35· Relating Fisheries to the Environment in the Gulf of Guinea: Information, Causality and Long-Term Memory

ROY MENDELSSOHN

Pacific Fisheries Environmental Group,
Southwest Fisheries Center, NMFS, NOAA,
P.O. Box 831, Monterey, CA 93942, USA.

Abstract

"Optimal transformations" are new statistical methods that greatly improve our understanding of non-linear relationships that may exist between environmental and ecological data. The effect of "optimal transformations" of environmental data is examined for tuna catches, and for the small pelagic fishery in the Gulf of Guinea. In the frequency domain it is shown that the "optimal transformation" changes the spectrum of the environmental variable so that it is more similar to the spectrum of CPUE (Catch Per Unit of Effort). When there is more than one environmental variable, methods are presented for determining the unique information in each variable that is relevant to the analysis, and for determining if the environmental variables carry information that is not contained in the CPUE series itself. The analysis makes clear why for tuna in some areas wind is most important, while in other areas wind and SST are important. For the small pelagics, the analysis explains why only SST (Sea Surface Temperature) is important for predicting the combined species CPUE, but SST and salinity are both important in predicting the CPUE for *S. maderensis* only. Evidence is presented that the environmental variables important to the fisheries in the Gulf of Guinea have "long-term memory". This has implications for our ability to accurately predict into the future the consequences of our management actions.

Résumé

Les «transformations optimales» sont de nouvelles méthodes statistiques qui améliorent beaucoup notre compréhension des relations non-linéaires qui peuvent exister entre des données environnementales et écologiques. L'effet des «transformations optimales» des données environnementales est examiné pour les captures de thons et de petits pélagiques dans le Golfe de Guinée. Dans le domaine fréquentiel les transformations optimales changent le spectre de la variable environnementale de telle façon qu'il apparaît plus semblable à celui des CPUE (Capture Par Unité d'Effort). Quand plus d'une seule variable environnementale est considérée, des méthodes sont présentées pour déterminer l'information, pertinente pour l'analyse, contenue dans chacune des variables et pour déterminer si les variables environnementales apportent de l'information par rapport aux séries de CPUE. L'analyse révèle pourquoi, pour le thon, le vent est la variable la plus importante dans certaines zones, tandis que pour d'autres il s'agit du vent et de la SST (température de surface de la mer). Pour les petits poissons pélagiques, l'analyse montre pourquoi la SST est importante pour expliquer les CPUE toutes espèces confondues alors que c'est la SST et la salinité qui permettent de prédire la CPUE de *S. maderensis*. On montre enfin que les variables environnementales importantes pour les pêcheries, dans le Golfe de Guinée, possèdent une «mémoire à long terme». Ceci a des implications pour notre faculté à prédire les conséquences des actions d'aménagement.

Introduction

Fisheries management models traditionally have concentrated on the equilibrium behavior of catch and effort in isolation from the environment in which the fish live. In the Gulf of Guinea, as elsewhere, there is mounting evidence that the ocean environment plays a key role in the observed dynamics of fish stocks. A list of references documenting this statement can be found throughout the papers in this volume. I mention in particular Mendelssohn and Roy (1986), and Mendelssohn and Cury (1987, 1989), because the techniques and results of those papers are directly relevant to the concerns of this paper.

Studies of the ocean environment's effects on fish dynamics have occurred at several levels. At the lowest level, studies attempt to determine the direct mechanisms that cause mortality, reproductive success, or growth in a fish. For example, Lasker's work (1978) on the survival of larvae suggests that the concentration of food sources within the first several days of life are essential for survival of the larvae. Such information, if true, provides understanding of the mechanisms by which the ocean affects the fish, and thereby possible limitations of any higher level models. To make practical use of this information, however, would require knowing where the larvae are located and measuring the food concentrations right after birth, a formidable and expensive task.

The next level of study has traditionally looked at contemporaneous environmental conditions in relation to fish catch or fish recruitment. (This approach has been extended, for example, in the papers cited above, by considering the dynamics of both the environment and the fish in space and time). In most of these studies, cross-correlations between variables, such as cross-correlations between sea surface temperature (SST) and catch-per-unit effort (CPUE), have been used to explore the relationships between fish populations and the ocean environment.

It is clear that at this "higher" level, the variables being used no longer directly affect the fish. Rather, they act as surrogates for other variables, or for processes in the ocean, that do directly act on the fish. This is a subtle and a seemingly insignificant distinction, but it is one that will be important for the statistical discussion that follows. For, if the environmental variables being used are not causal, then it makes sense to view the modeling process, particularly in terms of forecasts, as deciding if a variable to be included in the model contains any additional, unique information about the future of the fishery dynamics over and above the information contained in the series itself or in other series to be included in the model.

The cross-correlation between two series at a given lag is not independent of the cross-correlations at other lags nor of the auto-correlations of either series. The most extreme example of this is if each series has an independent deterministic sinusoidal component. Then there is a sinusoid in the cross-correlation function. As a less extreme example, two independent series, each with a large lag-one autocorrelation, over a short period of observation will often have a significant sample cross-correlation at lag-one. This will particularly be true if the underlying series have been time aggregated (Granger 1980).

In the examples above, pre-filtering would have removed some of this spurious correlation. The effect of pre-filtering is to remove some of the information that a series carries about itself. When this removes the cross-correlations, it is tantamount to saying that there is no additional information in the new series. Pre-filtering has its own dangers, but similar characteristics can be searched for in a true multivariate setting. There has been resistance in fisheries to such an approach, because "regressing" CPUE, for example, against lagged values of itself, does not have a clear, causal explanation. If we drop the idea that our models are causal, then it is clear why some such approach makes sense: if lagged values of CPUE can predict future values just about as well as a new variable, then the past history of the CPUE series has integrated into itself any information the new series might have. The past history is not necessary causal, but in terms of information content, it is sufficient.

As I will show, it is possible to resolve some of these issues if we assume that all relationships between CPUE and its own history or between CPUE and the environment is linear. Unfortunately, dating back to some of the earliest Japanese work on fish and the ocean, and to the earlier work on tuna (see Sund et al. 1981 and refences therein for example) the evidence is that the environment affects fish in a nonlinear fashion. The tuna work, based on short term conditions around concentrations of catch, suggests a "window-like" relationship between the environment and tuna CPUE. More recently, using long time-series of data and empirically derived techniques, Mendelssohn and Mendo (1987) have shown the nonlinearities between the environment and recruitment for anchoveta off Peru. Mendelssohn and Cury (1987, 1989) using similar techniques, have shown the nonlinearities in the relationships between CPUE for small pelagics off the Ivory Coast and environmental variables. Cury and Roy (1989) have shown "window-like" relationships between recruitment for eastern boundary current pelagics and transport or turbulence in the particular region.

Such nonlinearites affect the analysis greatly. An environmental variable may not show unique information because it is being examined on the wrong scale. For example, on a linear scale, a window-like relationship will show lack of fit. While I suggest some methods for circumventing this problem, they are ad hoc solutions at best. Using fisheries from the Gulf of Guinea, I will examine ways to determine which variables will be important to include in models, and I will show in specific examples how the analysis changes due to finding a more appropriate scale for the variables.

If the environment is important in fisheries dynamics, then it is to be expected that some of the properties of the ocean's dynamics may be inherited by the fishery dynamics. Recent years has seen a boom in the literature on "chaos" and the related ideas of fractals and fractal dimensions. Probably a more accurate nomenclature is "sensitivity to initial conditions." There is, however, a feature of these concepts that is more relevant to fisheries science. The analyses in this paper assume stationarity (or more exactly joint stationarity) in the series to be studied. Or else it is assumed that the nonstationarity is of a form that can be easily removed, such as a deterministic trend. The physical motivation for these assumptions is that the effect of anything that happens now dies off relatively rapidly. Hurst (1951) noticed that riverflows appeared to have the opposite characteristic - there appeared to be long-term memory in the system. This "Hurst effect", and achieving stationarity by fractional differencing, are closely related to the concepts of chaos and fractals (see for example Mandelbrodt 1983). If the ocean's dynamics exhibit long-term memory, this can have significant implications for how far in the future we can ever expect to accurately forecast fish catches. At the end of this paper I briefly examine the question of long-term memory in the environmental series from the Gulf of Guinea.

Methods

In this section I briefly review some of the relevant literature on "causality" in time-series, and on empirical methods for determining nonlinear transformations of variables in the form of "optimal" additive models. I also discuss in more detail the specific computational techniques used in the next section.

Information between two or more time series

Pierce (1979) developed a time-series analogue to the R^2 statistic of regression analysis for the case of bivariate time series with no feedback. Pierce's statistic measures how much better a second series helps predict a given series over that available from the past history of the series. Parzen (see discussion section of Geweke 1982) pointed out that this amounted to studying the "innovations of the innovation series", a comment which will be made more concrete below. Geweke (1982, 1984) extended these concepts to the true multivariate case with feedback. Other related papers are Gelfand and Yaglom (1959), Caines and Chan (1975), Gustafson, Ljung and Soderstrom (1977) and Anderson and Gevers (1982).

In this paper I follow Gersch (1986), and the presentation closely follows the discussion of that paper. The background for developing a quantitative measure of the information in one time series about another time series (Gelfand and Yaglom 1959) is the measure of the amount of information between two continuously distributed vector random variables X and Y (Shannon 1948, Woodward 1953),

$$I_{x,y} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\ln[f_{X,Y}(x,y)/f_X(x)f_Y(y)]\,dx\,dy \tag{1}$$

where $f_{X,Y}(x,y)$ is the joint probability distribution function of X and Y, and $f_X(x), f_Y(y)$ are the marginal probability distributions of X and Y. Equation (1) is the negative entropy of the "true" distribution $f_{X,Y}$ with respect to the assumed distribution $f_X f_Y$. More formally, equation (1) is the amount of information, on the average, per observation, to reject the null hypothesis that the random vectors X, Y are independent.

In order to use equation (1) simplifying assumptions are necessary. Gelfand and Yaglom (1959) assume that the time series are stationary and jointly normally distributed. Let $w(t)$ be an $(r+q)$ vector-valued time series that consists of an r-vectored time series $x(t)$ and a q-vectored time series $y(t)$, that is $w(t) = (x(t), y(t))^T$, and the notation z^T denotes the transpose of the vector z. Then the spectral density matrix of $w(t)$ at a frequency λ can be partitioned as

$$S_{ww}(\lambda) = \begin{bmatrix} S_{xx}(\lambda) & S_{xy}(\lambda) \\ S_{yx}(\lambda) & S_{yy}(\lambda) \end{bmatrix}$$

Then the Shannon-Gelfand-Yaglom (SGY) measure of information is

$$I_{x,y} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \frac{|S_{ww}(\lambda)|}{|S_{xx}(\lambda)||S_{yy}(\lambda)|}\,d\lambda \tag{2}$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \frac{|S_{xx}(\lambda)|}{|S_{xx}(\lambda) - S_{xy}(\lambda)S_{yy}^{-1}(\lambda)S_{xy}^{*}(\lambda)|} \, d\lambda \tag{3}$$

$$= -\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln[1 - W_{xy}^2(\lambda)] \, d\lambda \tag{4}$$

where S^* denotes the conjugate transpose of the matrix S, and W^2 is the multiple coherence at the frequency λ. In equation (3), if X and Y are independent, then the denominator and numerator are identical and the integral is zero, as would be expected. Equation (4) expresses the information in terms of the extension of the coherence between two series, the multiple coherence.

For arbitrary time series, the SGY measure partitions into

$$I_{x,y} = I_{x \to y} + I_{y \to x} + I_{x.y} \tag{5}$$

where $I_{x \to y}$ is the amount of feedback between the series $x(t)$ and $y(t)$, and $I_{x.y}$ is the amount of instantaneous feedback between the two series (Geweke 1982). The measures in equation (5) can be expressed in terms of linear regressions. Let $|W$ denote conditioning on the present and past of a series, $|W^-$ denote conditioning on the past of a series. Then the series $w(t)$ can be written in equivalent moving average (MA) and autoregressive (AR) forms

$$w(t) = \sum_{i=0}^{\infty} A(i)\epsilon_{t-i}; \qquad Var\{\epsilon_t\} = \Sigma(W|W^-) \tag{6}$$

$$w(t) = \sum_{i=1}^{\infty} B(i)w(t-i) + \epsilon_t. \tag{7}$$

The component series $x(t)$ and $y(t)$ have AR representations

$$x(t) = \sum_{i=1}^{\infty} E_{1i}x(t-1) + u_{1t}; \qquad Var\{u_{1t}\} = \Sigma(X|X^-) \tag{8}$$

$$y(t) = \sum_{i=1}^{\infty} G_{1i}x(t-1) + v_{1t}; \qquad Var\{v_{1t}\} = \Sigma(Y|Y^-) \tag{9}$$

The component series can also be expressed in terms of two different ARMAX models that depend on different present and past information

$$x(t) = \sum_{i=1}^{\infty} E_{2i}x(t-i) + \sum_{i=1}^{\infty} F_{2i}y(t-i) + u_{2t}; \qquad Var\{u_{2t}\} = \Sigma(X|X^-, Y^-) \tag{10}$$

$$y(t) = \sum_{i=1}^{\infty} H_{2i}x(t-i) + \sum_{i=1}^{\infty} G_{2i}y(t-i) + v_{2t}; \qquad Var\{v_{2t}\} = \Sigma(Y|X^-, Y^-) \tag{11}$$

$$x(t) = \sum_{i=1}^{\infty} E_{3i}x(t-i) + \sum_{i=0}^{\infty} F_{3i}y(t-i) + u_{3t}; \qquad Var\{u_{3t}\} = \Sigma(X|X^-, Y) \tag{12}$$

$$y(t) = \sum_{i=0}^{\infty} H_{3i}x(t-i) + \sum_{i=1}^{\infty} G_{3i}y(t-i) + v_{3t}; \qquad Var\{v_{3t}\} = \Sigma(Y|X, Y^-) \tag{13}$$

Then the information measures in equation (5) can be written in terms of the innovations of these ARMAX models (Geweke 1982) as

$$I_{x,y} = \ln \frac{|\Sigma(X|X^-)||\Sigma(Y|Y^-)|}{|\Sigma(W|W^-)|} \tag{14}$$

$$I_{x \to y} = \ln \frac{|\Sigma(Y|Y^-)|}{|\Sigma(Y|X^-, Y^-)|} \tag{15}$$

$$I_{y \to x} = \ln \frac{|\Sigma(X|X^-)|}{|\Sigma(X|X^-, Y^-)|} \tag{16}$$

$$I_{x.y} = \ln \frac{|\Sigma(X|X^-, Y^-)|}{|\Sigma(W|W^-)|} \tag{17}$$

This is the innovations of the innovation series mentioned above.

The amount of information and feedback can be computed in either the time domain or the frequency domain. An important additive decomposition is (Geweke 1982)

$$I_{y\to x} \geq \int_{-\frac{1}{2}}^{\frac{1}{2}} I_{y\to x}(\lambda)\, d\lambda \tag{18}$$

with similar relationships for the other measures. (The frequency domain representation is useful for examining environment influences on CPUE, as estimates at different frequencies asymptotically are independent. In the time-domain, it is necessary to adjust for the past history of CPUE. In the frequency domain, only the relationships between CPUE and the environment need be examined).

Model the time series $x(t)$ in two ways - regress it on its own past and on its own past and the past of the series $y(t)$. Let $S_x(\lambda)$ be the power spectrum of $x(t)$ from the regression on its own past, and let $S_{x|x,y}(\lambda)$ be the contribution to the power spectrum of x from its own past when it has been regressed on both x and y. Then

$$I_{y\to x}(\lambda) = \ln \frac{|S_x(\lambda)|}{|S_{x|x,y}(\lambda)|} \tag{19}$$

If, for example, $y(t)$ does not influence or feedback on $x(t)$ then the inner term in equation (19) is one, and the measure is zero as expected. Geweke (1984) shows that these results continue to hold under conditioning or partial regression on another time series $z(t)$.

To practically implement this, suppose we observe three univariate time series $x(t), y(t)$ and $z(t)$. Assume that there is pairwise linear dependence between the series, that is

$$I_{x,y} \neq 0, \qquad I_{x,z} \neq 0, \qquad I_{y,z} \neq 0. \tag{20}$$

Assume also that each pair of time series is conditioned on the excluded third time series, with the following results,

$$I_{x,y|z} \neq 0, \qquad I_{x,z|y} \neq 0, \qquad I_{y,z|x} = 0. \tag{21}$$

This implies that the the $x(t)$ series uniquely explains the linear relationships between the three series.

Suppose spectral density matrices have been estimated at each frequency and are partitioned as above. Then the coherence between x and y is,

$$W^2_{xy}(\lambda) = \frac{|S_{xy}(\lambda)|}{S_{xx}(\lambda)S_{yy}(\lambda)} \tag{22}$$

The partial spectral densities (conditioned on the excluded variables) are (Brillinger 1981)

$$S_{xx|z}(\lambda) = S_{xx}(\lambda)(1 - W^2_{xz}(\lambda)) \tag{23}$$

$$S_{yy|z}(\lambda) = S_{yy}(\lambda)(1 - W^2_{yz}(\lambda)) \tag{24}$$

$$S_{xy|z}(\lambda) = S_{xy}(\lambda) - \frac{S_{xz}(\lambda)S_{zy}(\lambda)}{S_{zz}(\lambda)} \tag{25}$$

Then the partial coherence between x and y after removing the effects of z is

$$W^2_{xy|z}(\lambda) = \frac{|S_{xy|z}(\lambda)|^2}{S_{xx|z}(\lambda)} \tag{26}$$

Brillinger (1981) gives multivariate extensions for each of these terms. The partial coherences are the frequency domain analogues of partial correlations in regression theory.

The conditions in equations (20, 21) for x to be "causal" can be estimated by

$$W^2_{xy}(\lambda) \neq 0, \quad W^2_{xz}(\lambda) \neq 0, \quad W^2_{yz}(\lambda) \neq 0 \tag{27}$$

$$W^2_{xy|z}(\lambda) \neq 0, \quad W^2_{xz|y}(\lambda) \neq 0, \quad W^2_{yz|x}(\lambda) = 0 \tag{28}$$

Gersch (1986) and Brillinger (1981) give significance tests for each of these quantities. There are a variety of methods for estimating spectral density matrices. In this paper, I have used the autoregressive estimates of Akaike (1980). There is a close connection between AR spectral estimates and Maximum Entropy Spectral Estimation (Ulrych and Bishop 1975).

The AR spectral estimate is calculated by fitting multivariate autoregressive models to the data up to a maximum lag. Akaike's AIC criterion is used to pick the "best" model. Assume this "best" model has m lags, and that the estimated coefficient matrices are given by A_i at lag i. Also let the innovations covariance matrix be denoted by U. Then the spectral density matrix at frequency λ is given by,

$$S(\lambda) = A(\lambda)^{-1} U A^*(\lambda)^{-1} \tag{29}$$

and $A(\lambda) = I - \sum_j^m A_j \exp(-2\iota\lambda j)$ is the Fourier transform of the coefficient matrices.

Estimating nonlinear transformations

The standard linear regression model estimates a function of the form

$$y(i) = a + \sum_j b(j)x_j(i). \tag{30}$$

If it is believed that the relationship between the y's and the x's is nonlinear but of an unknown form, then an acceptable first approximation might be the additive model

$$T(y(i)) = \tilde{a} + \sum_j \tilde{b(j)} s_j(x_j(i)) \tag{31}$$

where $T(\cdot)$ and the $s_j(\cdot)$ are unknown functions. Breiman and Friedman (1985) have derived an algorithm, the "Alternating Conditional Expectations" algorithm, or ACE, to estimate models of the form of equation (31).

Briefly, for one independent variable, they show this problem is equivalent to finding the transformations that have the highest possible correlation. The quantities that need to be solved for can be written in terms of conditional expectations. The algorithm proceeds by alternately estimating the conditional expectation after removing the influence of the other variables. The conditional expectations are estimated using scatterplot smoothers. The result, therefore, is not a function that can be written down, but rather an empirical transformation of each observed point. The transformation can be observed by plotting the transformed values against the original values. There is a close connection between the ACE algorithm and the power method of calculating the maximum eigenvalue of a matrix (see discussion section of the article).

The ACE algorithm has been used in a fisheries context by Mendelssohn and Cury (1987, 1989), by Mendelssohn and Mendo (1987), Mendelssohn and Husby (1992), and Cury and Roy (1989). What is noted in each of these papers is that the overall form of the transformations generally have close, obvious physical interpretations. Transformations of CPUE, as would be expected from the ratio of two seperate series, tends to be close to a log scale.

Fractional differencing and long-term memory

My discussion follows Porter-Hudak (1982). Hurst (1951) examined the flows of rivers. Let $x_1, x_2, \ldots, x_T$ be the historical sequence of flows. Then the cumulative flows up to time t are

$$Su_t = \sum_{j=1}^{t} x_j, \qquad t = 1, 2, \ldots, T. \tag{32}$$

Related to reservoir construction is the sequential range

$$R = \max_t Su_t - \min_t Su_t. \tag{33}$$

Hurst looked at the normalized sequential range (or "rescaled adjusted range") where Su_t has the mean flow removed and the quantity is divided by the standard deviation of the flows. If the flows are independent over large time scales then one would expect that $R/S \sim (T/2)^{\frac{1}{2}}$. However, Hurst found that for a wide range of river flows $R/S \sim (T/2)^H$ where H is in the range (.6, .8). Mandelbrot (Mandelbrot and Van Ness 1968, Mandelbrot 1971) explained this in terms of fractional Gaussian noise (fGn). Let $B(s)$ be Brownian motion, a stochastic process such that $B(s+u) - B(s)$ are $N(0,1)$ and independent. Then fractional Brownian motion takes the form

$$B_H(t) = \int_{-\infty}^{t} (t-a)^{H-\frac{1}{2}}\, dB(s) \qquad -\infty < t < \infty \tag{34}$$

or

$$B_H(t) - B_H(0) = \int_{-\infty}^{t} (t-s)^{H-\frac{1}{2}}\, dB(s) - \int_{-\infty}^{0} (-s)^{H-\frac{1}{2}}\, dB(s). \tag{35}$$

Discretizing the equation yields

$$b_t(H) = B_H(t) - B_H(t-1). \tag{36}$$

Fractional Gaussian noise can be described by its autocorrelation function

$$C(s) = (\frac{1}{2})(|s-1|^{2H} - 2|s|^{2H} + |s+1|^{2H}). \tag{37}$$

Fractional Gaussian noise also exhibits the self-similarity property, that is

$$X(t\lambda) \sim (\lambda)^a X(t). \tag{38}$$

A related concept is fractional differencing (Granger and Joyeux 1980, Hosking 1981)

$$(1-B)^d X_t = \epsilon_t \tag{39}$$

where d is possibly nonintegral, B is the backshift operator, and ϵ_t are independently distributed as $N(0, \sigma^2)$. Note that this is not a nonintegral lag operator, but rather an infinitely lagged polynomial in B whose weights die out at a rate given in the autoregressive representation equation (41). The spectral density of this model is

$$f(\lambda) = \frac{\sigma_\epsilon^2}{2\pi}(2(1 - cos(\lambda))^{-d} \qquad \lambda \neq 0. \tag{40}$$

Granger and Joyeux (1980) derive the following autoregressive representation of a fractionally differenced process

$$\sum a_j X_{t-j} = \epsilon_t; \qquad a_j = \frac{\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)} \qquad j > 1. \tag{41}$$

Porter-Hudak (1982) and Geweke and Porter-Hudak (1983) show that the power spectrums of the error term from an fGn process and a fractionally differenced process only differ by a short-memory component. Thus both models are estimating essentially the same long-term memory component. Porter-Hudak (1982) and Kashyap and Eom (1988) give methods for estimating the fractional differencing parameter d, based on regressions between the log of the periodogram of the observed series and the log of the theoretical spectrum of a fractionally differenced series. The most important property is that long-term memory models have a value of d in the range $(0, \frac{1}{2})$; processes such that $d \geq 1/2$ have infinite variance, and those with $d \leq 0$ are short-term memory models (Granger and Joyeux 1980, Hosking 1981).

Yellowfin Tuna in the Gulf of Guinea

In this section and the next one I make real some of the ideas in the last section, using yellowfin tuna catches in the Gulf of Guinea and catches of small pelagics off the Ivory Coast. Mendelssohn and Roy (1986) studied the space-time dynamics of yellowfin CPUE in the Gulf of Guinea in relation to SST and wind. They divided the northern part of the Gulf of Guinea into 11 areas (fig. 1).

Within each area a local model was estimated that also estimated the missing data (in this instance, whenever there was no fishing in the area there was missing data), using an extension of the EM algorithm (Dempster, Laird and Rubin 1977) suggested by Shumway and Stoffer (1982). Frequency domain principal components and canonical correlations were used to examine phase relationships between the different areas at given frequencies. Mendelssohn and Roy (1986) suggest that yellowfin CPUE follows the propogation of SST in what appears to be a remotely-forced, Kelvin wave type behavior.

The local models suggested that the onset of upwelling previous to fishing was an important key to fishing success. The combined local and global results suggested that SST and wind were surrogate variables for an oceanic process that probably tended to concentrate food for the predatory tunas. There was evidence that the actual relationships were nonlinear, but this was not explored in the paper. The question of which variables in each area were most important in the local models was addressed in at best an ad hoc manner. It is these last two questions that I explore here.

I restrict myself to the coastal areas, areas 7 through 11, where the bulk of the fishing effort has been. It is worth emphasizing that the spectral results given here are average results - at any given time period a variable may be having much more influence on the system than represented in the average statistics.

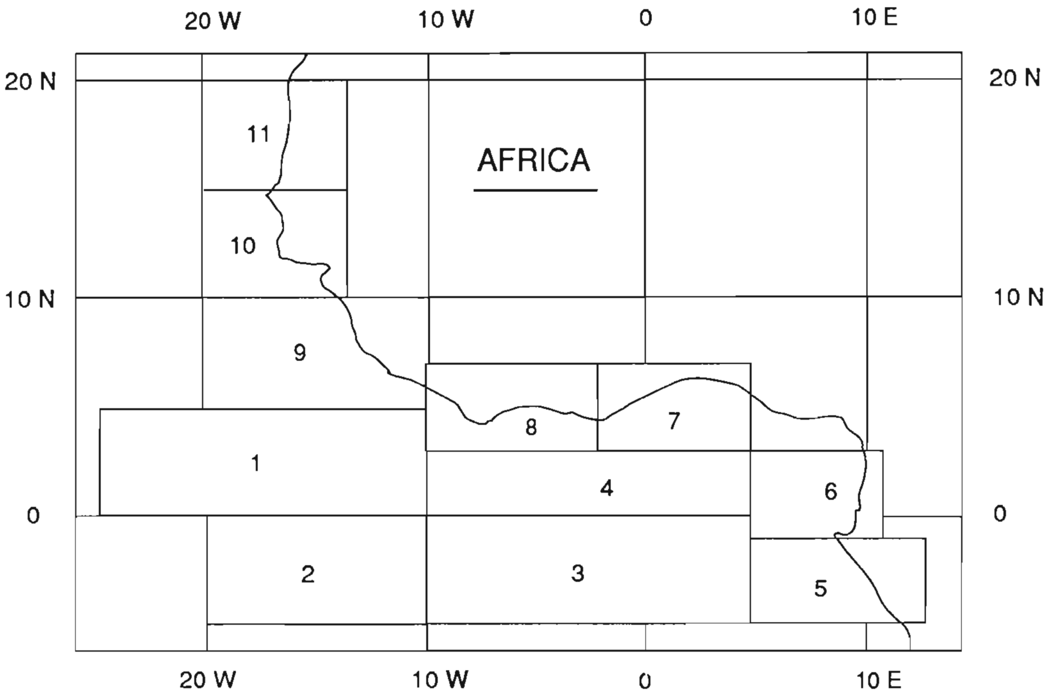

Figure 1: Areas of study for Yellowfin Tuna CPUE.

Area 7

The optimal transformations were estimated arbitrarily up to a lag of two fortnights. The estimated transformation for any variable can be different at each lag, so that ideally each variable at each lag could be treated as a separate variable in the analysis. However, I have arbitrarily used LNCPUE = ln(CPUE + 1.) as the transformation of CPUE, and the estimated transformation at a lag of 1 fortnight for all the environmental variables.

A simple, convenient method for determining the effect of the transformations on a variable is to compare the power spectra of the original and transformed series. This is particularly useful in this situation as it relates nicely to the suggested measures of information in section 2.1 (fig. 2). The LNCPUE series has a smoother spectrum than does CPUE with the peak moved from roughly a period of 7 weeks to a period of 5 weeks.

The estimated transformation for SST (fig. 3) is nearly linear at a lag of one fortnight, but shows a window-like structure with a peak around $27^{\circ} - 28^{\circ}$ C. Despite the near linearity of the transformation at a lag of 1 fortnight, it has a pronounced effect on the spectrum (fig. 4). The peak in the spectrum has been shifted from a period of roughly two months to a period of almost 5 weeks, much closer to the period of LNCPUE.

The estimated tranformation of the north-south wind component is highly nonlinear (not shown), but the effect of the transformation on the spectrum is almost identical to that of SST. The spectrum of the transformed variable is smoother with a peak identical to that of SST. Both the SST series and the north-south wind series have low coherences with CPUE at all frequencies (fig.5-6) while the transformed variables have significant and significantly higher coherences. However, while the coherences and partial coherences for the north-south wind are essentially identical, the partial coherences for SST are greatly lower then the coherences. Carrying out the full analysis in the last section suggests that the significant environmental variable in area 7 is the north-south wind, and that the coherence with SST is due to coherence and feedback between SST and the wind.

Area 8

As in area 7, the spectrum of LNCPUE is shifted over to the higher frequencies with a period of about 5 weeks, but in this case the peak is also much narrower (fig. 7). The estimated transformation for the east-west wind (not shown) is hyperbolic in shape, and its affect on the power spectrum is almost identical to that of CPUE. The estimated

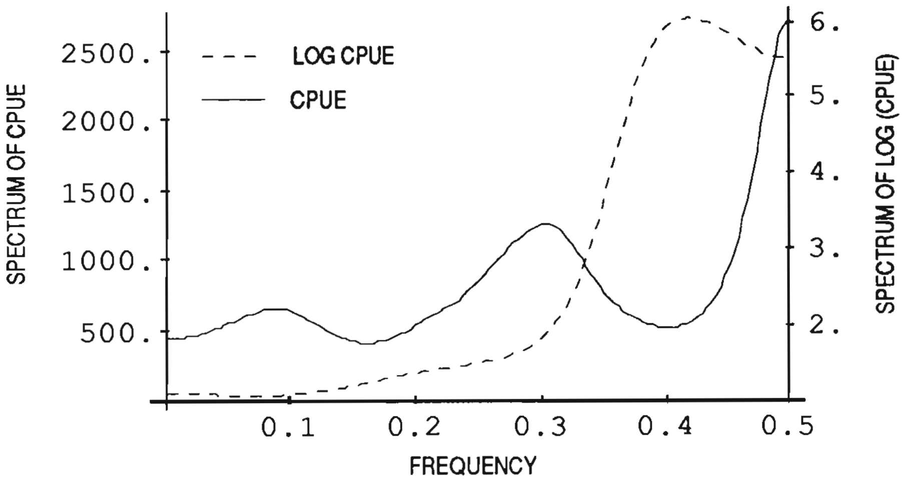

Figure 2: Power Spectrum of CPUE and LNCPUE in area 7.

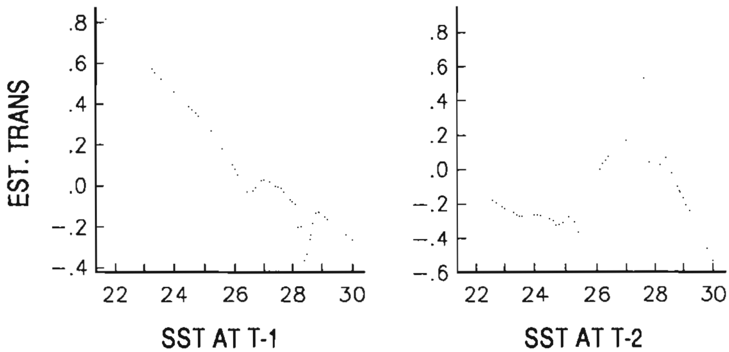

Figure 3: Estimated transformation of SST in area 7 from the ACE algorithm at lags of one and two fortnights.

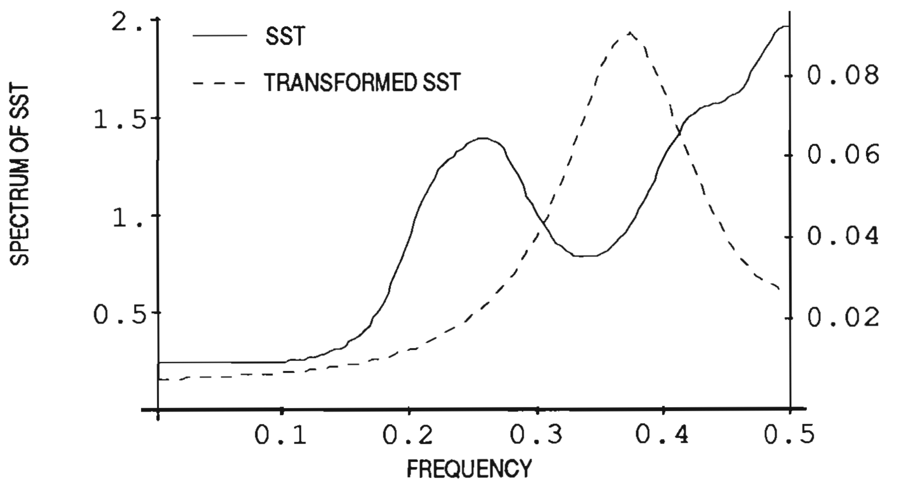

Figure 4: Estimated power spectrum of raw and transformed SST in area 7.

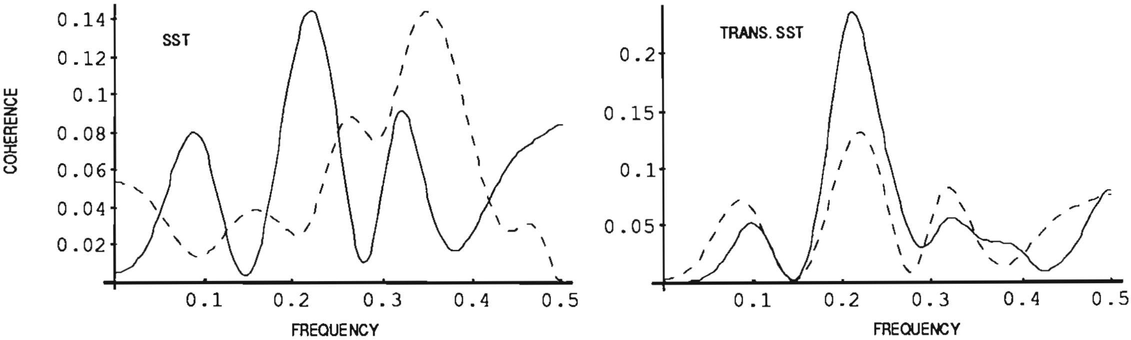

Figure 5: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with SST and LNCPUE with transformed SST in area 7.

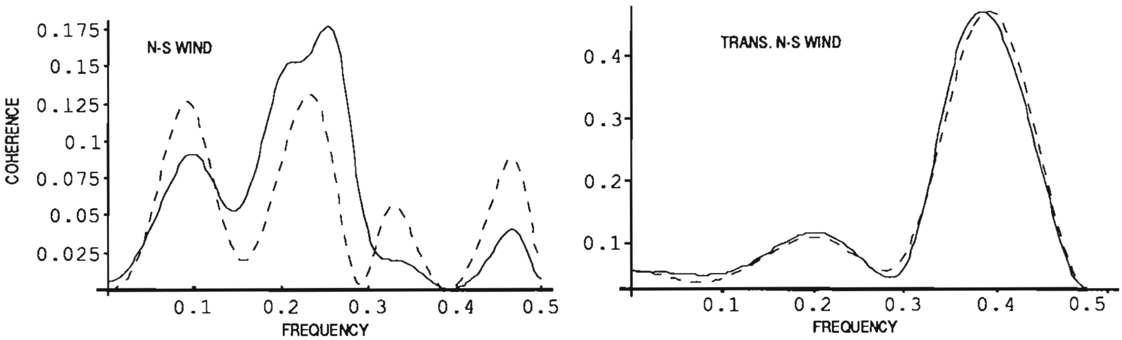

Figure 6: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with north-south wind and LNCPUE with tranformed north-south wind in area 7.

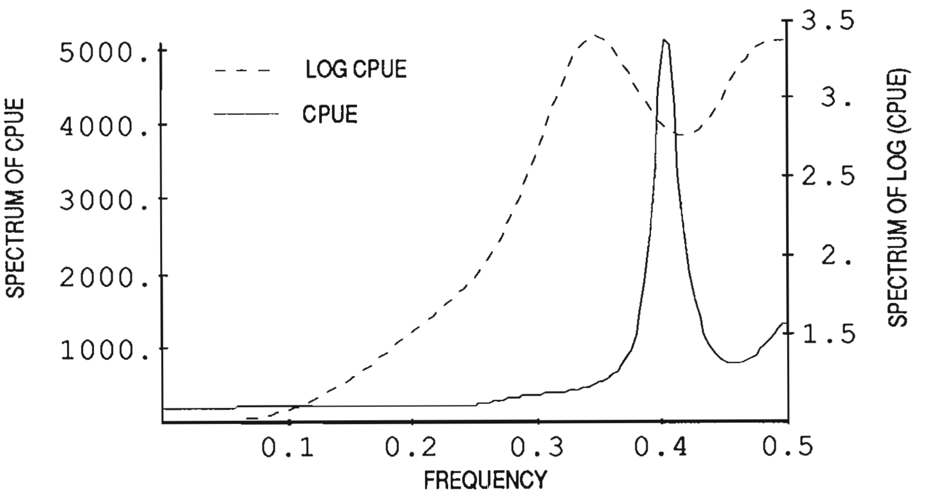

Figure 7: Power spectrum of CPUE and LNCPUE in area 8.

transformation of the north-south wind is again highly non-linear (not shown), and the power spectrum is transformed much as with CPUE and the east-west wind.

The transformation of SST at a lag of 1 fortnight is almost window-shaped (fig. 8) The transformation smooths the power spectrum shifting the power into the highest frequencies at a period of around 1 month. While the coherences of both raw and untransformed east-west wind are high, the partial coherences of both at all frequencies are non-significant. (This is true for areas 9 through 11 also, so this variable will not be examined further). The coherence between CPUE and SST, particularly with the raw series, is particularly significant at a period of roughly 5 weeks (fig. 9). The partial coherences between CPUE and SST are essentially insignificant. The partial coherences between LNCPUE and SST are insignificant except at a period of 1 month.

The coherences and partial coherences between CPUE and north-south wind are almost identical (fig.10), while with the transformed data, one peak at a period of 1.5 months is sharply reduced. However, the overall results are the same as with area 7 - north-south wind is the predominant environmental variable, working at relatively high frequencies, and with an additional effect due to SST on the transformed scale at a period of one month. The transformed data has in effect separated out the frequencies at which the wind and SST contain important information about CPUE. However, at most frequencies, the observed coherence between SST and CPUE appears to be due to feedback and coherence between SST and the wind.

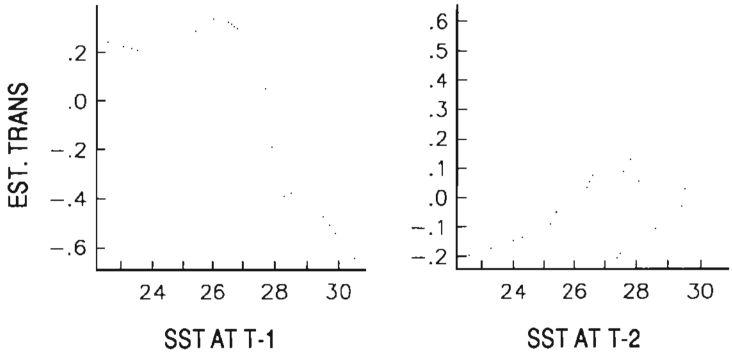

Figure 8: Estimated optimal transformation of SST at a lag of one and two fortnights in area 8.

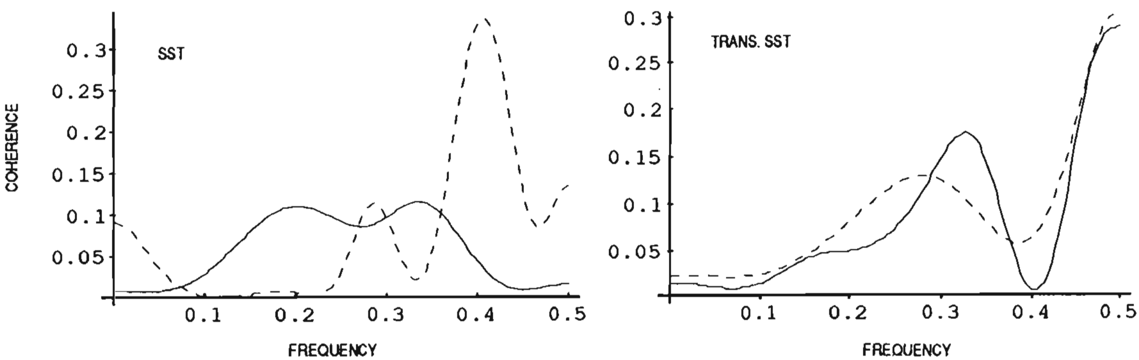

Figure 9: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with SST and LNCPUE with transformed SST in area 8.

Area 9

The spectra for CPUE and LNCPUE in this area are nearly identical with a sharp peak at a period of slightly greater than one month (not shown). The transformation of the north-south wind (fig. 11) is window-shaped, and transforms the power spectrum to a single peak at the same frequency as that of CPUE (fig. 12).

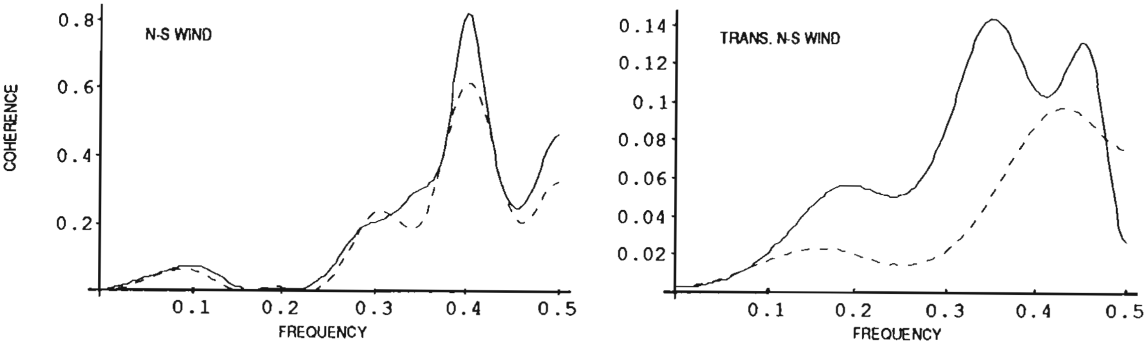

Figure 10: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with north-south wind and LNCPUE with transformed north-south in area 8.

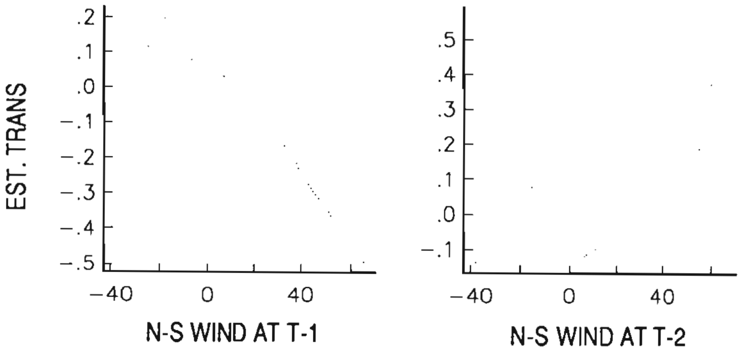

Figure 11: Estimated transformation for north-south wind in area 9.

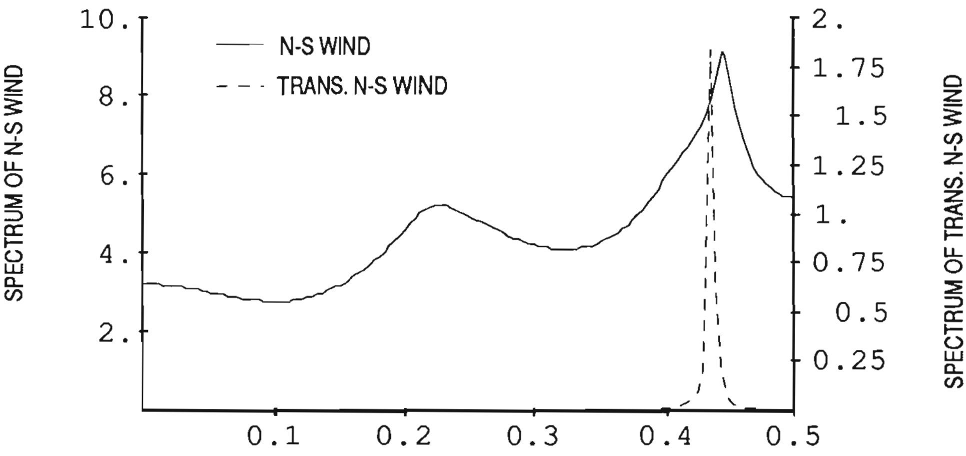

Figure 12: Estimated spectra of raw and transformed north-south wind in area 9.

The coherences and partial coherences between CPUE and north-south wind are similar (fig. 13), but with the transformed data a very significant peak in the coherence disappears in the partial coherence. A similar result can be seen for SST (fig. 14). Here the high-frequency peak in the coherence disappears, while a lower frequency peak remains. This is evidence that at the peak frequency, SST and the north-south wind contain nearly identical information about the environment.

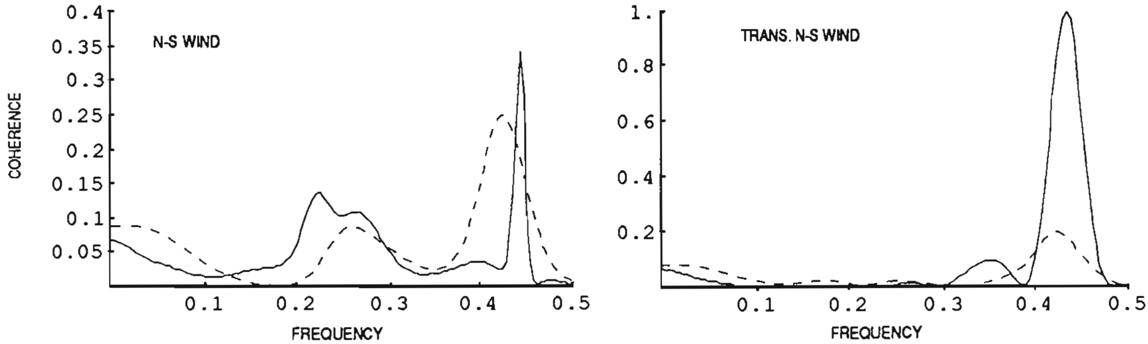

Figure 13: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with north-south wind and LNCPUE with transformed north-south wind in area 9.

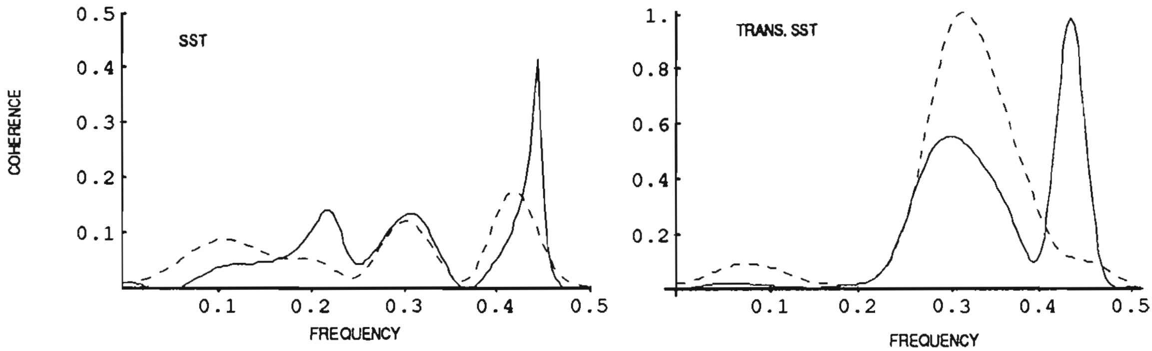

Figure 14: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with SST and LNCPUE with tranformed SST in area 9.

Areas 10 and 11

In area 10, the estimated transformations for both the north-south wind and SST are nearly linear (not shown). For both variables, however, the effect on the estimated power spectrum is to shift a single peaked spectrum from a frequency of .45 to a frequency of .4 (not shown). For the raw data, the peaks in the coherences with either environmental variable disappear in the partial coherences (fig. 15 and 16). As these peaks occur at the same frequencies for both SST and the north-south wind, this implies that both variables contain similar information about the ocean in terms of CPUE for yellowfin. On the transformed scale, the coherences and partial coherences for LNCPUE with transformed SST are nearly identical, while those with the north-south wind are reduced at the peak frequency. However, at this peak frequency the partial coherence with the north-south wind is still significant, so that both wind and SST contain unique information when the data are appropriately transformed.

The situation is very similar in area 11, where SST and north-south wind both are significant at a period of roughly 2 months. Each variable contains some overlapping information at this frequency, but the partial coherences are also significant though smaller than the coherences.

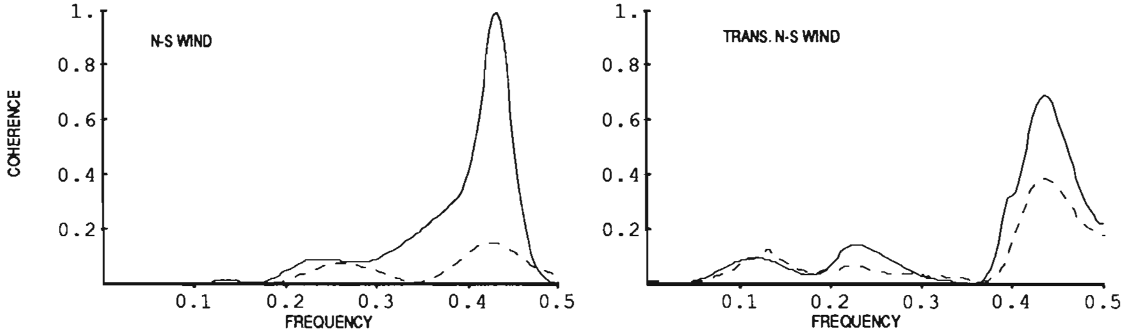

Figure 15: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with north-south wind and LNCPUE with transformed north-south wind in area 10.

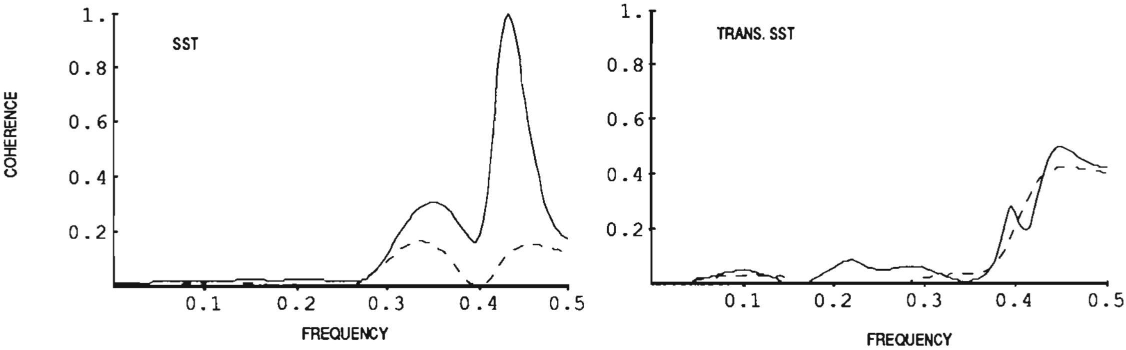

Figure 16: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with SST and LNCPUE with transformed SST in area 10.

Small pelagics off the Ivory Coast

Mendelssohn and Cury (1988) estimate forecasting models and nonlinear transformations for the explanatory variables for modeling the CPUE of small pelagics off the Ivory Coast. Mendelsson and Cury (1989) also study the dominant species in the catch, *Sardinella maderensis*, in time, as well as in space and time. Their results suggests that the factors affecting the aggregate CPUE are not the same as those affecting *S. maderensis* by itself. In this section I reexamine this question using the techniques of this paper.

I have arbitrarily estimated the optimal transformations for the explanatory variables from a model which includes lags up to 4 fortnights. Thus, each model contains lagged values of CPUE, lagged values of SST at 5 meters depth from a shore station (SST5M) and salinity at 5 meters depth (SAL5M) from the same station.

Log of CPUE again is used as the transformation for CPUE. LNCPUE shifts the spectrum from a peak at around a period of 1 month to a peak centered around 5 weeks (fig. 17). The estimated transformations for SST5M and SAL5M (fig. 18,19) shift their power spectra in a similar manner so that all three series have a peak in their spectrum at a period of roughly 5 weeks (fig. 20,21).

The coherences and partial coherences for SST5M with CPUE are nearly identical (fig. 22). The coherences make clear how the transformation has shifted the spectrum so that it is closely aligned with LNCPUE.

The coherences for CPUE with SAL5M show a strong coherence at a period of 1 month, but this disappears in the partial coherence (fig. 23). On the transformed scale, while the coherences and partial coherences are similar, they are not particularly significant. This supports the claim in Mendelssohn and Cury (1987) that for combined CPUE, SST5M was sufficient for prediction and that no extra information was contained in the SAL5M series.

When examining CPUE for *S. maderensis* only, the picture is different. The coherences and partial coherences for SST5M with CPUE are nearly identical, with a significant peak at a period of roughly 2 months (fig. 24,25). For SAL5M, the coherences and partial coherences are identical expect at this frequency, where a significant peak in the coherence is eliminated. The most significant coherence for CPUE with SAL5M is now at a period of roughly a month.

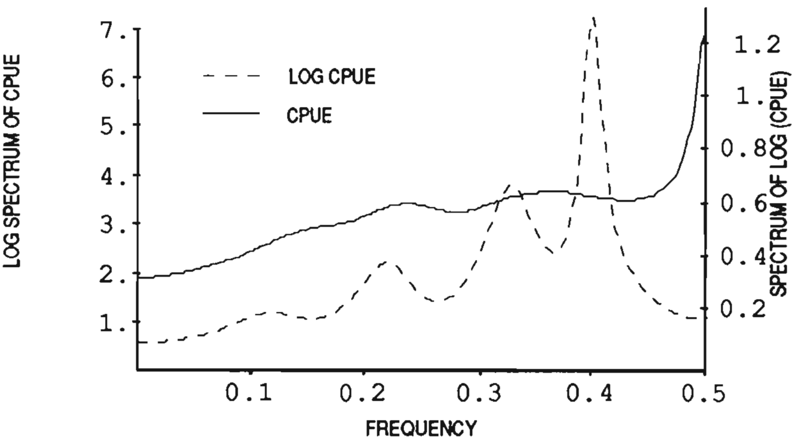

Figure 17: Estimated power spectrum for CPUE and LNCPUE for small pelagics off the Ivory Coast.

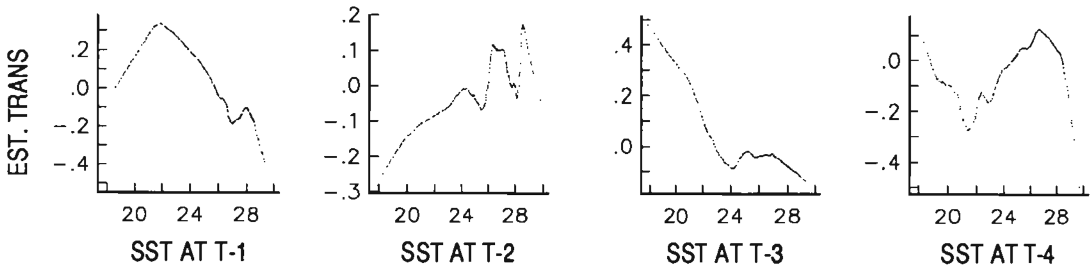

Figure 18: Estimated transformations for SST5M off the Ivory Coast.

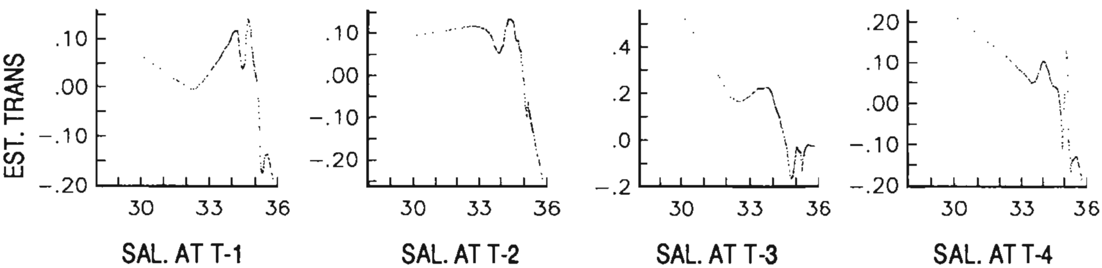

Figure 19: Estimated transformations for SAL5M off the Ivory Coast.

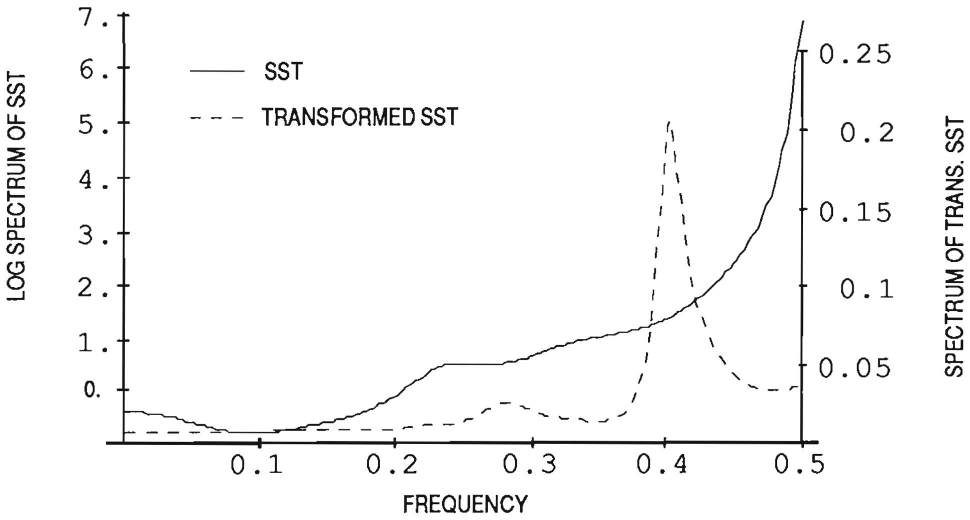

Figure 20: Estimated power spectrum for SAL5M off the Ivory Coast.

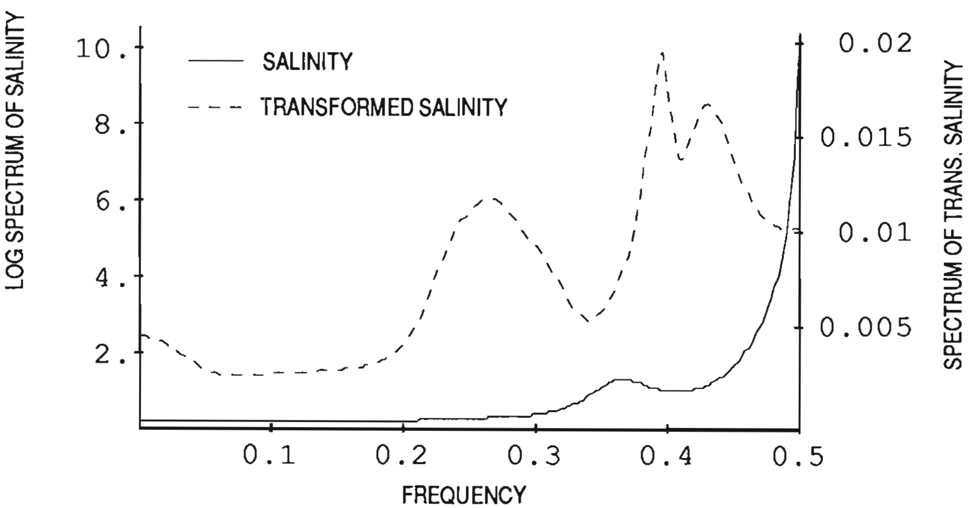

Figure 21: Estimated power spectrum for SAL5M off the Ivory Coast.

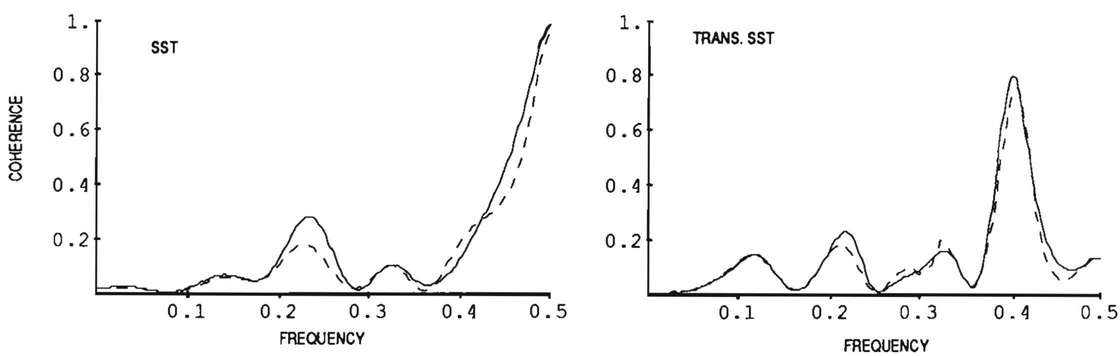

Figure 22: Coherences (solid lines) and partial coherences (dashed lines) for CPUE with SST5M and LNCPUE with transformed SST5M for small pelagics.

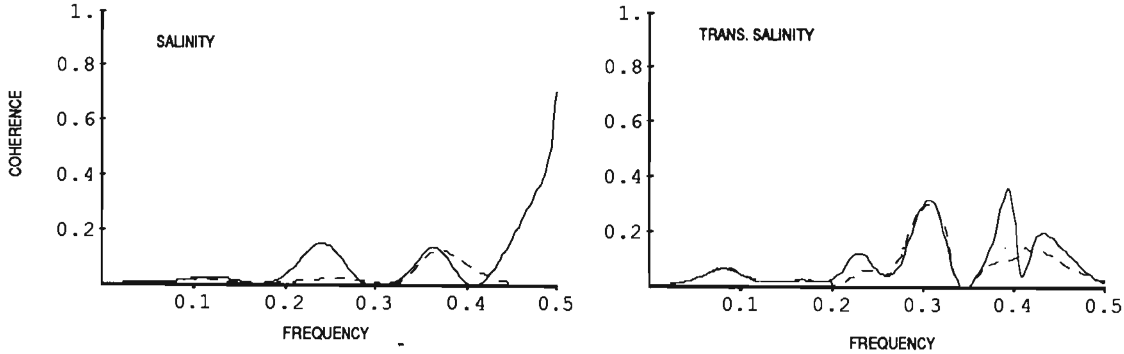

Figure 23: Coherences (solid lines) and partial coherences (dashed line) for CPUE with SAL5M and LNCPUE with transformed SAL5M for samll pelagics.

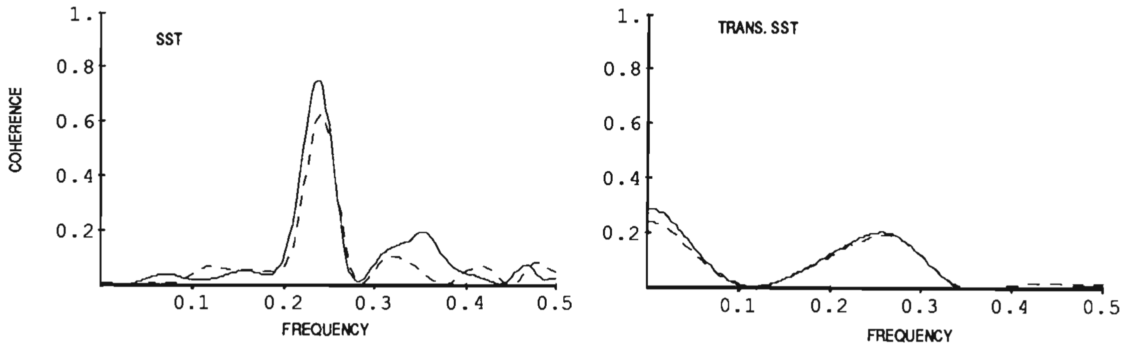

Figure 24: Coherences (solid lines) and partial coherences (dashed line) for CPUE with SST5M and LNCPUE with transformed SST5M for *S. maderensis*.

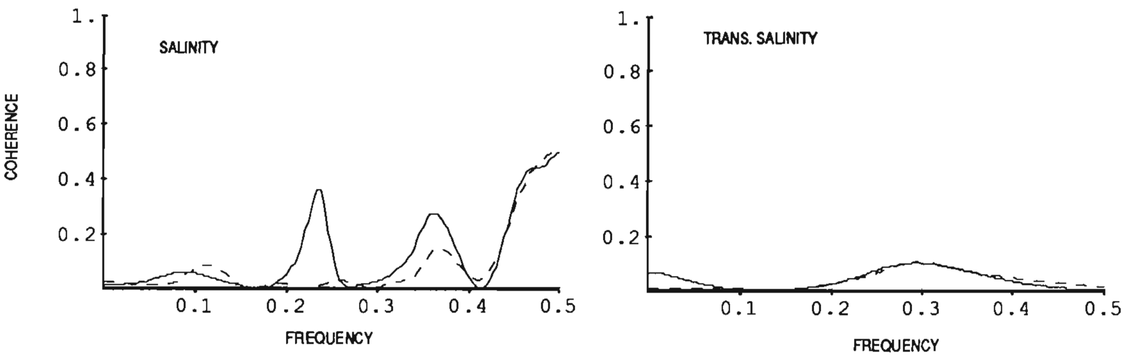

Figure 25: Coherences (solid lines) and partial coherences (dashed lines) between raw and transformed SAL5M and raw and transformed CPUE for *S. maderensis*.

Thus, for *S. maderenis*, the partial coherences show SST5M and SAL5M both containing significant information about CPUE, but at different frequencies. That SAL5M is important for *S. maderensis* but not for combined CPUE was asserted by Mendelssohn and Cury (1989) using time-domain techniques. Here I show further evidence of the claim.

The low coherences in the transformed data for *S. maderensis* is due to the fact that I have arbitrarily used the transformation at a lag of 1 fortnight. Particularly for SAL5M, it is clear from the analysis of Mendelssohn and Cury (1989) that salinity has its greatest effects at much longer lags. There are significant and difficult management questions raised by the fact that the dynamics of the combined CPUE (that faced by the fishermen) are much different from that of the dominant species in the catch. Clearly, a forecast of the aggregate could become misleading, and it may be preferable to aggregate forecasts instead.

Long-term memory in the ocean environment in the Gulf of Guinea

In this section I briefly consider if there is long-term memory in the ocean environment in the Gulf of Guinea. The presence of long-term memory may have profound implications for fisheries management. If long-term memory exists in the ocean, then small changes in the ocean, rather than dying out and causing bounded changes, persist and their effects may multiply through time. Large changes, such as El Nino type events, can have lasting consequences. If the ocean exerts a significant influence on the dynamics of a fish population, then it can be expected that fish populations will display a similar property. This would limit how far into the future we can reasonable expect to understand the consequences of management regimes on the fishery.

I examine the question of the presence of long-term memory by calculating estimates for the fractional differencing parameter d in equations (38) and (39) for several of the environmental series. I use the shore station SST and salinity series (SST5M and SAL5M) as probably reflecting the environmental conditions relevant to the important near-shore fisheries. I also use SST and the two wind component series from area 9, which is the area of the Intertropical Convergence Zone. This area most likely will reflect much of the dynamics of the Gulf. The salient feature here is that short-term memory series have a value between $(-1/2, 0)$, long-term memory models have a value of d between $(0, 1/2)$, and series with a value of d greater than $1/2$ have infinite variance.

The estimates of d (table 1) suggest that all of the series except the east-west component of the wind have a strong long-term memory component.

	Method of Estimation	
Series	Geweke & Porter-Hudak	Kashyap & Eom
SST5M	0.4170	0.4426
SAL5M	0.4241	0.4486
Area 9 SST	0.6169	0.6728
Area 9 NS Wind	0.4036	0.4312
Area 9 EW Wind	0.0021	-0.0070

Table 1: Estimates of the fractional differencing parameter for selected environmental series in the Gulf of Guinea.

There is some evidence that SST may even be an infinite variance series. If this is so, climatologies and other mean-like statistics in this region would have very little meaning. SST in the Gulf of Guinea appears to reflect important processes in the Gulf of Guinea that have major influences on fish dynamics, and my results suggest that these processes may have infinite variance. While much more research must be done in this area before any definitive statements can be made, the initial evidence is troubling indeed.

References

Akaike, H. 1980. On the identification of state space models and their use in control. *In*: D. Brillinger and G. Tiao (eds.), Directions In Time Series. Proceedings of the IMS Special Topics Meeting on Time Series Analysis, Ames, Iowa 1978. Institute of Mathematical Statistics, Hayward.

Anderson, B.D.O. and M.R. Gevers. 1982. Identifiability of linear stochastic systems operating under linear feedback. Automatica, 18:195-213.

Brillinger, David R. 1981. Time Series: Data Analysis and Theory, (Expanded Edition). San Francisco, Holden-Day.

Breiman, L., and J.H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. J. Am. Stat. Assoc., 80: 580-619.

Caines, P.E. and C.W. Chan. 1975. Feddback between stationary stochastic processes. IEEE Transactions on Automatic Control AC-20:498-508.

Cury, P. and C. Roy. 1989. Optimal environmental windows and pelagic fish recruitment success in upwelling areas. Can. J. Fish. Aquat. Sci., Vol. 46 (4): 670-680.

Dempster, A.P., N.M. Laird and D.B. Rubin. (1977). Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc., Ser. B. 39:1-38.

Gelfand, I.M. and A.M. Yaglom. 1959. Calculation of the amount of information about a random function contained in another such function. American Mathematical Society Translations, 12: 199-246.

Gersch, Will. 1986. Modeling multivariate covariance nonstationary time series and their dependent structure. Unpublished manuscript, Dept. of Information and Computer Sciences, University of Hawaii.

Geweke, John. 1982. The measurement of linear dependence and feedback between multiple time series. (With discussion). J. Am. Stat. Assoc., 77:304-351.

Geweke, John. 1984. Measures of conditional linear dependence and feedback between time series. J. Am. Stat. Assoc., 79:907-915.

Geweke, J. and S. Porter-Hudak. 1983. The estimation and application of Long Memory Time Series Models. J. Time Series Anal., 4:221-238.

Granger, C.W.G. 1980. Long memory relationships and the aggregation of dynamic models. J. Econometrics, 14:227-238.

Granger, C.W.G. and R. Joyeux. 1980. An introduction to long memory time series models and fractional differencing. J. Time Series Anal., 1:15-29.

Gustafson, I., L. Ljung, and T. Soderström. 1977. Identification of processes in closed loops - Identifiability and accuracy aspects. Automatica, 13:59-75.

Hosking, J.R.M. 1981. Fractional differencing. Biometrika, 68:165-176.

Hurst, H.E. 1951. Long term storage capacity of reservoirs. Trans. Amer. Soc. Civil. Engrs., 116:770-799.

Kashyap, R.L. and Kie-Bum Eom. 1988. Estimation in long-memory time series models. J. Time Series Anal., 9:35-42.

Lasker, R. 1978. The relation between oceanographic conditions and larval anchovy food in the California Current: identification of factors contributing to recruitment failure. Rapp. P.-V. Reun. Cons. Int. Explor. Mer, 173:212-230.

Mandelbrot, B.B. 1971. A fast fractional Gaussian noise generator. Water Resources Research, 7:543-553.

Mandelbrot, B.B. 1983. The Fractal Geometry of Nature. W.H. Freeman and Company, New York, 468p.

Mandelbrot, B.B. and J.W. Van Ness. 1968. Fractional Brownian motion, fractional noises and applications. SIAM Review, 10:422-437.

Mendelssohn, R. and P. Cury. 1987. Forecasting a fortnightly abundance index of the Ivoirian coastal pelagic species and associated environmental conditions. Can. J. Fish. Aquat. Sci., 44:408-421.

Mendelssohn, R. and P. Cury. 1989. Temporal and Spatial Dynamics of a Coastal Pelagic Species, *Sardinella Maderensis* Off the Ivory Coast. Can. J. Fish. Aquat. Sci., 46(10):1686-1697.

Mendelssohn, R. and D. Husby. 1992. Space-Time Analysis and Forecasting of Albacore Catch-per-unit-effort in Relation to the Ocean Environment. Manuscript in preparation.

Mendelssohn, R. and C. Roy. 1986. Environmental influences on the French, Ivory Coast, Senegalese and Moroccan tuna catches in the gulf of Guinea. p.170-188, *in*: Special Publication of the International Skipjack Year Program, ICCAT, Madrid.

Mendelssohn, R. and J. Mendo. 1987. Exploratory analysis of anchoveta recruitment off Peru and related environmental series. *In*: D. Pauly and I. Tsukayama (eds.), The Peruvian Anchoveta and Its Upwelling Ecosystem: Three Decades of Change. ICLARM, Manila, p.294-306.

Pierce, David A. 1979. R^2 measures for time series. J. Am. Stat. Assoc., 74:901-910.

Porter-Hudak, Susan L. 1982. Long-Term Memory Modelling - A Simplified Spectral Approach. Ph.D. Thesis, Department of Statistics, University of Wisconsin, Madison.

Shannon, C.E. 1948. A mathematical theory of communication. Bell Systems Technical Journal, 27:379-423, 623-656.

Shumway, R.H. and D.S. Stoffer. 1982. An approach to time series smoothing forecasting using the EM algorithm. J. Time Series Anal., 3:253-264.

Sund, P.N., M. Blackburn and F. Williams. 1981. Tunas and their environment in the Pacific ocean: A review. Oceanogr. Mar. Biol. Ann. Rev., 19:443-512.

Ulrych, T.J. and T.N. Bishop. 1975. Maximum entropy spectral analysis and autoregressive decomposition. Rev. Geophys. Space Phys., 13:183-200.

Woodward. P.M. 1953. Probability and Information Theory with Applications to Radar. McGraw-Hill, New York.