombreuses difficultés. Le maillage spatial rigide (grille  $2^\circ \times 2^\circ$ ) ne permettait pas de prendre en compte la topographie de la côte. Les rails de navigations sur lesquels la densité des observations est maximale se trouvaient parfois distribués sur deux éléments de la grille. Le problème majeur résulte cependant du calcul des moyennes mensuelles réalisées par le NCAR et qui intègre sans discernement l'ensemble des données disponibles.



COURANT DE CALIFORNIE



une table de conversion échelle beaufort-noeuds (ou m/s) est utilisée. Malheureusement, cette table de conversion élaborée au début du siècle, sous-estime les vents faibles et modérés. Comme le nombre de données de vent mesurées à l'aide d'un anémomètre s'accroît régulièrement depuis les années 1950 (fig. 5), le vent moyen calculé à partir de l'ensemble des données disponibles (estimées à partir de l'échelle Beaufort et mesurées à partir d'un anémomètre), présente un biais systématique se traduisant par une tendance positive, totalement fictive (fig. 5).

Utiliser les données pré-traitées par le NCAR sous forme de moyennes mensuelles présentait de nombreux attraits, cependant l'utilisation de ce type de produit s'est avérée dangereuse pour des études climatiques du fait des nombreux biais existants. Seule l'accès aux données individuelles permet de contourner certaines de ces difficultés. Après cette première tentative à partir des fichiers mensuels, la décision fut donc prise de développer notre base de données à partir des fichiers au format CMR5 qui permettent de conserver l'accès aux enregistrements individuels. Le volume important des fichiers au format CMR5 (3,2 Gb) constituait une contrainte majeure qui a pu être surmontée par l'émergence de nouveaux périphériques de stockage de grande capacité (disque optique). Pour la diffusion à travers le réseau, la solution adoptée fut le transfert des fichiers sur cinq CD-ROM dont une centaine d'exemplaires fut produite. Le transfert sur support optique s'est effectué après avoir restructuré et homogénéisé l'ensemble des fichiers CMR5. Un tri par secteur géographique fut réalisé. La structure choisie est la suivante : l'ensemble du globe terrestre est découpé en 648 carrés de  $10^\circ$  de latitude par  $10^\circ$  de longitude. Les carrés  $10^\circ \times 10^\circ$  sont ensuite subdivisés en 25 carrés de  $2^\circ \times 2^\circ$ . A chaque carré  $10^\circ \times 10^\circ$  est associé un répertoire contenant 25 fichiers, chaque fichier rassemble les données dans un carré de  $2^\circ \times 2^\circ$ . Cette structure permet de limiter la taille des fichiers à prendre en compte lors d'une interrogation de la base de données dans une région donnée : les seules données à traiter seront celles correspondant à l'ensemble des fichiers  $2^\circ \times 2^\circ$  qui délimitent la région sélectionnée. Avec une telle structure, un utilisateur travaillant régulièrement dans une zone donnée peut envisager de ne prendre en compte qu'une portion relativement réduite de la base de données (par exemple, un utilisateur travaillant dans le Golfe de Guinée pourra s'affranchir des fichiers du Pacifique et de l'Atlantique Nord).

En parallèle avec le développement de la base de données, un module d'interrogation a été développé. Les tâches allouées à ce module ont volontairement été limitées à la sélection, l'extraction et le calcul de séries temporelles suivant différents pas d'espace et de temps. Cela privilégie bien sur la rapidité d'accès aux données mais surtout

cela laisse à l'utilisateur le libre choix des outils pour la visualisation et le traitement des séries. D'autre part, il aurait été inutile de se lancer dans une telle direction alors qu'il existe sur le marché de nombreux logiciels scientifiques extrêmement performants pour réaliser ces tâches. La sélection des données suivant différents critères : 1-critères géographiques (latitude-longitude) et temporels (année-mois), 2-critères relatifs à l'origine (bateaux, bouées, ...) et le type de données (données de vent estimées ou mesurées), 3-critères statistiques. Le dialogue entre l'utilisateur et la machine se fait à travers une interface dont les principales fenêtres de dialogue sont présentées sur les figures 6-A à 6-C.

Dans un premier temps, l'interface a été développée sur Macintosh, une version DOS a ensuite été mise au point mais, du fait de la complexité du code informatique, un système de développement spécifique a du être acquis auprès d'une compagnie privée, le coût de la licence multi-utilisateurs étant prohibitif, nous n'avons pu distribuer ce produit. Une nouvelle version destinée aux compatibles PC, utilisant l'environnement Windows 95, est en cours de développement au Centre ORSTOM de Brest et sera disponible début 1997.

The screenshot shows a 'Settings' dialog box with the following sections and controls:

- Space Settings:**
  - North Latitude: 34
  - East Longitude: -136
  - South Latitude: 30
  - West Longitude: -140
  - Space Interval: 2 Degrees (dropdown)
- Time Settings:**
  - Start Year: 1980
  - End Year: 1982
  - Time Interval: Year (dropdown)
- Selected Parameters:**
  - List: SST, Air Temperature, Dew Point, Wind Components
  - Parameter Options button
- File Locations:**
  - Data Directory: MAC HD:CODE:COADS Data:
  - Output Directory: MAC HD:CODE:Results:
  - Output File Prefix: FILE1
- Output Format:**
  - ASCII
  - HDF
  - Map
  - NetCDF
- Buttons:** Quit, Extract

Figure 6.A : Fenêtre principale de l'interface.

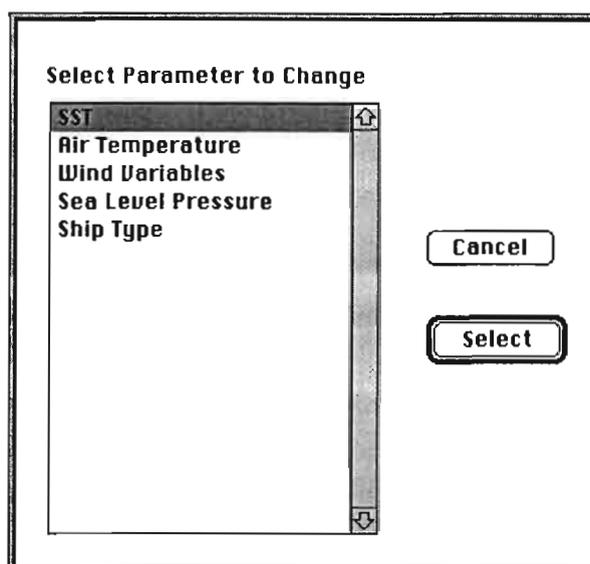


Figure 6 B : Fenêtre correspondant à la sélection des paramètres et le type de plate-forme : température de surface de la mer, de l'air, vent, pression atmosphérique et plate-forme.

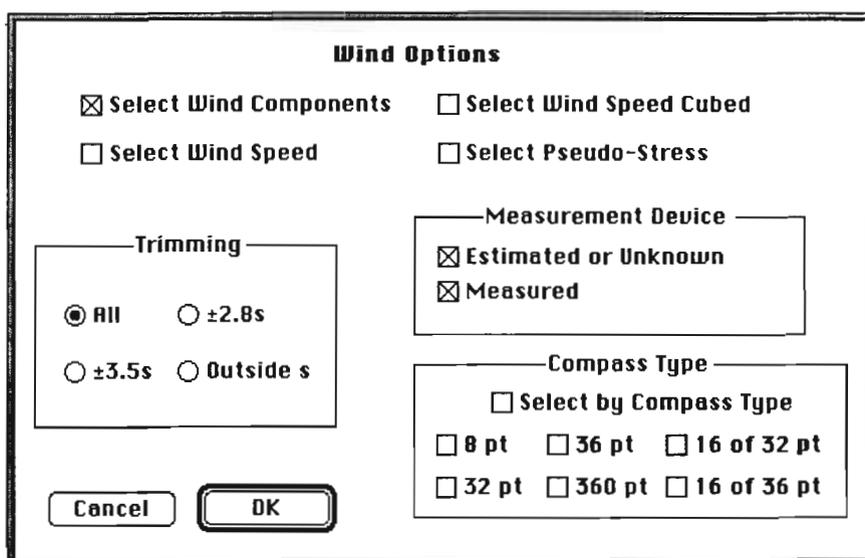


Figure 6-C : Fenêtre correspondant à la sélection des options pour le vent : composantes Nord-Sud et Est-Ouest, vitesse scalaire, variables dérivées (pseudo tension du vent, vent au cube), type de mesure (estimé, mesuré et type de compas) et indicateurs statistiques.

#### IV. Questions et méthodes d'analyse

L'étude de la variabilité climatique de l'environnement marin dans les systèmes d'upwelling est importante du fait de l'influence déterminante des fluctuations climatiques sur l'instabilité des espèces

pélagiques. Caractériser les changements qui surviennent réellement dans l'environnement océanographique est devenu un problème crucial. C'est une question difficile qui nécessite de considérer avec attention ce que l'on entend exactement par le mot "changement" et la manière dont les changements apparaissent dans les données océanographiques.

La détection des changements dans les données océanographiques a pris une nouvelle tournure depuis l'hypothèse émise par Bakun (1990) stipulant que le réchauffement global pourrait avoir un effet positif sur la tension des vents parallèles à la côte et en conséquence produire une intensification des upwellings côtiers. Les séries temporelles de tension de vent le long du bord Est des océans semblent montrer l'existence d'une tendance positive depuis le début des années 1950. Cette tendance semble être un trait commun partagé par ces différentes séries couvrant l'Atlantique et le Pacifique. Toutes autres manifestations de "changement" paraissent plus locales ou transitoires. Si ce phénomène s'avérait être effectivement généralisé, on devra lui imputer une cause générale avec des conséquences peut-être problématiques pour l'avenir à long terme de certaines pêcheries.

L'analyse des changements climatiques des upwellings côtiers s'articule autour d'un certain nombre de questions:

- le cycle saisonnier est-il constant ou changeant dans le temps ?
- observe-t-on réellement des tendances de long-terme ?
- les changements sont-ils permanents ou transitoires ?
- existe-t-il des ruptures de tendances, à partir de quel seuil ?

Pour expliquer ces changements, les océanographes cherchent également à savoir si des chocs autonomes, tels ceux produits par les phénomènes de type ENSO, ont des effets permanents ou transitoires sur la dynamique des océans, si les changements saisonniers ou les tendances convergent d'une quelconque façon pour différentes variables océanographiques, ou encore, analysant les évolutions communes d'une variable sur un large ensemble géographique, à partir de quel niveau de glissement est-il possible de parler de zone de rupture ou de transition ceci pouvant correspondre, par exemple, à un seuil écologique et établir un barrage écologique aux migrations de poissons.

Pour comprendre le comportement d'une variable et formuler quelques hypothèses, il est habituel de distinguer dans l'évolution générale d'une série différents mouvements caractéristiques: tendanciel, cyclique, saisonnier et résiduel. Cela suppose donc que l'on soit capable d'identifier et d'extraire ces diverses composantes d'une donnée observée et synthétique par nature. Détecter des "tendances" et statuer sur l'existence d'un "changement saisonnier" est loin d'être facile, les concepts de "tendance" ou de "saisonalité" n'ayant pas même de définition statistique précise et formelle. Il est maintenant usuel, lorsqu'on analyse une série temporelle tendancielle, de s'interroger au préalable sur le caractère stochastique ou déterministe de la tendance. Dans les échantillons finis, ces deux types de tendance ont la même apparence. Une batterie de tests, dits "tests de racine unité", ont été mis au point pour distinguer ces deux cas préalablement à toute représentation. En effet, les propriétés statistiques de ces séries et les modèles à utiliser sont radicalement différents. Sur les séries réelles, l'existence conjointe de ces deux types de tendance ou encore, plus fréquemment, l'existence de retournement ou de rupture de tendance, compliquent sérieusement la mise en oeuvre de ces tests et la fiabilité de leurs résultats.

En dépit d'une longue histoire de l'étude des saisonnalités, la littérature sur ce sujet offre une grande variété de définitions. La définition la plus simple est celle d'une saisonnalité déterministe autorisant la moyenne d'une série à varier à chaque saison. Ce modèle produit une saisonnalité fixe et régulière. Les autres définitions de la saisonnalité tels que la saisonnalité stochastique stationnaire, les processus saisonniers "intégrés", ou encore les processus périodiques sont plus intéressants à détecter dès lors que l'on cherche à mettre en évidence l'existence de "changements saisonniers" et à les caractériser. Les processus déterministes permettent quelques variations mais pas de véritables changements du cycle saisonnier, tandis que les processus périodiques et stationnaires stochastiques autorisent des changements transitoires et les processus intégrés sont caractéristiques de changements permanents du cycle saisonnier.

Comme pour la composante tendancielle, les propriétés statistiques de ces grands types de saisonnalité sont fondamentalement différentes, mais, dans les échantillons finis et en absence de chocs aléatoires, ils peuvent apparaître très similaires. Identifier les propriétés statistiques d'une série observée permet donc de caractériser les changements qui surviennent dans l'environnement océanographique. La méthode consiste à utiliser des tests de racines unités saisonnières. Les tests actuellement disponibles ne sont pas toujours satisfaisants et sont encore l'objet d'une recherche active en économétrie.

Si l'existence de racines unités dans une série chronologique est important pour comprendre quel type de tendance affecte le comportement de cette série et son mouvement saisonnier, cela a aussi d'intéressantes conséquences lorsqu'on examine conjointement un ensemble de séries. Des séries non-stationnaires dans leur tendance de long-terme ou dans leur cycle saisonnier ont parfois la propriété de devenir stationnaires lorsque on les modélise conjointement. On dit alors qu'elles sont cointégrées. Cette notion de cointégration est intéressante, car bien que statistique, elle peut fournir des interprétations utiles dans un cadre plus théorique. Les modèles de cointégration permettent de séparer, dans un ensemble de séries, les composantes tendancielle ou saisonnières qui leur sont communes de celles qui restent particulières à chaque série.

Une étape-clé dans la compréhension de la dynamique des upwellings est de pouvoir comprendre et distinguer les liens qui existent entre des évolutions ou des tendances communes qui affectent de la même manière une large zone géographique (changement global) et les mouvements cycliques, stationnaires ou non, qui restent propre à une zone particulière (dynamique et changement local). Considérés dans un contexte spatial, les modèles de cointégration peuvent permettre d'analyser la variabilité à différentes échelles spatiales et temporelles.

Ces notions et modèles d'analyses des séries chronologiques, dont on n'a présenté ici qu'une petite sélection, ont été développés ces dix à quinze dernières années par des économistes travaillant essentiellement dans le domaine de la finance. Ces méthodes sont encore très peu utilisés par d'autres disciplines, elles peuvent pourtant, à l'évidence, être d'un grand intérêt dans de multiples domaines. Les rares transferts disciplinaires qui se sont faits très récemment concernent les modèles Espace-Etat et l'utilisation du filtre de Kalman à des études climatologiques car seule l'existence de larges bases de données avec des séries de taille suffisamment longue permet de les mettre en oeuvre.

## **V. Quelques résultats**

L'utilisation de la base de données COADS au sein de CEOS a été intensive. COADS a été sollicitée dans de nombreux domaines : études sur la climatologie des upwellings, sur la variabilité du recrutement en Californie et au Chili, sur les stratégies de reproduction dans les zones d'upwelling, sur les bouleversements écologiques dans le Golfe de Guinée, ...

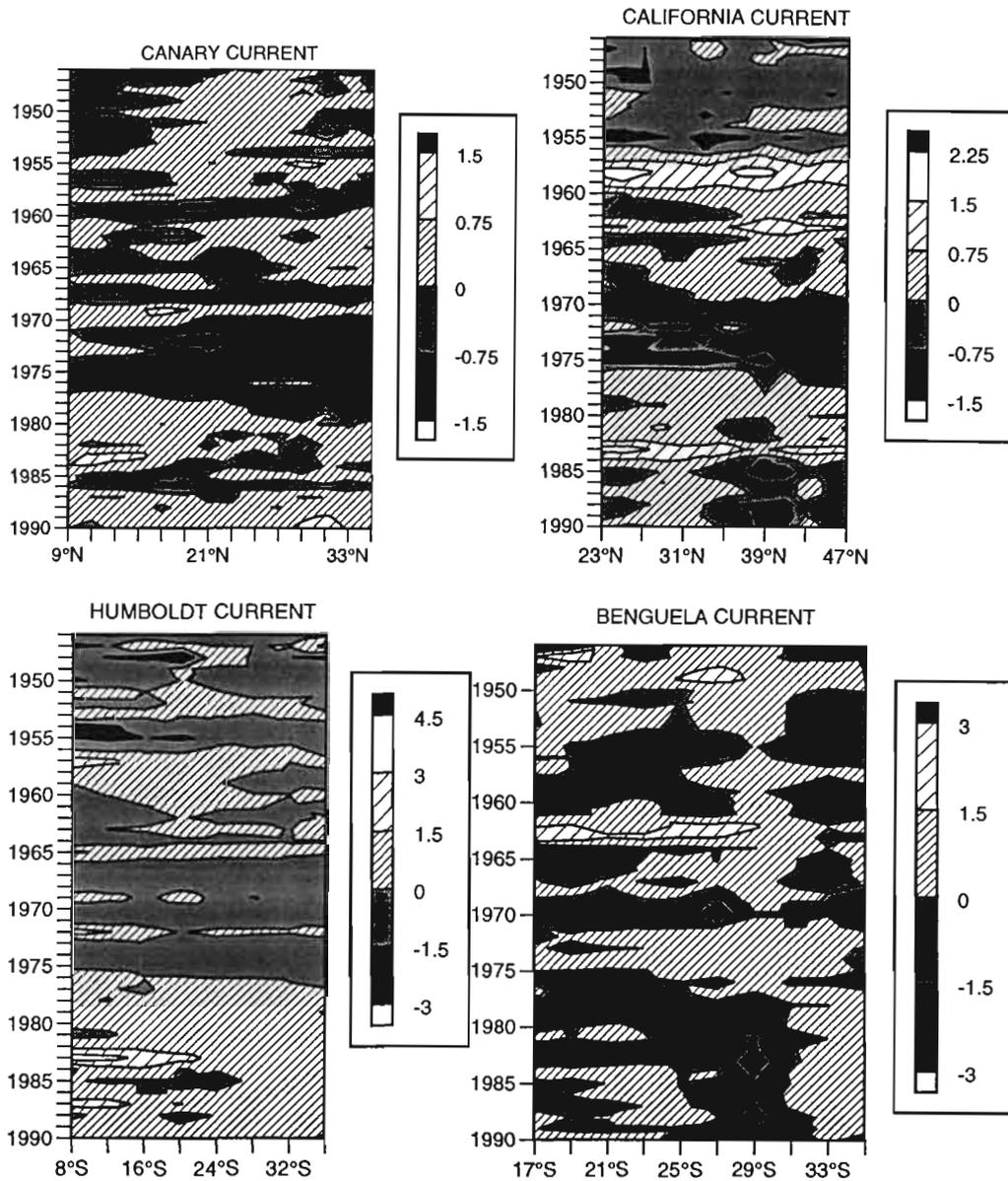


FIGURE 7 : Evolution spatio-temporelle des anomalies de température de surface dans les quatre grandes zones d'upwelling déterminées d'après les données COADS. Un élément remarquable est la présence d'une anomalie froide au cours des années 1970-1976 qui s'est développée simultanément dans trois de ces écosystèmes (Canaries, Californie et Humboldt). On peut aussi noter que l'amplitude de la variabilité interannuelle est faible dans le courant des Canaries par rapport à celle observée dans les autres régions qui sont dominées par des événements chauds de type ENSO (adapté de Laurent, 1996).

La première étape fut d'évaluer la qualité des informations contenues dans COADS et de cerner le degré de confiance que l'on pouvait accorder aux séries temporelles obtenues à partir des extractions

réalisées. La mise en évidence des biais introduits par la prise en compte de données issues des bouées devant les côtes californiennes (fig. 4) et les tendances artificielles dues à l'augmentation des données de vent collectées à l'aide d'un anémomètre (fig. 5) sont quelques exemples, parmi d'autres, du travail réalisé dans ce domaine.

Des études comparatives de la variabilité climatique des upwellings ont été réalisées. Pour ce type d'étude, l'atout majeur de COADS est de disposer d'une base de données homogènes sur l'ensemble des océans, permettant ainsi des études comparatives inter-écosystèmes. La figure 7 illustre le type d'analyses et de comparaisons qu'il a été possible de réaliser. Dans le même ordre d'idée, une synthèse des relations entre des paramètres de l'environnement et les stratégies de reproduction des poissons pélagiques, a pu être menée sur l'ensemble des écosystèmes d'upwelling (fig. 8).

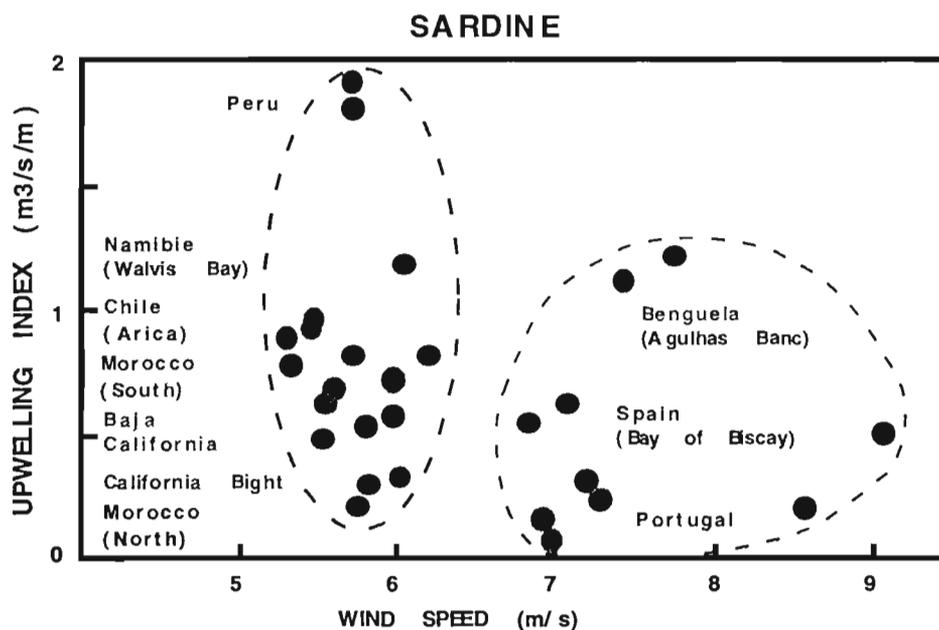


Figure 8: Vitesse du vent et indice d'upwelling observés au cours des saisons de ponte pour la sardine dans différents écosystèmes côtiers (d'après Shin, 1995).

Plusieurs travaux s'intéressant à la variabilité du recrutement dans différentes régions ont pu être réalisés à l'aide de COADS, permettant ainsi une généralisation de concept de la fenêtre optimale (Cury *et al.*, 1995; Roy *et al.*, 1992; Roy *et al.*, 1995). Concernant les bouleversements écologiques apparus au cours des années 1980 en Côte d'Ivoire, l'hypothèse émise par Pezennec et Bard (1992) a pu être confirmée (Roy, 1995).

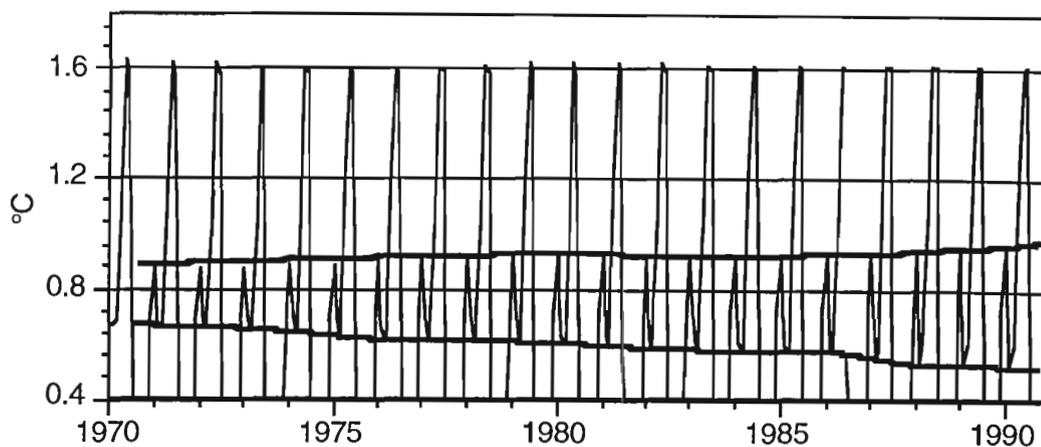


Figure 9 : Zoom de la série temporelle de TSM devant la Côte d'Ivoire (composante saisonnière) montrant l'intensification de l'upwelling au cours de la Petite Saison Froide depuis les années 1970. Cette intensification a contribué à l'établissement d'une seconde saison de ponte en début d'année, favorisant ainsi l'accroissement très important de la population de sardinelles (d'après Roy, 1995).

## VI. Conclusion

L'ensemble des travaux ayant fait appel à COADS ne peut être détaillé ici, ces quelques exemples permettent d'illustrer quelques uns des domaines qui ont été abordés. De la climatologie à l'halieutique, les domaines d'applications de COADS à travers le réseau CEOS ont été extrêmement variés. L'apport de la base données COADS a été très important pour CEOS et ce jeu de données a stimulé de manière significative les activités scientifiques menées à travers le réseau CEOS.

Une centaine d'exemplaires du jeu de CD-ROM ont été produits. La distribution à travers la communauté scientifique a été assurée par la NOAA pour le Pacifique et par l'ORSTOM pour l'Atlantique et l'Océan Indien. Nous avons reçu de nombreuses demandes (plus d'une cinquantaine) pour cet outil, émanant de chercheurs individuels, de laboratoires de recherche ou d'organismes internationaux. Un frein important à la diffusion du produit a été de ne pouvoir diffuser une version PC de l'interface d'interrogation (les CD-ROM sont par contre accessibles à partir de nombreuses plateformes: PC, Macintosh ou Unix). Cette lacune sera comblée prochainement (début 1997).

La mise à jour de la base de données n'a pas été envisagée de manière concrète. Depuis le démarrage du projet, le NCAR a diffusé de nouveaux produits. Ce n'était évidemment pas du ressort de CEOS d'assurer la mise à jour des CD-ROM et de l'interface. Une activité minimale a cependant été assurée. La nouvelle version de COADS

pour la période 1980-1993 est accessible sur disque à la NOAA (PFEG-Monterey) et au centre ORSTOM de Brest. Le programme d'interrogation a été également modifié pour s'adapter au nouveau format des données. Un financement reste cependant à obtenir pour assurer le transfert sous CD-ROM et leur diffusion. Ces activités ne sont cependant pas du ressort de CEOS mais plutôt des activités d'animation et de valorisation scientifique de l'ORSTOM.

## **VII. Bibliographie**

- Bakun, A. 1990. *Global climate change and intensification of coastal ocean upwelling*. *Science*, 12 January 1990, Vol. 247: 198-201.
- Bakun A., V. Christensen, C. Curtis, P. Cury, M.H. Durand, D. Husby, R. Mendelssohn, J. Mendo, R. Parrish, D. Pauly and C. Roy. 1993. *The Climate and Eastern Ocean Systems project*. *Naga*, 15(4):26-30.
- Cury, P. and C. Roy, (eds). 1991. *Pêcheries Ouest-africaines: variabilité, instabilité et changement*. ORSTOM, PARIS, 525p.
- Cury P., C. Roy, R. Mendelssohn, A. Bakun and D. M. Husby. 1995. *Moderate is better: exploring nonlinear climatic effect on the Californian northern anchovy*. In R.J. Beamish [ed.] *Climate change and northern fish populations*. Can. Spec. Publ. Fish. Aquat. Sci. 121, p41-424.
- Diaz H.F., K. Wolter and S.D. Woodruff. 1992. *Proceedings of an international COADS Workshop, Boulder, Colorado, January 13-15 1992*. US Department of Commerce, NOAA, 390p.
- Diaz H.F. and H.J. Isemer, 1994. . *Proceedings of an international COADS winds Workshop, Kiel, Germany, 31 May-2 June 1994*. . US Department of Commerce, NOAA - Institut für Meereskunde, Kiel.
- Fletcher J.O., J. Slutz and S.D. Woodruff, 1983. *Towards a Comprehensive Ocean Atmosphere Data set*. *Trop. Ocean-Atmos. Newsllett.*, 20, 13-14.
- Folland C.K. and J. Hsiung, 1987. *Correction of seasonally varying biases in uninsulated bucket sea surface temperature data using a physical model*. *Met. Office Synoptic Climatology Branch Memo No. 154*.

- Folland C.K. and D.E. Parker. 1990. *Observed variations of sea surface temperature.*  
In : *Climate-Ocean interaction*, M.E. Schlesinger ed., NATO workshop,  
Oxford, Kluwer Academic Publishers, 21-52.
- Houghton J.T., Jenkins G.J. and J.J. Ephraums, eds. 1990. *Climate change, the IPCC  
scientific assesment.* Cambridge University Press. 365p.
- Jones P.D. and T.M.L. Wigley, 1992. *Corrections to pre-1941 SST measurements for  
studies of long-term changes in SSTs.* In : *Proceedings of the internatio-  
nal COADS workshop, Boulder, Colorado, 13-15 January 1992*, H. F.  
Diaz, K. Wolter and S.D. Woodruff Ed., US Department of Commerce,  
NOAA. 227-237.
- Laurent C., 1996. *Détermination de la variabilité climatique (1945-1990) dans deux  
principales zones d'upwelling à partir de la base de données COADS.*  
Rapport de DEA. UBO Brest. 66p.
- National Climate Data Center, 1968. *TDF-11 reference manual.* NCDC, Asheville,  
N.C., USA.
- Pezennec O. and F.X. Bard, 1992. *Importance écologique de la petite saison d'upwel-  
ling ivoiro-ghananéenne et changements dans la pêche de Sardinella  
aurita.* Aquat. Living Ressour., 5, 249-259.
- Roy C. 1995. *The Cote d'Ivoire and Ghana coastal upwellings : dynamics and  
Change.* In : *Dynamique et usage des ressources en sardinelles de  
l'upwelling côtier du Ghana et de la Côte d'Ivoire.* F.X. Bard et K.A.  
Koranteng eds., ORSTOM éditions Paris. p346-361.
- Roy C., P. Cury et S. Kifani, 1992. *Pelagic fish recruitment success and reproductive  
strategy in upwelling areas : environmental compromises.* In : *Benguela  
Trophic Functioning.* Payne, A.I.L., Brink, K.H., Mann, K.H. and R.  
Hilborn (Eds). S. Afr. J. mar. Sci. , 12 : 135-146.
- Roy C., C. Porteiro and J. Cabanas. 1995. *The Optimal Environmental Window Hy-  
pothesis in the ICES Area: The Example of the Iberian Sardine.* in : *Dy-  
namics of upwelling in the ICES area.* ICES coop. res. rep., 206. p57-65.

- Shin Y.J. 1995. *Stratégies de reproduction des Clupéoidés des zones d'upwelling côtier: tentative de généralisation. Rapp. DEA. ENSA Rennes.68p.*
- Slutz, R.J. S.J. Lubker, J. D. Hiscox, S.D. Woodruff, R. L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms. 1985. *Comprehensive Ocean-atmosphere Data Set; Release 1. NOAA Environmental Research Laboratories, Climate Research Program, Boulder, CO, 268.*
- Woodruff S.D. 1986. *Proceedings of a COADS Workshop, Boulder, Colorado, January 22-24 1986. NOAA Environmental Research Laboratories, Climate Research program, Boulder, Colo., 218p.*
- Woodruff, S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer. 1987. *A Comprehensive Ocean-atmosphere Data Set. Bull. Amer. Meteor. Soc., 68, 1239-1250.*
- Woodruff S.D., S.J. Lubker, K. Wolter, S.J. Worley and J.D. Elms. 1993. *Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-1992. Earth System Monitor. US. Dept. of Commerce. NOAA. 40p.*