

REVUE DE STATISTIQUE APPLIQUÉE

BERNARD BURTSCHY

JACQUES SERVAIN

Modélisation statistique de la température de surface de la mer

Revue de statistique appliquée, tome 34, n° 3 (1986), p. 21-47

http://www.numdam.org/item?id=RSA_1986__34_3_21_0

© Société française de statistique, 1986, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

MODÉLISATION STATISTIQUE DE LA TEMPÉRATURE DE SURFACE DE LA MER

Bernard BURTSCHY ⁽¹⁾ et Jacques SERVAIN ⁽²⁾

*(1) Ecole Nationale Supérieure des Télécommunications,
46, rue Barrault, 75634 Paris Cedex 13.*

*(2) Laboratoire d'Océanographie Physique,
Université de Bretagne Occidentale,
6, avenue Victor Le Gorgeu, 29200 Brest.*

Présente Affiliation : Antenne ORSTOM, IFREMER
BP 337
22273 Brest-Cedex

RÉSUMÉ

Des séries mensuelles de température océanique et atmosphérique à l'interface air-mer sont analysées sur une zone de l'Atlantique Nord-Est, proche du plateau continental européen. L'objectif est d'établir la meilleure technique statistique pour prévoir l'évolution de la température de surface de la mer. Les séries temporelles utilisées couvrent les années 1947-1982.

Différents jeux de coefficients de modélisation (méthodes ARIMA univariées et à fonctions de transferts) sont calculés sur la période 1947-1979; la performance prévisionnelle (pour 1 à 3 mois d'avance) est testée sur les années 1980 à 1982. Afin de minimiser l'erreur de prévision de la température de surface de l'océan, il est nécessaire de quantifier les fluctuations saisonnières de l'impact des anomalies mensuelles passées à l'aide d'une méthode originale basée sur le transfert saisonnier optimal. Cette erreur est réduite pour les mois où la température de surface de la mer est particulièrement stable (hiver-été). La forte inertie thermique du paramètre océanique explique que l'auto-transfert soit généralement plus important que le transfert croisé entre les deux températures. Des situations inverses se produisent toutefois pendant les mois d'établissement ou de destruction de la thermocline saisonnière. Ce dernier résultat peut s'expliquer par une sensibilité prononcée de l'océan superficiel à la variabilité atmosphérique durant ces périodes. D'autre part, le caractère stochastique de la saisonnalité (par rapport au caractère déterministe) est mis en évidence. Pour la prévision de l'anomalie mensuelle de surface de la mer avec un mois d'avance, la méthode optimum proposée permet une réduction de moitié de l'écart-type résiduel.

Mots clés : Modélisation statistique, Coefficients de transfert, Variation saisonnière, Température de surface.

Key words: Statistical modelling, Transfer coefficients, Seasonal variation, Sea surface temperature.

INTRODUCTION

Dans le Nord-Est de l'océan Atlantique, le contenu calorifique superficiel est principalement régi par le bilan thermique à l'interface air-mer. La part due à l'advection est réduite; elle représente moins de 15 % aux abords du plateau continental. Le bilan thermique étant lui-même essentiellement radiatif, la température de surface de la mer suit une évolution dominée par un cycle saisonnier. En se référant à un signal annuel moyen calculé sur une longue période, on peut définir les anomalies successives de la température de surface. Les séries chronologiques ainsi composées représentent la variabilité thermique interannuelle.

Dans le golfe de Gascogne, les courants de surface sont faibles et tourbillonnaires, ce qui sous-tend une modification lente des caractéristiques hydrologiques des masses d'eau. Cette stationnarité, associée à la grande inertie calorifique de la mer, permet un caractère conservatif des anomalies thermiques de l'océan superficiel (KRAUS et MORRISON, 1966; REYNOLDS, 1979). Une propriété similaire n'existe évidemment pas pour les masses d'air atmosphériques. Toutefois, lorsque l'on étudie le système océan-atmosphère de manière couplée, il est parfois possible de déceler certains phénomènes conservatifs au niveau de l'interface. La liaison thermodynamique s'effectue dans les deux sens et il s'ensuit des actions et des rétro-actions entre les deux milieux pouvant entretenir pendant quelques mois certaines anomalies climatiques (FRANKIGNOUL et HASSELMAN, 1977; SERVAIN, 1980).

Nous nous proposons, dans un but prévisionnel de l'anomalie mensuelle de la température de surface de la mer, d'exploiter statistiquement les propriétés conservatives de celle-ci, en s'aidant simultanément de sa liaison avec la température de l'air. Pour cet objectif, nous analysons les séries chronologiques formées à partir d'observations de ces deux paramètres pendant la période 1947-1982 sur une zone de l'Atlantique proche des côtes françaises. Différentes méthodes sont utilisées, mettant en jeu l'analyse des autorégressions simple et croisée des deux températures de surface. L'impact saisonnier du couplage thermique océan-atmosphère est mesuré. Il nous permet d'affiner un modèle statistique de prévision de la température de surface de la mer. Dans la discussion finale, nous établissons les aboutissements d'une telle étude.

1. DISCUSSION DES PARAMÈTRES UTILISÉS

Les données de climatologie de surface utilisées ici proviennent de la collecte des rapports météorologiques de navires marchands de toutes nationalités. Une première source d'informations fut achetée au Environmental Data Service, Asheville N.C. (U.S.A.) et regroupe des données de 1947 à 1972. Une seconde source nous a été communiquée par le Service de Climatologie Marine de la Météorologie Française et complète les observations jusqu'en décembre 1982. La zone retenue, volontairement restreinte (2° de latitude par 2° de longitude), est située aux abords du plateau continental le long du rail de navigation Brest-La Corogne (Fig. 1).

Les relevés météorologiques des navires comportent nombre d'informations, concernant notamment les températures de surface de la mer et de l'air, le vent et la pression barométrique. Une étude préliminaire (SERVAIN, 1979) a montré cependant qu'à une échelle de temps de l'ordre du mois, la pression et la force du vent étaient peu corrélées avec la température de surface de la mer. Il reste qu'en zone océanique proche du littoral européen les deux seuls paramètres, vraiment représentatifs de la liaison climatologique moyenne mensuelle entre l'océan et l'atmosphère, sont les températures de surface de la mer et de l'air (Fig. 2).

Les évolutions moyennes annuelles des deux températures de surface suivent un cycle similaire, avec un maximum thermique en été qui apparaît en

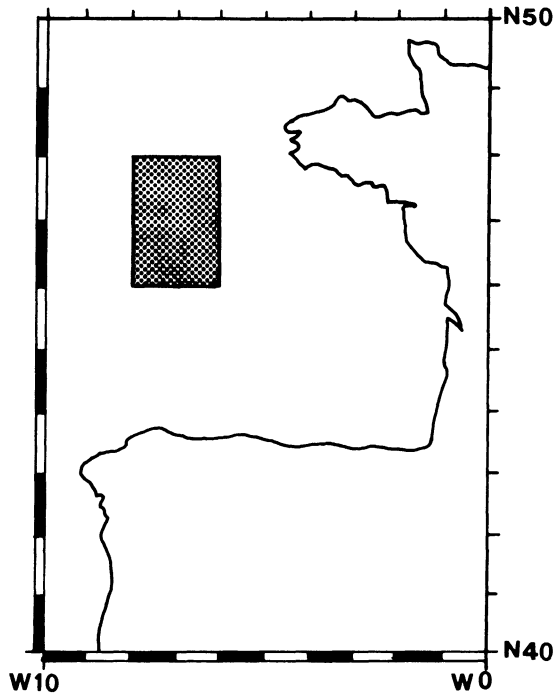


Figure 1. — La zone géographique utilisée est limitée en latitude par les parallèles 46°N, 48°N et en longitude par les méridiens 6°W, 8°W.

phase (mi-août) et un minimum thermique en hiver décalé d'environ un mois entre le signal de la température de l'air (mi-février) et celui de la température de la mer (mi-mars). De fin avril à septembre la mer est légèrement plus froide que l'air. C'est pendant cette période que l'océan accumule son énergie calorifique en provenance de l'atmosphère. Pendant le reste de l'année, le signal thermique de la surface de la mer est supérieur à celui de l'air, favorisant ainsi le transfert énergétique de l'océan vers l'atmosphère par le flux de chaleur sensible et latent.

Dans la suite du texte, on notera par M et A les séries se rapportant respectivement à la température de surface de la mer et à la température de surface de l'air. L'écart-type de l'anomalie mensuelle de la température de surface de la mer ($S_M \neq 0.6^\circ\text{C}$) est pratiquement toujours inférieur à l'écart-type équivalent pour la température de l'air ($S_A \neq 0.8^\circ\text{C}$). Ceci marque le caractère plus stationnaire de la thermodynamique marine par rapport à celle de l'atmosphère. Nous notons toutefois une modulation saisonnière dans le rapport S_M/S_A . Celui-ci dépasse rarement 0.5 en hiver, alors qu'en été il avoisine l'unité. Ce dernier résultat s'explique par des mouvements atmosphériques moins prononcés en saison estivale, provoquant une diminution de la variance de la température de l'air. Par contre, durant cette saison, le contenu thermique de l'océan est concentré sur une couche de mélange peu épaisse au-dessus de la thermocline, ce qui facilite une plus grande variabilité de la température de la surface de la mer et, nous le verrons par la suite, une plus grande sensibilité aux variations rapides de l'atmosphère.

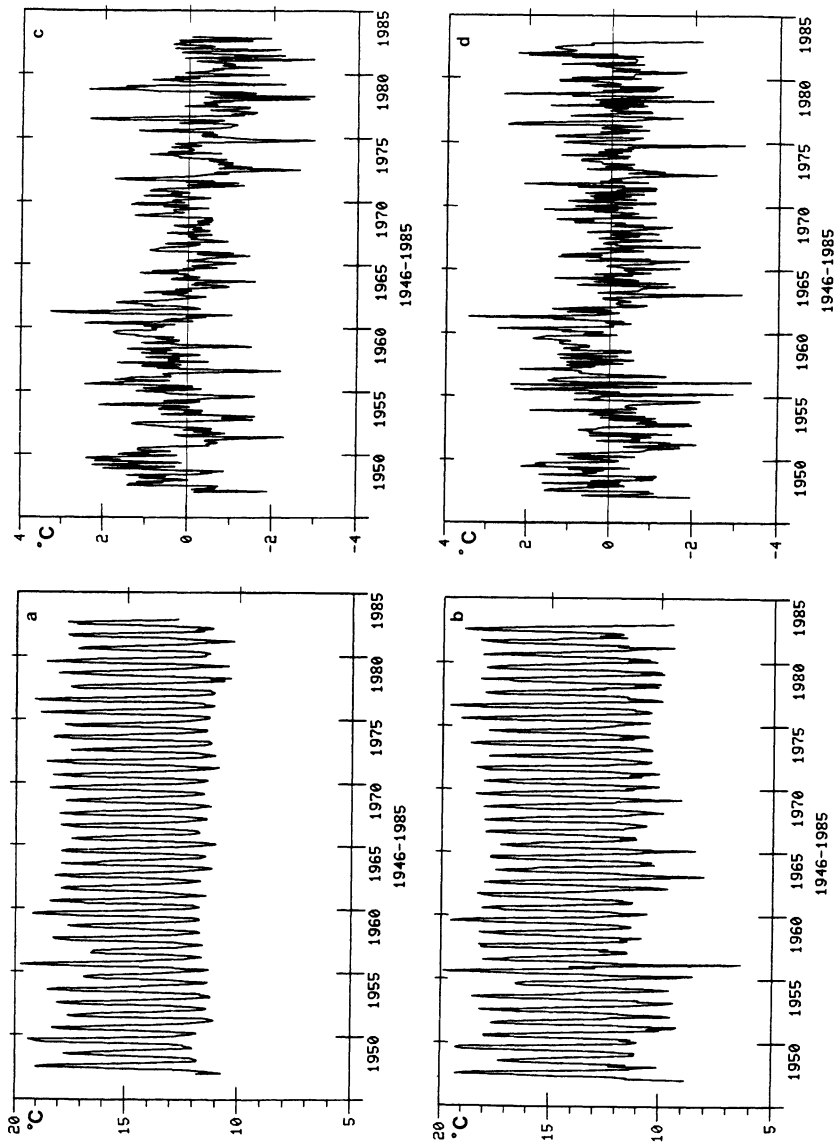


Figure 2. — Séries brutes de la température de surface de la mer (a) et de l'air (b).
 Séries des anomalies normées de la température de surface de la mer (c) et de l'air (d).
 Les tirets sur l'axe temporel indiquent le mois de décembre de l'année en regard.

2. LA MÉTHODOLOGIE D'ANALYSE

2.1. L'analyse des séries chronologiques

Une série chronologique est un ensemble d'observations ordonnées dans le temps; en se restreignant aux séries chronologiques discrètes, finies, avec des observations régulièrement et également espacées, on notera les séries $Z_1, Z_2, \dots, Z_t, \dots$

Il est assez naturel d'analyser la manière dont un instant t est lié à ses p instants précédents représentés par $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$ ce qui revient à calculer l'autorégression (d'ordre p) de la série :

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + a_t$$

où a_t représente l'erreur au temps t .

Ce modèle peut être généralisé (BOX et JENKINS, 1968) en introduisant les réactions aux q perturbations subies par le modèle aux instants $t - 1, t - 2, \dots, t - q$ et représentées par les erreurs $a_{t-1}, a_{t-2}, \dots, a_{t-q}$:

$$Z_t = \varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p} + a_t - \theta_1 a_{t-1} \dots - \theta_q a_{t-q}$$

Ce modèle est appelé un modèle ARMA(p, q) i.e. AutoRegressive Moving Average ou encore AutoRégression à Moyenne Ajustée. Les ordres p et q du modèle sont déterminés grâce à l'examen des coefficients d'autocorrélation de la série : l'autocorrélation $\rho_{ZZ}(k)$ avec décalage k de la série Z mesure la corrélation entre les observations Z_t et Z_{t-k} .

L'analyse d'un tel modèle, impliquant une seule série Z , est nommée analyse univariée.

La liaison entre deux séries Z et X peut être modélisée par deux équations :

$$Z_t = \sum_{i=0}^{\infty} a_i X_{t-i} + \sum_{j=1}^{\infty} b_j Z_{t-j} + e(Z_t)$$

$$X_t = \sum_{i=0}^{\infty} c_i Z_{t-i} + \sum_{j=1}^{\infty} d_j X_{t-j} + e(X_t)$$

Où $e(Z_t)$ représente l'erreur du modèle de la série Z (de même pour X).

Dans le cas le plus général, ces deux équations ne sont pas indépendantes l'une de l'autre et il faut les estimer simultanément selon un modèle bivarié.

L'étude de la liaison entre deux séries implique un nouvel outil d'analyse, la corrélation croisée $\rho_{ZX}(k)$ avec décalage k . Mais l'interprétation directe de la corrélation croisée est impossible car elle dépend des autocorrélations des deux séries (GRANGER et NEWBOLD, 1974). Les autocorrélations sont éliminées en calculant la corrélation croisée sur les résidus des séries après modélisation univariée de chacune d'entre elles (HAUGH, 1976; HAUGH et BOX, 1977) : cette nouvelle corrélation est appelée corrélation croisée « blanchie », alors que la corrélation croisée issue du calcul direct sur les deux séries, sans transformation aucune, est nommée corrélation croisée « non blanchie ».

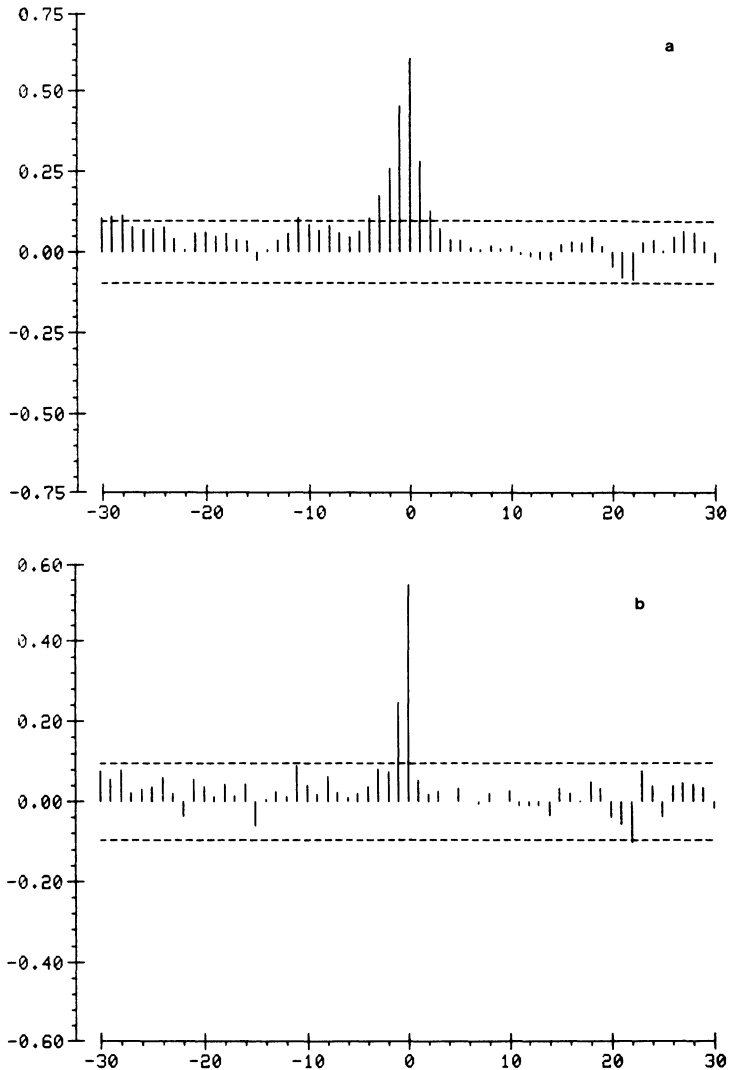


Figure 3. — Fonctions de corrélations croisées « non blanchie » (a) et « blanchie » (b) entre les séries des anomalies de température de surface de la mer et de l'air. La ligne en pointillé indique le seuil de l'intervalle de confiance à 95 %. La partie gauche des tracés est relative à l'influence de la température de l'air sur la température de l'eau. La partie droite représente l'influence inverse.

Il peut arriver qu'une seule des deux influences (soit X sur Z, soit Z sur X) soit significative et qu'il n'y ait pas de « feedback » : à l'examen des corrélations croisées blanchies $\rho_{ZX}(k)$, on constaterait alors la nullité (au sens statistique) de ces corrélations soit pour $k > 0$, soit pour $k < 0$. La nullité pour $k = 0$ impliquerait l'absence de liaison instantanée entre Z_t et X_t (i.e. sans décalage) et la nullité pour tous les décalages, l'absence de liaison.

Dans le cas où une seule des deux influences est significative (par exemple X sur Z), l'équation obtenue s'apparente à une régression, qu'on appellera le « transfert ». Exemple de transfert :

$$Z_t = a X_{t-1} + b Z_{t-1} + e_t$$

La figure 3 montre un exemple de corrélations croisées non blanchie et blanchie entre les deux séries étudiées. Il est bien évident que seule la partie inférieure de la figure (la corrélation blanchie) nous est utile pour l'interprétation, la corrélation non blanchie ne figurant qu'à titre pédagogique. On observe l'existence d'une importante corrélation instantanée (0.54) entre M_t et A_t . D'autre part, la liaison entre M_t et A_{t-1} est significative (0.25) et l'on note l'absence de tout « feedback » c'est-à-dire de liaison entre, par exemple, M_{t-1} et A_t . Ceci autorise une modélisation par une équation unique du style $M_t = f(M_{t-1}, A_t, A_{t-1})$.

2.2. Le problème de la saisonnalité des séries

La forte saisonnalité des chroniques étudiées ici complique singulièrement l'analyse sur le plan statistique. Aussi est-il d'usage dans ce cas d'éliminer a priori la variation saisonnière afin d'en extraire « l'anomalie ». La définition même de « l'anomalie » est délicate. Il peut s'agir en effet, soit d'une différence d'amplitude entre le signal d'une année spécifique et celui de l'année moyenne, soit d'un déphasage entre ces deux signaux, soit plus généralement d'une juxtaposition des deux cas précédents. S'il n'existe pas de méthode parfaite dans le domaine de la désaisonnalisation, certaines techniques sont plus efficaces que d'autres selon les circonstances. Celle de BOX et JENKINS (1970) traite dynamiquement la saisonnalité par différentiation :

$$\nabla_{12} Z_t = Z_t - Z_{t-12}$$

La saisonnalité est dans ce cas traitée stochastiquement et la série, auparavant non stationnaire, apparaît beaucoup plus simple à traiter. Les dangers d'une telle simplification ont été mis en évidence par SLUTSKY (1927) : en particulier, s'il existe des tendances ou des saisonnalités déterministes, celles-ci seront complètement masquées et des modèles inadéquats peuvent être identifiés. La vérification de tels symptômes se fait aisément avec le modèle AR(2) qui a des propriétés périodiques lorsque les racines du polynôme associé sont complexes. Soit le modèle AR(2) :

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + a_t$$

$$Z_t (1 - \varphi_1 B - \varphi_2 B^2) = a_t$$

où B est l'opérateur de retard : $B Z_t = Z_{t-1}$.

Si les racines du polynôme associé en B sont complexes ($\varphi_1^2 + 4 \varphi_2 < 0$) les caractéristiques du signal définies par sa fréquence f_0 et par son amortissement d sont données (BOX et JENKINS, 1970) par :

$$d = \sqrt{-\varphi_2}$$

$$\cos 2 \pi f_0 = \frac{|\varphi_1|}{2 \sqrt{-\varphi_2}}$$

S'agissant d'une série mensuelle (période 12 donc $f_0 = 1/12$) dont l'amortissement est constant ($d = 1$), on a

$$\varphi_2 = -1$$

$$\cos\left(\frac{2\pi}{12}\right) = \frac{\sqrt{3}}{2} \quad \text{d'où} \quad \varphi_1 = \sqrt{3}$$

Le modèle saisonnier s'écrira sous la forme :

$$(1 - \sqrt{3} \cdot B + B^2) Z_t = a_t$$

$$\text{où } B Z_t = Z_{t-1}$$

Si le calcul des racines du polynôme d'un modèle saisonnier fait apparaître des racines complexes proches de 1 en valeur absolue, il est légitime de différencier. Mais si deux racines ou plus sont supérieures à 1 en valeur absolue, il est fort probable qu'une composante déterministe soit présente.

Pour éliminer la variation saisonnière, une autre méthode consiste à soustraire à chaque terme de la série la moyenne mensuelle correspondante calculée sur l'ensemble de l'historique (HIPEL et Mc LEOD, 1980) :

$$W1_{i,j} = Z_{i,j} - M_j$$

$$\text{où } M_j = \frac{1}{k} \sum_i Z_{i,j} \quad j = 1,12$$

Cette méthode, appelée « méthode des moyennes », est particulièrement simple mais possède deux inconvénients majeurs : elle ne tient pas compte des différences entre variances mensuelles et elle suppose que la série n'a pas de tendance.

Le premier inconvénient est aisé à supprimer : il suffit de diviser par l'écart type du mois correspondant :

$$W2_{i,j} = \frac{Z_{i,j} - M_j}{E_j}$$

où E_j représente l'écart type du mois j .

$$E_j^2 = \frac{1}{k} \sum_i (Z_{i,j} - M_j)^2$$

Cette méthode nécessite, pour une série mensuelle, 12 estimations de moyenne et 12 estimations d'écart type, soit 24 paramètres. L'estimation de ce grand nombre de paramètres peut être réduite en les approximant par les séries de Fourier (PARZEN et PAGANO, 1979). Les formules permettant d'estimer les coefficients de Fourier et leur application à l'estimation des M_j et E_j sont bien connues dans la littérature (CROLEY et RAO, 1977). L'utilisation de toutes les composantes de Fourier est équivalente à l'estimation des M_j et E_j .

S'il peut paraître légitime de considérer a priori l'absence de tendance, il ne faut pas oublier que toute la notion d'anomalie en dépend. Cette absence de tendance constitue une hypothèse qu'il est possible de vérifier à l'aide d'une méthode de désaisonnalisation telle que X11 (SHISKIN *et al.*, 1967) : X11, méthode de désaisonnalisation très utilisée en économie, recherche la

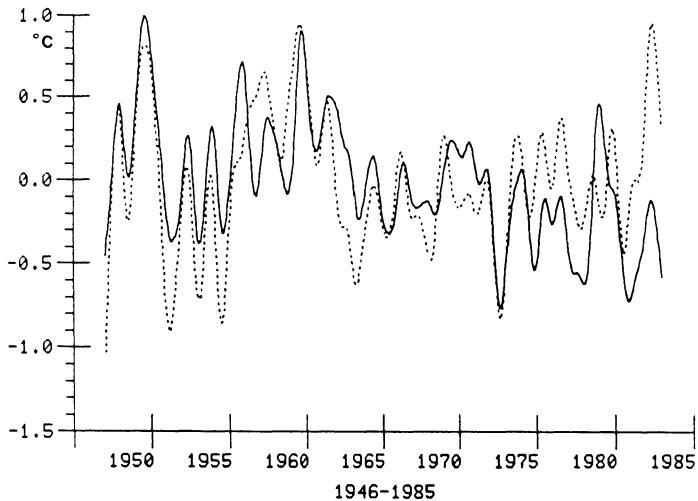


Figure 4. — Tendence (par la méthode X11) de la température de surface de la mer (trait continu) et celle de l'air (trait discontinu).
Les tirets sur l'axe temporel indiquant le mois de décembre de l'année en regard.

tendance d'une série par applications répétées de filtres pondérés basés sur la moyenne mobile. Dans le tracé de la tendance des deux températures de surface par X11 (Fig. 4), une relation étroite dans le domaine des basses fréquences est mise en évidence. On note en particulier une forte oscillation proche de deux ans (l'Oscillation Quasi-Biennale, OQB, déjà remarquée de nombreuses fois lors d'études antérieures, ANGELL et KORSHOVER, 1974; ZVEREV, 1977; SERVAIN, 1979). Les deux séries semblent aussi être modulées par une fluctuation dont la période serait une dizaine d'années. Pour la température de surface de la mer, nous notons une légère décroissance proche de quatre dixièmes de degrés entre les années 1950 et les années 1980. Ce refroidissement fait suite au lent réchauffement séculaire qui a vu son apogée vers 1950 (FIEUX et STOMMEL, 1975; SERVAIN, 1976). En revanche il n'existe apparemment aucune tendance linéaire entre les deux extrémités de la série de la température de l'air. On se gardera d'analyser les transferts entre séries désaisonnalisées, les mauvaises identifications étant fréquentes (NEWBOLD, 1980). La saisonnalité fait partie intégrante du transfert.

La recherche de la saisonnalité pose un problème plus général : il s'agit de déterminer qu'elles sont les parts respectives des composantes déterministes ou stochastiques de cette saisonnalité (HAVENER et SWAMY, 1981), étant bien entendu que la méthode des moyennes est un moyen de supprimer la composante déterministe en préservant la saisonnalité. On appellera dans la suite *anomalie centrée* la série dont on a retiré les moyennes mensuelles respectives et *anomalie normée* lorsque de surcroît elle a été divisée par l'écart-type correspondant. Les séries qui n'auront pas subi ce traitement seront identifiées dans la suite sous le nom de séries *brutes*, séries que l'on sera amené à différencier pour préserver leur stationnarité.

Dans le but de vérifier le déterminisme de la composante saisonnière et d'en mesurer toutes les conséquences, toutes les méthodes seront testées sur les trois types de séries (brutes, centrées, normées).

2.3. Modèles univariés et à fonctions de transferts

L'objectif de l'étude est l'analyse et la prévision de la série de la température de surface de la mer M. Dans une première étape, l'analyse univariée de cette série sera effectuée par la méthode ARIMA (AutoRegressive Integrated Moving Average). Cette analyse sera effectuée, comme expliqué ci-dessus, sur la série brute et les deux séries d'anomalies (centrée et normée). Le test prévisionnel de l'extrapolation permettra d'établir une première base d'analyse.

La série A est susceptible d'être explicative de M. La fonction de corrélation croisée sera utilisée afin de détecter les décalages et la forme d'influence. Cependant, un calcul correct de la corrélation croisée nécessite que les deux séries soient préalablement « blanchies » de leur autocorrélation (HAUGH et BOX, 1977) ce qui induit une analyse univariée de A. Le modèle à fonctions de transfert obtenu sera testé prévisionnellement et confronté au modèle univarié pur.

3. LES RÉSULTATS UNIVARIÉS ET À FONCTIONS DE TRANSFERTS

3.1. Les résultats univariés

Les deux séries brutes M et A ont été rendues stationnaires à l'aide d'une différenciation saisonnière dont l'utilisation paraît s'imposer sans ambiguïté à l'examen visuel des autocorrélations. Le caractère stochastique de la saisonnalité (cf. 2.2) est confirmé par le calcul des racines du polynôme associé au modèle correspondant : les racines sont proches de 1 en valeur absolue. Aussi le raisonnement portera non sur la série de départ mais sur l'accroissement saisonnier :

$$\nabla_{12} Z_t = Z_t - Z_{t-12}$$

Les séries des anomalies ne présentent aucune difficulté d'identification. Il en est tout autrement des modèles sur séries différenciées : les modèles, beaucoup plus complexes à identifier, sont aussi plus riches (tableau I). Ils font très nettement apparaître l'oscillation quasi-biennale de 24 mois citée ci-dessus.

Si l'on examine les deux modèles (plus simples) relatifs à l'anomalie normée, on obtient respectivement pour M et A les modèles suivants :

$$\begin{aligned} M_t &= 0.62 M_{t-1} + a_t & S_a &= 0.78 \\ A_t &= 0.38 A_{t-1} + a_t & S_a &= 0.92 \end{aligned}$$

S_a représente l'écart-type résiduel (calculé sur les a_t).

TABLEAU I
La modélisation univariée

<p style="text-align: center;"><u>Ia Température de surface de la mer M</u></p> <p style="text-align: center;"><u>Série brute différenciée</u></p> $M_t = 0.63 M_{t-1} - 0.78 M_{t-12} + a_t - 0.33 a_{t-12} - 0.66 a_{t-24}$ <p style="text-align: center;">réduction écart-type : 0.85 à 0.48</p> <p style="text-align: center;"><u>Anomalie Centrée</u></p> $M_t = 0.63 M_{t-1} - 0.09 a_{t-12} + a_t$ <p style="text-align: center;">réduction écart-type : 0.61 à 0.47</p> <p style="text-align: center;"><u>Anomalie Normée</u></p> $M_t = 0.62 M_{t-1} + a_t$ <p style="text-align: center;">réduction écart-type : 1.00 à 0.78</p>
<p style="text-align: center;"><u>Ib Température de surface de l'air A</u></p> <p style="text-align: center;"><u>Série brute différenciée</u></p> $A_t = 0.35 A_{t-1} - 0.87 A_{t-12} + a_t - 0.15 a_{t-12} - 0.83 a_{t-24}$ <p style="text-align: center;">réduction écart-type : 1.15 à 0.79</p> <p style="text-align: center;"><u>Anomalie Centrée</u></p> $A_t = 0.35 A_{t-1} + a_t$ <p style="text-align: center;">réduction écart-type : 0.83 à 0.78</p> <p style="text-align: center;"><u>Anomalie Normée</u></p> $A_t = 0.38 A_{t-1} + a_t$ <p style="text-align: center;">réduction écart-type : 1.00 à 0.92</p>

Ces deux modèles sont des modèles autorégressifs d'ordre 1 avec des inerties différentes. Une variation de 1° à l'instant t se répercutera de la façon suivante les mois suivants :

	M	A
t	1	1
t + 1	0.62	0.38
t + 2	$(0.62)^2 = 0.38$	$(0.38)^2 = 0.14$
t + 3	$(0.62)^3 = 0.24$	$(0.38)^3 = 0.05$

Une meilleure réduction de la variance et une meilleure performance prévisionnelle de la température de surface de la mer en découlent naturellement, ce qui rejoint le phénomène physique d'une plus grande capacité de mémorisation de l'océan.

3.2. Les fonctions de transferts

La question posée est alors de déterminer si la connaissance de A permet d'améliorer significativement la prévision de M. Nous avons vu précédemment qu'après examen des corrélations croisées blanchies entre les deux séries, il pouvait être envisagé une modélisation de M suivant l'équation $M_t = f(M_{t-1}, A_t, A_{t-1})$.

Deux conceptions s'opposent alors. La conception descriptive modélise cette liaison afin de bien tenir compte de la nature de l'influence (tableau IIa).

Ainsi pour l'anomalie normée, nous aurons la relation :

$$M_t = 0.47 A_t - 0.21 A_{t-1} + 0.54 M_{t-1} + a_t \quad S_a = 0.64$$

Le gain en écart type résiduel est significatif car il est passé de 0.78 (univarié) à 0.64. La prévision ne s'est cependant guère améliorée car pour prévoir M_{t+1} , il faut prévoir A_{t+1} ce qui n'est guère efficace en raison de la faible inertie de A. L'autre conception (tableau IIb) est plus prévisionnelle. Lors du calcul de la prévision de M à l'instant t + 1, on ne connaît pas la valeur de A_{t+1} . On identifie alors un modèle *décalé* entre M_t et A_{t-1} en ignorant le terme de liaison instantané A_t . Il ne s'agit certes pas du meilleur modèle au sens de la description mais la corrélation croisée est significative. On a ainsi le modèle suivant pour l'anomalie normée :

$$M_t = 0.11 A_{t-1} + 0.59 M_{t-1} + a_t \quad S_a = 0.77$$

La réduction de variance est à peine sensible et la performance prévisionnelle guère meilleure. Force est donc de constater que si l'impact instantané de A_t sur M_t est significatif, il est peu efficace du point de vue prévisionnel à cause de la faible inertie de A. Ceci illustre bien le fait que ce qui est efficace pour la modélisation ne l'est pas forcément pour la prévision. En conséquence, le modèle basé sur les fonctions de transferts, s'il permet de bien comprendre la nature de l'impact de la température de l'air, n'est pas compétitif d'un point de vue prévisionnel par rapport au modèle univarié pur.

TABLEAU II
Les équations du transfert

<p><u>I Ia Transfert instantané</u></p> <p><u>Série brute différenciée</u></p> $M_t = 0.37 A_t - 0.14 A_{t-1} + 0.58 M_{t-1} + a_t$ <p>Ecart-type résiduel : 0.39</p> <p><u>Anomalie Centrée</u></p> $M_t = 0.33 A_t - 0.15 A_{t-1} + 0.54 M_{t-1} + a_t$ <p>Ecart-type résiduel : 0.39</p> <p><u>Anomalie Normée</u></p> $M_t = 0.47 A_t - 0.21 A_{t-1} + 0.54 M_{t-1} + a_t$ <p>Ecart-type résiduel : 0.64</p>
<p><u>I Ib Transfert décalé</u></p> <p><u>Série brute différenciée</u></p> $M_t = 0.09 A_{t-1} + 0.52 M_{t-1} + a_t$ <p>Ecart-type résiduel : 0.48</p> <p><u>Anomalie centrée</u></p> $M_t = 0.08 A_{t-1} + 0.59 M_{t-1} + a_t$ <p>Ecart-type résiduel : 0.47</p> <p><u>Anomalie normée</u></p> $M_t = 0.11 A_{t-1} + 0.59 M_{t-1} + a_t$ <p>Ecart-type résiduel : 0.77</p>

4. LE TRANSFERT SAISONNIER

4.1. Vers une nouvelle forme de transfert

Comme toutes les méthodes appliquées à des séries saisonnières, la méthode des fonctions de transfert a l'inconvénient de ne pas individualiser les mois et, par conséquent, de calculer une sorte de comportement moyen.

TABLEAU III
Le transfert saisonnier

	A_t	A_{t-1}	M_{t-1}	S_a
Eq. base	0.47	-0.21	0.54	0.64
Mois 1	0.32	-0.04	0.63	0.77
2	0.27	-0.12	0.64	0.79
3	0.36	-0.11	0.61	0.77
4	0.65	-0.07	0.32	0.57
5	0.74	-0.10	0.39	0.50
6	0.68	0.05	0.26	0.56
7	0.82	-0.03	0.46	0.52
8	0.63	-0.47	0.68	0.52
9	0.81	0.07	0.26	0.37
10	0.56	0.17	0.27	0.59
11	0.22	0.10	0.56	0.71
12	0.23	0.03	0.67	0.71

Or l'examen approfondi des résidus des modèles précédents montre que l'influence de l'anomalie varie selon le mois considéré. Cette constatation conduit à réestimer les divers coefficients de l'équation de transfert mois par mois.

Le tableau III montre la variation mensuelle des coefficients de l'équation de base de l'anomalie normée :

$$M_t = 0.47 A_t - 0.21 A_{t-1} + 0.54 M_{t-1} + a_t$$

Les coefficients des divers modèles varient dans de grandes proportions et l'examen des écarts-types résiduels montre que certains mois sont mieux modélisés que les autres.

Le tableau IV montre la variation mensuelle des coefficients de l'équation décalée de l'anomalie normée :

$$M_t = 0.11 A_{t-1} + 0.59 M_{t-1} + a_t$$

4.2. La régression pas à pas à saisonnalité évolutive

Plusieurs coefficients des tableaux III et IV ne sont pas statistiquement significatifs. D'autre part, l'application des fonctions d'autocorrélations et d'autocorrélations partielles sur les séries résiduelles montre que certaines ne sont pas aléatoires. Ceci signifie concrètement que ce ne sont pas seulement

TABLEAU IV
Le transfert saisonnier décalé

	A_{t-1}	M_{t-1}	S_a
Eq. base	0.11	0.59	0.77
Mois 1	0.06	0.61	0.81
2	-0.02	0.55	0.82
3	0.19	0.58	0.84
4	0.02	0.63	0.81
5	0.26	0.37	0.83
6	0.48	0.18	0.82
7	0.42	0.27	0.79
8	-0.25	0.92	0.69
9	0.38	0.24	0.83
10	0.26	0.43	0.78
11	0.08	0.62	0.74
12	0.06	0.71	0.73

TABLEAU V
Transferts saisonniers optimaux

	Equation	(S_a)
Mois 1	$0.31 M_{t-1} + 0.38 M_{t-9} + 0.38 M_{t-2}$	(0.70)
2	$0.48 M_{t-1} + 0.29 M_{t-9}$	(0.77)
3	$0.80 M_{t-1} - 0.56 M_{t-2} + 0.42 M_{t-11}$	(0.68)
4	$0.62 M_{t-1} + 0.38 M_{t-3}$	(0.70)
5	$0.65 M_{t-2}$	(0.78)
6	$0.66 A_{t-1}$	(0.81)
7	$0.92 M_{t-1} - 0.46 A_{t-2} - 0.26 M_{t-10}$	(0.68)
8	$0.68 M_{t-1} - 0.26 M_{t-12}$	(0.65)
9	$0.57 A_{t-1}$	(0.83)
10	$0.26 A_{t-1} + 0.43 M_{t-1}$	(0.78)
11	$0.69 M_{t-1} - 0.18 M_{t-4}$	(0.71)
12	$0.73 M_{t-1}$	(0.72)

les coefficients qui varient selon les mois, la structure du modèle restant fixe, mais qu'il faut aussi remettre en cause la forme de l'équation.

En fait rien n'empêche d'identifier une équation différente pour chaque mois selon la méthodologie des transferts si ce n'est que cette approche est fastidieuse : pour des données mensuelles, il faudrait identifier et estimer 12 équations. Cette approche est possible car nous avons la chance de disposer d'une série suffisamment longue. Les différentes équations de ce modèle, que l'on pourrait appeler modèle à équations variables, sont résumées dans le tableau V.

Comme les séries sont toutes fondamentalement autorégressives, une automatisation devient possible grâce à l'autorégression pas à pas optimale telle qu'elle est préconisée par GRANGER et NEWBOLD (1977, p. 176). Cela revient à chercher parmi tous les décalages le sous-ensemble significatif. Pour des raisons physiques évidentes, nous avons limité la recherche à 12 décalages mensuels pour M et 3 pour A. Des décalages supérieurs n'auraient aucun sens et n'introduiraient que des liaisons purement fortuites. Notons que cette méthode originale donne les mêmes résultats qu'une identification équation par équation en ayant l'avantage de l'automatisme.

Cette approche peut encore être améliorée grâce à une utilisation judicieuse de l'analyse en composantes principales (les composantes mineures, cf. Annexe).

4.3. Discussion des transferts saisonniers optimaux

En comparant les écarts-types des erreurs (S_a) donnés par les tableaux IV et V, on observe que notre méthode d'optimisation des équations du transfert apporte une amélioration sensible sept mois sur douze, cinq pendant la saison froide (janvier à mai) et deux pendant la saison chaude (juillet et août). La valeur de S_a diminue de 0.11 pour les mois de janvier, avril et juillet, elle diminue de 0.16 pour le mois de mars. La modélisation optimale de la température de surface de la mer a lieu pour les mois où la situation thermique superficielle est bien établie, c'est-à-dire mars ($S_a = 0.68$) et août ($S_a = 0.65$). Il n'est pas étonnant de constater que cette température de surface demeure difficile à modéliser pendant les mois de forte transition thermique, à savoir juin ($S_a = 0.81$) et septembre ($S_a = 0.83$).

Une schématisation des principaux caractères des transferts saisonniers optimaux est visualisée sur la figure 5. Une lecture attentive en apporte les enseignements suivants :

— L'autorégression du premier ordre de la température de surface de la mer est prépondérante 9 mois sur 12. Les 3 mois restants sont mai, juin et septembre, c'est-à-dire pendant l'établissement et la disparition de la thermocline. L'autorégression d'ordres supérieurs (2 ou 3) demeure importante pendant les mois d'hiver et de printemps. Une influence notable est signalée entre les anomalies de la température de surface de la mer des mois d'avril-mai et les anomalies thermiques de l'hiver suivant (janvier, février, mars). Tous ces points sont sans doute à relier au fait que la structure

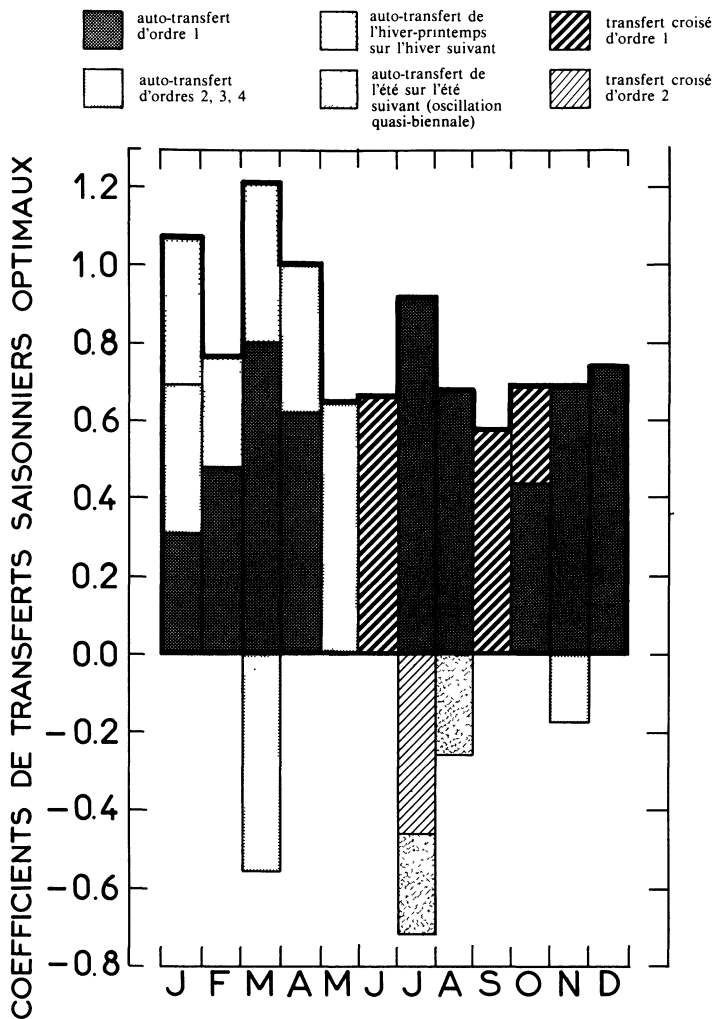


Figure 5. — Schématisation des coefficients de transferts saisonniers optimaux pour la modélisation mensuelle de la température de surface de la mer.

thermique de la mer est globalement homogène sur plusieurs dizaines de mètres pendant la saison froide. Si de fortes anomalies de température sont générées durant cette période, le volant thermique de la couche superficielle de l'océan augmente. Une telle propriété peut ainsi faciliter la transmission d'une partie de ces fortes anomalies de température lors des mois suivants, avant et après l'apparition de la thermocline saisonnière, cette dernière jouant ici le rôle d'écran thermique pendant la période estivale.

— L'oscillation quasi-biennale est présente sous la forme de coefficients de transferts négatifs entre la température de surface de la mer d'un été et celle de l'été de l'année suivante. Ce résultat est conforme à des études plus anciennes (DAVIS, 1967; WRIGHT, 1968) qui tendaient à démontrer

une plus grande sensibilité de la saison estivale à cette période de deux ans. Une analyse en composantes principales menée sur les différents mois de l'année (non montrée ici) nous a permis de vérifier cette hypothèse, notamment pendant la période 1950-1965.

— Le fait le plus intrigant est sans doute l'apparition de coefficients de transferts très significatifs entre la température de l'air et la température de la mer pour la modélisation de cette dernière pendant les mois de juin-juillet et de septembre-octobre. Il arrive même qu'en juin et en septembre ces seuls coefficients soient présents, reléguant par là même tout transfert de la température de la mer sur elle-même. Toutefois, ceci ne doit pas faire illusion. Nous avons vu précédemment qu'il était malaisé de modéliser la température de surface de l'océan pendant les mois de transition thermique. Il est clair que pendant ces périodes, la structure thermique superficielle de la mer est fragile. Elle perd beaucoup de sa qualité de mémorisation, ceci n'étant pas compensé qu'en partie par une plus grande sensibilité à l'influence de la variabilité thermique atmosphérique.

5. LA PERFORMANCE PRÉVISIONNELLE

5.1. Les indicateurs

La modélisation de séries chronologiques simultanées et de leurs liaisons est intéressante car elle permet de mieux cerner leur comportement. Mais si la description est un des objectifs de l'étude, le but ultime reste la performance prévisionnelle. Se pose alors inévitablement la définition des critères de mesure, des périodes d'estimation et de tests prévisionnels.

Notons Z_t la valeur réelle de la série à l'instant t et $Z(t; t + i)$ la prévision faite à partir de l'observation t (i.e. on connaît la série jusqu'à t) pour l'instant $t + i$.

L'erreur commise à $t + i$ d'une prévision effectuée à t s'écrit :

$$E(t; t + i) = Z_{t+i} - Z(t; t + i)$$

Deux paramètres sont à définir. Le premier est le nombre P de mois en amont pour la prévision. Le deuxième paramètre à définir est la longueur de la période de test de la prévision : soit $[1, T]$ la période de modélisation, par exemple [janvier 1947, décembre 1979]. La période de test prévisionnel sera [janvier 1980, décembre 1982]. La difficulté de jugement de la prévision provient du fait qu'on ne sait jamais si la valeur réelle de comparaison de prévision est une valeur que l'on pourrait qualifier de « normale » ou si, au contraire, elle est un peu exceptionnelle : dans ce cas, il est difficile de déterminer qui, de la prévision ou de la valeur de référence, est mauvaise. Toute l'ambiguïté du jugement de la qualité d'une prévision tient en ce dilemme.

Aussi, dans la suite, trois périodes seront choisies pour tester la prévision :

[janvier 80, décembre 80] : 12 mois

[janvier 80, décembre 81] : 24 mois

[janvier 80, décembre 82] : 36 mois

Trois horizons prévisionnels seront considérés : les prévisions à 1, 2 et 3 mois à l'avance.

Les indicateurs prévisionnels s'écriront sous la forme :

$$\sum_{t=T+1}^{T+H} \sum_{i=1}^P f(E(t; t+i))$$

P : horizon prévisionnel (1, 2, 3 mois)

H : la période test (12, 24, 36 mois).

Ainsi constitué, l'indicateur est un indicateur glissant. Dans le cas d'une prévision trois mois à l'avance, en décembre 79, on prévoira janvier, février et mars 80 et on calculera la fonction d'erreur; puis en janvier 80, on prévoira février, mars, avril etc.

Il ne reste plus maintenant qu'à définir la fonction d'erreur f . La mesure la plus communément utilisée est la Moyenne de la Somme des carrés des Erreurs (MSE), la fonction f représentant le carré de l'erreur :

$$MSE = \frac{1}{P \cdot H} \sum_{t=T+1}^{T+H} \sum_{i=1}^P E^2(t; t+i)$$

Cet indicateur, utilisé dans la plupart des travaux théoriques et appliqués, ne permet pas l'utilisation de tests car ils supposent que les erreurs aient les propriétés d'un échantillon (en particulier vis-à-vis de l'indépendance des erreurs), ce qui n'est jamais le cas.

Une alternative pour l'évaluation de la performance prévisionnelle est fournie par la Moyenne des valeurs Absolues des Erreurs (MAE), la fonction f étant la fonction valeur absolue.

$$MAE = \frac{1}{P \cdot H} \sum_{t=T+1}^{T+H} \sum_{i=1}^P |E(t; t+i)|$$

Cet indicateur privilégie une fonction linéaire des erreurs au lieu d'une fonction quadratique, c'est-à-dire une moins grande sensibilité aux grandes erreurs.

Enfin, une Moyenne des Erreurs est souvent calculée afin de fournir un indicateur de dérèglement ME :

$$ME = \frac{1}{P \cdot H} \sum_{t=T+1}^{T+H} \sum_{i=1}^P E(t; t+i)$$

Dans la mesure où il n'existe pas de mesure simple et universellement acceptée, tous les tests prévisionnels ci-dessous seront jugés à l'aide de ces trois critères.

5.2. La prévision à une période

Le tableau VI résume les prévisions obtenues avec un horizon de un mois à l'avance sur la période [janvier 1980, décembre 1982], soit 36 mois. Les indicateurs prévisionnels, définis en 5.1, sont la Moyenne de la Somme des carrés des Erreurs (MSE), la Moyenne des valeurs Absolues des Erreurs (MAE), et la Moyenne des Erreurs (ME). Les modèles ont été estimés sur la période [janvier 1947, décembre 1979]. A la lecture du tableau VI, on peut formuler les remarques suivantes :

— la première remarque est la plus grande efficacité prévisionnelle des modèles basés sur un traitement stochastique de la saisonnalité (c'est-à-dire en différenciant la série brute) par rapport aux modèles utilisant la notion d'anomalie (tableau VIa) : les indicateurs prévisionnels de la série brute (MSE = 0.29, MAE = 0.41 pour la mer) sont systématiquement meilleurs que pour les deux séries d'anomalies. Cette conclusion n'est pas a priori évidente car les phénomènes physiques s'accordent généralement mieux d'un traitement déterministe de la tendance et de la saisonnalité, les modèles ainsi obtenus étant beaucoup plus simples à identifier. Elle rejoint les conclusions de PLOSSER (1979) sur d'autres types de séries, montrant une plus grande efficacité prévisionnelle des modèles stochastiques (sauf cas particuliers d'inversibilité) au détriment d'une plus grande difficulté d'identification du modèle.

— La seconde remarque est le très faible gain obtenu en intégrant par une fonction de transfert la variable explicative atmosphérique par rapport à une modélisation univariée (tableau VIb) : les indicateurs MSE et MAE restent pratiquement constants entre la partie a et la partie b du tableau VI, ce qui rejoint la remarque du paragraphe 3.2 notant le peu d'intérêt d'introduire une variable explicative par fonction de transfert classique.

— Mais dès que l'on module cette fonction de transfert par mois (tableau VIc), le gain, inégal selon les mois comme il a été montré ci-dessus, devient significatif. L'indicateur MSE passe ainsi de 1.17 à 0.95 pour le transfert saisonnier décalé et à 0.55 pour le transfert saisonnier optimal (tableau VI d). Nous sommes, à ce niveau, à la limite de prévisibilité de la série. N'oublions pas que les données proviennent de la collecte de relevés météorologiques de navires marchands de toutes nationalités et que ces relevés sont de qualité diverse; l'écart-type pour un mois de ces observations en nombres variables n'est jamais inférieur au degré.

— Enfin, l'indicateur ME, la moyenne des erreurs, c'est-à-dire un indicateur de dérèglement, est toujours négatif pour la température de la mer quelle que soit la méthode utilisée. Cela traduit la légère décroissance, constatée lors de l'étude de la tendance (Fig. 4), de quelques dixièmes de degrés entre les années 1950 et les années 1980 de la température de la mer.

5.3. Horizon prévisionnel et période de test

Les modèles décrits précédemment comportaient essentiellement un décalage du premier ordre soit sur l'autorégression soit sur le transfert. De

TABLEAU VI
 Comparaison des prévisions
 (Test sur janvier 80-décembre 82 avec l'anomalie normée)

	MSE	MAE	ME	S _{init}	S _a
<u>VIa MODELISATION UNIVARIEE</u>					
<u>Mer</u>					
Série brute dif	0.29	0.41	-0.15	0.85	0.48
An. centrée	0.31	0.42	-0.16	0.61	0.47
An. normée	1.15	0.75	-0.29	1.00	0.78
<u>Air</u>					
Série brute dif	0.64	0.65	0.01	1.15	0.79
An. centrée	0.64	0.65	0.03	0.83	0.78
An. normée	0.93	0.80	0.01	1.00	0.92
<u>VIb MODELE A FONCTION DE TRANSFERT</u>					
<u>Transfert instantané</u>					
Série brute dif	0.29	0.41	-0.15	0.85	0.39
An. centrée	0.32	0.32	-0.20	0.61	0.39
An. normée	1.17	0.77	-0.34	1.00	0.64
<u>Transfert décalé</u>					
Série brute dif	0.29	0.42	-0.11	0.85	0.48
An. centrée	0.32	0.43	-0.17	0.61	0.47
An. normée	1.17	0.76	-0.30	1.00	0.77
<u>VIc TRANSFERT SAISONNIER</u>					
<u>Transfert saisonnier instantané</u>					
An. normée	0.85	0.65	-0.10	1.00	0.55
<u>Transfert saisonnier décalé</u>					
An. normée	0.95	0.70	-0.13	1.00	0.71
<u>VId TRANSFERT SAISONNIER OPTIMAL</u>					
Série brute dif	0.25	0.39	-0.08	0.85	0.35
An. normée	0.55	0.54	-0.08	1.00	0.56

TABLEAU VII
Horizon prévisionnel et période de test

VIIa L'HORIZON PREVISIONNEL (Test sur la période Janvier 80 - Décembre 82)						
Prévision à	1 période		2 périodes		3 périodes	
	MSE	MAE	MSE	MAE	MSE	MAE
<u>Modèle univarié</u>						
Série brute dif.	0.29	0.41	0.35	0.48	0.35	0.49
An. normée	1.15	0.75	1.31	0.84	1.32	0.84
<u>Fonction de transfert décalée</u>						
Série brute dif.	0.29	0.42	0.35	0.48	0.36	0.48
An. normée	1.17	0.76	1.30	0.84	1.30	0.85
<u>Transfert saisonnier optimal</u>						
Série brute dif.	0.25	0.39	0.28	0.44	0.29	0.44
An. normée	0.55	0.54	0.66	0.68	0.71	0.67
VIIb L'INFLUENCE DE LA PERIODE DE TEST (Fonction de transfert décalée)						
Prévision à	1 période		2 périodes		3 périodes	
	MSE	MAE	MSE	MAE	MSE	MAE
<u>Période de test</u>						
1980-1982	1.17	0.76	1.30	0.84	1.30	0.85
1980-1981	0.92	0.60	1.02	0.71	1.08	0.76
1980	1.38	0.81	1.41	0.87	1.22	0.83

ce fait, la simple application de la formule du modèle fournit une prévision à une période à l'avance. Mais toute prévision à un horizon supérieur à une période doit intégrer, pour pouvoir être calculée, les prévisions déjà calculées, ce qui conduit automatiquement à une dégradation de la performance prévisionnelle.

Le tableau VIIa résume, pour les principaux modèles, les prévisions à une, deux et trois périodes à l'avance. La dégradation entre une et deux périodes à l'avance est nette; la faible différence entre les prévisions à deux

et à trois périodes à l'avance provient de l'effet de stabilisation des modèles autorégressifs aux ordres supérieurs (cf. 3.1 pour une illustration). On remarquera que cet effet est beaucoup moins net pour le modèle à transfert saisonnier optimal en raison de la variété des décalages selon les mois.

Le choix de la période de test est d'une grande importance pratique car les indicateurs de l'efficacité prévisionnelle sont affectés par l'existence de points irréguliers. Autant dire qu'il n'existe pas de période de test idéale. Tous les résultats présentés jusqu'ici étaient relatifs à la période [1980, 1982].

Le tableau VIIb montre la variation des indicateurs prévisionnels pour la méthode des fonctions de transferts décalées selon que l'on choisit comme période de test 1980 à 1982, 1980 à 1981 ou encore l'année 1980. Cette variation est non négligeable. Si elle n'affecte en rien les conclusions sur les diverses formes de modélisation effectuées qui restent valables, car fondées comme il se doit sur l'écart-type résiduel, elle illustre bien la difficulté de jugement de la qualité d'une prévision. Néanmoins, cette variation s'explique par les variations de la tendance déjà mises en évidence sur la figure 4.

CONCLUSION

La forte inertie thermique de la couche superficielle océanique dans l'Atlantique Nord-Est est confirmée. Une étroite relation instantanée est observée entre la température de surface de la mer et la température de l'air à l'interface océan-atmosphère. Une liaison significative est notée entre l'anomalie de température de l'air à l'instant t et l'anomalie de la température de surface de la mer à l'instant $t + 1$ mois. Le phénomène inverse n'apparaît pas. Les deux séries de longues durées utilisées (36 ans de 1947 à 1982) sont marquées par un cycle quasi-biennal et sont modulées par une plus longue période d'une dizaine d'années. On note une légère tendance linéaire négative de la température de surface de la mer (0.4°C entre les deux extrémités de la série temporelle). Cette tendance est absente sur la série de la température de l'air.

De multiples formes de modélisation de l'anomalie mensuelle de la température de surface de la mer ont été entreprises sur la période de référence 1947-1980. Trois types de modélisations ont été analysés. Le premier type consiste à prévoir la température de la mer en utilisant uniquement la connaissance du passé de cette série : la meilleure modélisation (et la meilleure prévision) est réalisée, non pas en se servant de la notion d'anomalie communément admise, mais en exploitant la série brute afin de garder le caractère stochastique de la tendance et de la saisonnalité.

La deuxième forme de modélisation s'apparente à la technique classique de régression en intégrant simultanément la connaissance de la série de la température de surface de l'océan et celle de la série de la température atmosphérique à l'interface air-mer. Cette modélisation s'avère peu convaincante, les progrès réalisés dans la prévision n'étant pas décisifs.

Une méthode originale faisant intervenir les coefficients de transferts saisonniers optimaux entre les températures de surface de l'air et de l'eau a permis d'améliorer très sensiblement la modélisation de l'anomalie mensuelle de la température de surface de la mer et a montré que la prise en compte du paramètre thermique pouvait se révéler judicieuse dans certains cas. Trois groupes de coefficients de transferts saisonniers significatifs ont été mesurés. Les deux premiers groupes intéressent une autorégression du paramètre océanique : l'un est relatif à la saison froide et les coefficients de transfert associés (d'ordre un à quelques mois, en majorité positifs), indiquent une forte influence des anomalies hivernales sur les mois suivants; l'autre groupe apparaît pour la saison estivale, le signe négatif de ces coefficients faisant état de l'oscillation quasi-biennale observée principalement durant cette époque de l'année. Une influence significative de l'anomalie de la température de l'air permet de relever un troisième groupe de coefficients de transfert saisonnier optimaux; ces coefficients dépassent le seuil de signification uniquement pour le printemps et l'automne, périodes pendant lesquelles la structure thermique de la couche superficielle océanique est la plus sensible aux variations de températures de l'atmosphère.

Les mois de forte stabilité thermique de la couche superficielle de l'océan, c'est-à-dire pendant l'hiver et l'été, sont assez facilement modélisables, la forte capacité calorifique de l'eau jouant ici le rôle de régulateur thermique. Les mois de transition thermique (établissement de la thermocline au printemps et destruction de celle-ci en automne) sont en revanche plus difficiles à modéliser.

La performance prévisionnelle pour un à trois mois en avance est testée sur la période 1980-1982. Les meilleurs résultats sont bien sûr obtenus pour l'horizon égal à un mois. Ceci s'explique d'autant plus que les différentes formes de modélisation utilisées sont principalement régies par un décalage du premier ordre. Si la dégradation de la prévision est nette entre l'horizon un mois et les horizons 2 et 3 mois, il n'existe en revanche que peu de différences entre ces deux derniers cas; on arrive ici à la limite de la prévisibilité de la série.

En définitive, lorsque l'on utilise l'horizon un mois à l'avance pour la prévision statistique de l'anomalie thermique de la surface de la mer, la méthode originale proposée mettant en jeu une modulation mensuelle des fonction de transfert, permet une réduction de moitié de l'écart-type résiduel. Cette modélisation est facile à mettre en œuvre (peu de coefficients à déterminer) et pourrait être adaptée en temps réel à partir des observations de navires météorologiques au point fixe sur l'Atlantique Nord-Est. Nous pensons que cette méthode pourrait aussi être testée utilement dans un proche avenir en exploitant les données satellites de la mesure de la température de surface de la mer.

ANNEXE

La méthode des transferts saisonniers

La méthode consiste à identifier et estimer une équation de transfert par mois. Plus précisément, et en se référant à l'exemple traité, la méthode revient à chercher une fonction, variant chaque mois dans son équation et ses valeurs de coefficients, de la forme :

$$M_t = f(M_{t-1}, \dots, M_{t-12}, A_{t-1}, \dots, A_{t-3})$$

Dans la mesure où les séries sont suffisamment longues, il suffit d'appliquer le principe de la fonction de transfert classique à chaque mois. La méthode aboutit sans ambiguïté au résultat final : mais elle est fastidieuse car l'identification est difficilement automatisable et il faut identifier la fonction pour chaque mois.

Un premier pas vers l'automatisation consiste à utiliser l'autorégression multiple optimale (GRANGER et NEWBOLD, 1977), c'est-à-dire une régression multiple avec sélection automatique des variables exogènes parmi les M_{t-j} ($j = 1, 12$) et les A_{t-k} ($k = 1, 3$). L'application numérique de cette méthode fournit des résultats très proches des résultats précédents, identiques pour la structure de l'équation, proches pour la valeur des coefficients. Une estimation des équations par un algorithme de type MARQUARDT fournit des résultats identiques.

Cette méthode s'avère difficilement généralisable au cas vectoriel, c'est-à-dire lorsque l'on dispose de plusieurs séries simultanément. Le principe de la méthode proposée utilise les composantes mineures (les vecteurs propres associés aux plus petites valeurs propres) d'une analyse en composantes principales (A.C.P.). L'utilisation classique de l'analyse en composantes principales est de résumer l'information grâce aux composantes majeures (les vecteurs propres associés aux plus grandes valeurs propres). Mais les composantes mineures ont aussi leur intérêt car elles permettent la détection des relations entre variables : ainsi l'introduction de variables redondantes, qui a pour conséquence directe l'inflation des plus grandes valeurs propres, peut être détectée dans les composantes mineures. Les vecteurs propres associés aux valeurs propres nulles définissent un système linéaire exact entre variables, système qui peut être résolu.

De même, les composantes principales associées aux valeurs propres proches de zéro permettent d'identifier les variables qui peuvent être expliquées avec peu d'erreurs par les autres variables. On se sert à cet effet des composantes normées qui possèdent des variables dont les valeurs sont relativement grandes, ces variables sont appelées variables à potentiel prédictif. La rotation par transformation orthogonale de HOUSEHOLDER permet de déterminer les coefficients de la régression.

Cette méthode permet d'identifier les modèles ARMA vectoriels en généralisant les théorèmes de TSAY et TIAO (1985) démontrés dans le cas univariés.

Elle permet aussi de résoudre, à l'aide d'une seule A.C.P., les divers transferts saisonniers. La matrice de l'A.C.P. se présente avec, en lignes, les différentes années et, en colonnes, les 12 mois des séries M_t , M_{t-1} , A_t , A_{t-1} , soit 48 variables. Avec les hypothèses de stationnarité des séries et d'ordres finis des AR les décrivant, la matrice des autocovariances (qui est une matrice de HANKEL) est parfaitement définie. Les vecteurs propres normalisés associés aux plus petites valeurs propres non nulles fournissent la solution. Dans l'exemple considéré, la solution est très proche des solutions de l'autorégression optimale mais avec une difficulté, celle du test des coefficients. Une reestimation des coefficients fournit la solution.

Remerciements :

Cette étude fut partiellement financée par les conventions de Recherche CNEXO-UBO n° 83-2839 et 84-3149.

BIBLIOGRAPHIE

- ANGELL J.K. and J. KORSHOVER (1974). — Quasi-Biennial and Long-Term Fluctuations in the Centers of Action. *Monthly Weather Review*, 102, 10, pp. 668-678.
- BOX G.E.P. and G.M. JENKINS (1968). — Some Recent Advances in Forecasting and Control, Part I, *Applied Statistics*, 17, 91-109.
- BOX G.E.P. et G.M. JENKINS (1970). — *Time Series Analysis, Forecasting and Control*, Rev. ed. 1976 (Holden Day, San Francisco, CA).
- CROLEY T.E. and K.N.R. RAO (1977). — *A Manual for Hydrologic Time Series Deseasonalization and Serial Independence Reduction*. Iowa Institute of Hydraulic Research.
- DAVIS N.E. (1967). — The summers of North-West Europe. *Meteorol. Mag.*, 96, pp. 178-187.
- FIEUX M. and H. STOMMEL (1975). — Preliminary Look at Feasibility of using Marine Reports of Sea Surface Temperature for Documenting Climatic Change in the Western North Atlantic. *Journal of Marine Research*, 33, pp. 83-95.
- FRANKIGNOUL C. and K. HASSELMAN (1977). — Stochastic Climate Models, Part. 2. Application to Sea Surface Temperature Anomalies and Thermocline Variability. *Tellus*, 29, pp. 289-305.
- GRANGER C.W.J. and P. NEWBOLD (1974). — Spurious Regression in Econometrics. *Journal of Econometrics*, pp. 111-120.
- GRANGER C.W.J. and P. NEWBOLD (1977). — *Forecasting Economic Time Series*. (Academic Press, New York).
- HAUGH L.D. (1976). — Checking the Independence of Two Covariance-Stationary Time Series: a Univariate Residual Cross Correlation Approach. *Journal of the American Statistical Association*, 71, pp. 378-385.

- HAUGH L.D. and G.E.P. BOX (1977). — Identification of Dynamic Regression (Dynamic Lag) Models Connecting Two Time Series. *Journal of the American Statistical Association*, 72, pp. 121-130.
- HAVENER A. and P.A.V.B. SWAMY (1981). — A random Coefficient Approach to Seasonal Adjustment of Economic Time Series, *Journal of Econometrics*, 15, pp. 177-209.
- HIPEL K.W. and A.I. Mc LEOD (1980). — Perspectives in Stochastic Hydrology in ANDERSON O.D. (ed.), *Times Series*, (North Holland; Amsterdam).
- KRAUS E.B. and R.E. MORRISON (1966). — Local interaction between the sea and the air at monthly and annual time scales. *Q.J.R. Meteorol. Soc.*, 92, pp. 114-127.
- NEWBOLD P. (1980). — A Note on Relations between Seasonally Adjusted Variables. *Journal of Time Series Analysis*, 1, 1, pp. 31-35.
- PARZEN E. and M. PAGANO (1979). — An Approach to Modelling Seasonally Time Series. *Journal of Econometrics*, 9, pp. 137-153.
- PLOSSER C.I. (1979). — Short-term Forecasting and Seasonal Adjustment, *Journal of the American Statistical Association*, V75, 365, pp. 15-24.
- REYNOLDS R.W. (1979). — A Stochastic Forcing Model of Sea Surface Temperature Anomalies in the North Pacific and North Atlantic. *Climatic Research Institute and Department of Atmospheric Sciences*, Oregon State University, Corvallis, Oregon 97331, p. 23.
- SERVAIN J. (1976). — Etude statistique de l'ensemble des données hydrologiques recueillies depuis le début du siècle en Mer Celtique. *Thèse de 3^e cycle*, Université de Bretagne Occidentale, Brest.
- SERVAIN J. (1979). — *Sur l'analyse statistique des données climatologiques de surface du Proche Atlantique*. Université de Bretagne Occidentale, Contrat CNEXO 79/2002, p. 101.
- SERVAIN J. (1980). — Analyse corrélative des températures de surface sur le Proche Atlantique. *Oceanolog. Acta*, 3, 1, pp. 3-6.
- SHISKIN J., A.H. YOUNG and J.C. MUSGRAVE (1967). — The X-11 Variant of the Census Method-II Seasonal Adjustment Program. *Technical Paper*, 15, U.S. Bureau of the Census.
- SLUTSKY E. (1927). — The Summation of Random Causes as the Source of Cyclic Processes, in *Russian translated in Econometrica*, 5, 1937, p. 105.
- TSAY R.S. and G.C. TIAO (1985). — Use of Canonical Analysis in Time Series Model Identification. *Biometrika*, 72, pp. 299-316.
- WRIGHT P.B. (1968). — A Widespread Biennial Oscillation in the Troposphere. *Weather*, 23, pp. 50-54.
- ZVEREV A.A. (1977). — Manifestation of Quasi-Biennial Cycles in Water Temperature Fluctuations in the North Atlantic. *Oceanology*, 17, 6, pp. 640-644.