Environmental Influences on the French, Ivory Coast, Senegalese and Moroccan Tuna Catches in the Gulf of Guinea

R. MENDELSSOHN

Pacific Environmental Group, SWFC, NMFS, NOAA, P.O. Box 831, Monterey, Ca. 93942, U.S.A.

AND CI ROY[1]

Antenne ORSTOM, Centre Océanologique de Bretagne, B.P. 337, 29273 Brest Cédex, France

Both local and broadscale dynamic relationships were studied between catch per unit of effort (CPUE) for yellowfin and skipjack tunas in the Gulf of Guinea, and sea-surface temperatures (SST), and wind speed. The results suggested why a particular temperature during one time period will lead to high CPUE, while at another it will lead to low CPUE. Our analysis was not restricted to where fishing actually occurred.

We used a recently developed algorithm to complete missing CPUE and environmental data. Models were developed for eleven separate sub-areas of the Gulf of Guinea. Details of the method are provided in an appendix. Results from these models suggested that the environmental variables under study reflected an oceanographic process involving upwelling and concentration of nutrients a month prior to good fishing, followed by arrival of relatively warmer waters two weeks prior to good fishing.

We calculated dominant modes of variability for each parameter in time and space by six week and one year periods. The dominant mode for SST showed a space-time movement that was consistent with a recently developed theory of remote forcing in the equatorial Atlantic.

The dominant modes for CPUE for both yellowfin and skipjack showed relatively little discernible pattern. However, when dominant modes were calculated between SST and CPUE for the same periods and species of fish, the CPUE exhibited the same movement in space and time as SST, at about the same speed. This suggested the predictability of CPUE was largely due to a broadscale oceanographic process, which may have been associated with remote forcing in that region.

On a étudié les relations de la dynamique locale et à grande échelle entre la prise par unité d'effort (CPUE) de l'albacore et du listao dans le golfe de Guinée, les températures de surface (SST) et la vitesse du vent. Les résultats montrent pourquoi une température déterminée durant une période donne une CPUE élevée, alors qu'à une autre période elle entraîne une CPUE faible. Notre étude ne s'est pas limitée aux lieux de pêche actuels.

Pour obtenir les données de CPUE et de milieu manquantes, nous avons utilisé un algorithme récemment développé. Des modèles ont été établis dans le golfe de Guinée et ceci dans onze sous-zones séparées. Les détails de cette méthode indiquent que les variables de milieu sous étude reflétaient un procédé océanographique comportant un affleurement et une concentration d'aliments un mois avant une pêche fructueuse, suivi de l'arrivé d'eaux relativement plus chaudes, deux semaines avant une bonne pêche.

Nous avons calculé les principaux modes de variabilité de chaque paramètre dans le temps et dans l'espace sur des périodes de six semaines et un an. Les modes dominants de SST montraient un déplacement spatio-temporel compatible avec la théorie récemment élaborée de la pression de facteurs éloignés (''remote forcing'') dans l'Atlantique équatorial.

Les modes dominants de la CPUE d'albacore et de listao montrent un schéma assez peu défini. Néanmoins, lorsque les modes dominants étaient calculés entre les SST et la CPUE pour les mêmes périodes et les mêmes espèces de poissons, la CPUE signalait le même déplacement spatio-temporel que les SST, à peu près à la même vitesse. Ceci indique que la prévisibilité de la CPUE est surtout due à un procédé océanographique à grande échelle, qui pourrait être associé à la pression de facteurs éloignés (''remote forcing'') dans cette région.

Se estudiaron las relaciones dinámicas locales y a gran escala entre captura por unidad de esfuerzo (CPUE), para rabil y listado en el Golfo de Guinea, temperaturas de la superficie del mar (SST) y velocidad del viento. Los resultados indican porqué una determinada temperatura durante un período conduciría a una CPUE alta, mientras que en otro período conduciría a una CPUE baja. Nuestro análisis no se limitó al lugar de pesca actual.

Para obtener los datos de CPUE y de medio ambiente que faltaban, hemos utilizado un algoritmo recientemente desarrollado. Se establecieron modelos en el Golfo de Guinea para once sub-áreas separadas. Los detalles de este método se facilitan en un apéndice. Los resultados obtenidos a partir de estos modelos indican que las variables del medio ambiente estudiadas reflejan un proceso oceanográfico que comprende afloramientos y concentración de materia nutritiva, con un mes de anterioridad a una buena captura, seguida por la llegada de aguas relativamente cálidas dos semanas antes de una pesca abundante.

Se calcularon modos dominantes de variabilidad para cada parámetro en tiempo y espacio durante períodos de seis semanas y de un año. El modo dominante para SST mostró un movimiento tiempo-área que guardaba coherencia con una teoría recientemente desarrollada de ''forzamiento'' por factores remotos en el Atlántico ecuatorial.

Los modos dominantes para la CPUE del rabil y listado mostraron modelos relativamente poco definidos. Sin embargo, cuando los modos dominantes se calculaban entre SST y CPUE para los mismos períodos y especies de peces, la CPUE indicaba el mismo movimiento en el espacio y en el tiempo que el SST, aproximadamente a la misma velocidad. Esto sugería que la capacidad de pronóstico de la CPUE se debía principalmente a un proceso oceanográfico a gran escala, que podría haber estado asociado al ''forzamiento'' por factores remotos en esa zona.

[1] This work was undertaken during a one-year sabattical leave with the Pacific Environmental Group at Monterey, California.

1. Introduction

Environmental influences on fish population dynamics are being considered increasingly important as features to understand, and to integrate, into fisheries management models (see for example, Bakun and Parrish 1980; Sharp MS). One approach to developing a basic understanding of environmental effects is to study the dynamic relations between oceanographic and meteorological patterns in time and space and those of Catch-Per-Unit-Effort (CPUE) in a particular fishery. Any such analysis assumes that CPUE is an accurate reflection of relative abundance of the fish; this may not be the case. As we will show, a considerable part of the space-time variation in CPUE may be due to changes in where, when and how much the fleet fishes, and therefore CPUE may not be a good measure of relative abundance.

Recent works on the dynamics of the eastern part of the tropical Atlantic Ocean have focused on the study of upwelling at the equator and along the north and east coasts of the Gulf of Guinea (O'Brien et al. 1978; Voituriez 1981; Picaut 1982). Based on these results, and considering that upwelling areas are favorable to the concentration and capture of tuna, we thought that it would be worthwhile to investigate the dynamic relationships between the environment and the tuna fishery, not only at the particular time and place where the tuna are caught but also in the larger space and time domains.

2. The Data Used

2.1 Environmental Data

The National Climatic Center (U.S.) routinely collects environmental data from merchant ships in the area of study. From these data files, Sea Surface Temperature (SST) and the east-west and north-south components of wind velocity were extracted and screened to eliminate bad data. The environmental data were extracted for the years 1969–1979, to coincide with available catch and effort data. As the fisheries data are available by one degree square, and by fortnight, comparable mean series were calculated for the SST data and the two wind component series. A check on the accuracy of the SST data was made by comparing the extracted series with the historical means given in the atlas by Hastenrath and Lamb (1977); when the observed value was more than three standard deviations from the mean value reported in the atlas, it was discarded from the calculations. To obtain more complete spatial coverage, SST from the historical Nansen, Mechanical Bathy Thermograph (MBT), and eXpendable Bathy Thermograph (XBT) data sets in the area were also extracted and combined with the merchant ship data. Nevertheless, the amount of data from merchant ships is greater than that in the oceanographic data file, so our spatial coverage is mostly determined by the merchant ship reports.

2.2 Fisheries Data

The fisheries data used in the study come from the French, Ivory Coast, Senegalese, and Moroccan (FISM) tuna fleets. These data, collected by the *Office de la Recherche Scientifique et Technique Outre-Mer* (ORSTOM), are by one degree square and by fortnight since 1969, and when data are not missing the following information is available:

- total catch of yellowfin, skipjack, and bigeye tuna and total effort by bait boats
- total catch of yellowfin, skipjack, and bigeye tuna and total standardized effort by medium-size purse seiners
- total catch of yellowfin, skipjack, and bigeye tuna and total standardized effort by large-size purse seiners.

The catches are reported in units of 100 kg and the effort in hours fished. We extracted from this file the total catch and effort for yellowfin and skipjack tuna, combined over the classes of purse seiners, for which the effort had been previously standardized.

3. The Spatial Grid

The oceanographic features of the eastern part of the tropical Atlantic ocean are dominated by upwelling during the winter season in each hemisphere. In the northeast part (north of 5°N), which is influenced by the climatology of the boreal hemisphere, the cold upwelled waters occur along the coast from December to May and are mainly driven by local winds and by advection. The occurrence of these cold upwelled waters is associated with a thermal front which shifts seasonally between 10°N and 20°N (Wooster et al. 1976). The southern part (south of 5°N) is influenced by the climatology of the austral hemisphere. The cold season here occurs from May to September, and upwelling occurs during this period along the equator and along both the north coast (Ivory Coast, Ghana) and the east coast (Congo, Gabon) of the Gulf of Guinea. The coastal upwelling along the north coast is separated from the open-ocean upwelling by a zone of oceanic convergence north of the equator. In contrast to the situation in the northern hemisphere, in the Gulf of Guinea the upwelling seems to be not only locally forced, but is also the result of remote forcing (Servain et al. 1982). A second upwelling of smaller amplitude occurs along the northern and eastern coasts of the Gulf of Guinea from December to January.

Taking into account these main oceanographic features and the spatial extent of the FISM tuna fleet, the eastern part of the tropical Atlantic ocean was divided into eleven large areas (Fig. 1). Dividing the region into smaller areas would have meant either too little catch and effort data in each area, or a very low density of data for the environmental parameters. After first using only eight areas in which we included large regions straddling the equator, we found we needed to have areas that were restricted to one side of the equator. (An exception is area 6. This area extends to 1° south to include the northern side of the front off Cap Lopez which is associated with the dynamics of the waters of the Bay of Biafra which are relatively warm and have low salinity). This decision meant that some areas contained fewer fisheries data than we would have preferred.

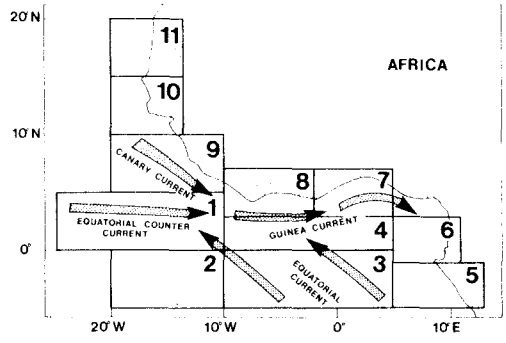

Figure 1. Spatial Grid used in this study, showing major currents in the Gulf of Guinea (after Ingham 1970).

Mean environmental and fisheries time series were calculated for each area, by fortnight, from the one degree square data. Mean CPUE was calculated as the mean of each fortnight's CPUE for each one degree square. Where there was fishing but no catch, the CPUE was defined to be zero. If there was no fishing, the CPUE was defined as missing data. The three environmental data series are missing little or no data in each area; however, the amount of missing CPUE data varies considerably among the eleven areas. The areas close to the coast have much more complete CPUE series than do the more oceanic areas, since the fleet has extended to the more oceanic areas only recently.

4. Comparison Between the Observed Time Series and the Estimated Values from the Local Models

4.1 Description of the Models

In the appendix we give a complete technical description of our approach to examining environmental influences in time and space on fish population dynamics. Briefly, the analysis consists of three steps: (1) estimating local (i.e. area specific) models that are used to fill in missing data points, necessary for the next two steps of the analysis; (2) calculating the dominant mode of variability in space and time of each variable, at a given frequency; (3) calculating the dominant mode of variability in space and time between two variables. In this section we examine the results of step (1).

A model of the form described in equation 1 in the appendix is estimated for each area. The model includes as variables: CPUE for both skipjack and yellowfin tuna, SST, and the north-south and east-west components of wind speed. The model is a simultaneous model in that each variable is assumed to be a linear function of all the variables of the model lagged at one and two fortnights previous to the present time period. Thus there are five equations for each area, and as the dynamics are assumed to interact, the five equations (twenty-five parameters) are estimated simultaneously. The resulting estimated parameter values are interpreted in a similar manner to multiple regression parameters.

As the results of our analyses are based on these completed time series, it is important as a first step to check how well the local models have "filled in" the missing observations. This can be done by comparing estimated and observed data when they both exist and, when there are missing data, by seeing if the estimated values agree with our qualitative knowledge of the fishery and of the environment. Further, the accuracy of the completed data was tested by artificially removing a segment of the CPUE data, treating the remaining series as the given data, then estimating the parameters of the local autoregressive models of order two (AR(2)) and the missing data points as above, and comparing the observed but artifically removed data with the corresponding estimated values produced by the model. (Only results for areas immediately relevant to the discussion are shown. Figures for all other areas are available upon request).

4.2 The Environmental Data

The environmental time series are relatively complete except for areas 5 and 6 (not shown) where there is incomplete coverage from 1971–1972, and for area 7 where the wind data are missing for all of 1969 (Fig. 2). Where we have data, the values predicted by the model are very close to the observed values except at very high peaks in SST and wind, where the predicted values tend to be less than the observed values (an example of this can be seen in area 8, where SST in April 1971 (Fig. 3) was as high as 30.5°C while the predicted values only reach a value slightly below 30°C). Otherwise, the differences

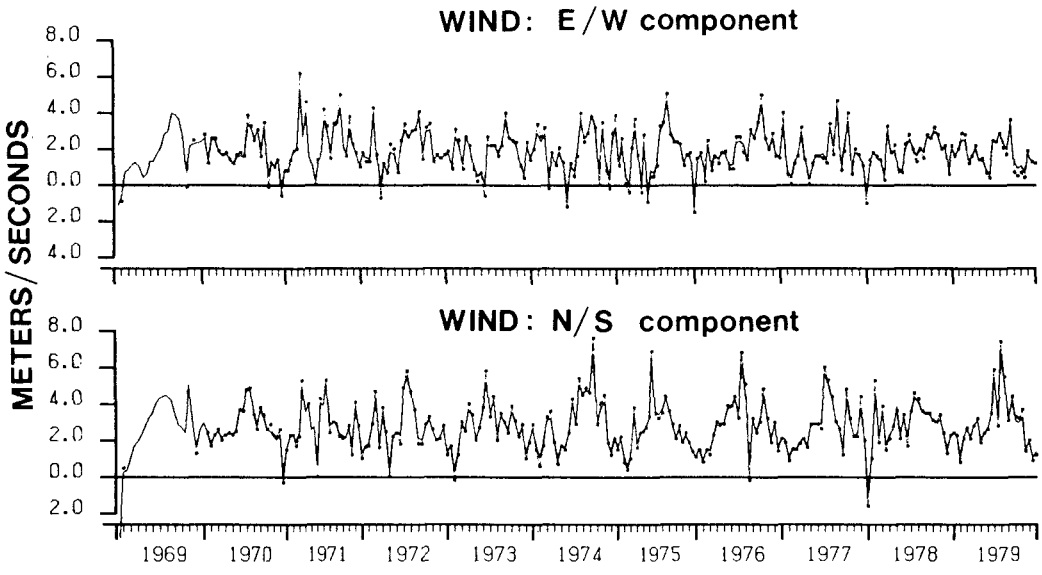

Figure 2. Observed wind speed (points connected by faint line) and wind speed predicted by the model (darker line) for area 7.

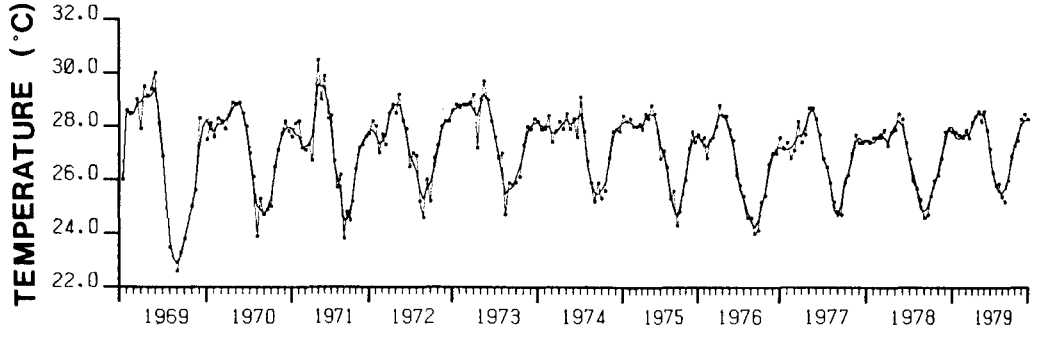

Figure 3. Observed Sea-Surface Temperatures (SST) (points connected by faint line) and those predicted by the model (darker line) for area 8.

between the observed and predicted values are barely discernible on the scale used in plotting the data.

For areas where there are many consecutive time periods with missing data, such as areas 5, 6, and 7, the predicted values again seem reasonable when we examine their qualitative properties. The predicted values during these periods of missing data are within the observed range of variation in the data, and have similar seasonal cycles as the observed part of the data (for example this is observed for area 6, from mid 1971 to early 1972). In area 7 (Fig. 2), where the wind data are missing for all of 1969, the fit seems reasonable except at the beginning of the year where there are unusually low values that are not consistent with observed values in surrounding areas. However, it is to be expected that the worst predictions would come at the beginning of the series if there are no data present, so this may partially explain the low predicted values.

4.3 The CPUE Data

In each of the eleven areas where data exist, there is very little difference between the predicted and observed values for CPUE of yellowfin and skipjack (e.g. Fig. 4). Again, the main difference appears to occur at maximum values of the peaks in CPUE. Even high frequency variations in CPUE, such as in area 5 (Fig. 4) are well predicted by the model. If there is some relationship between CPUE and the environment, this good fit is not surprising, because as the models used to complete the data were not designed to accurately forecast them, but rather to

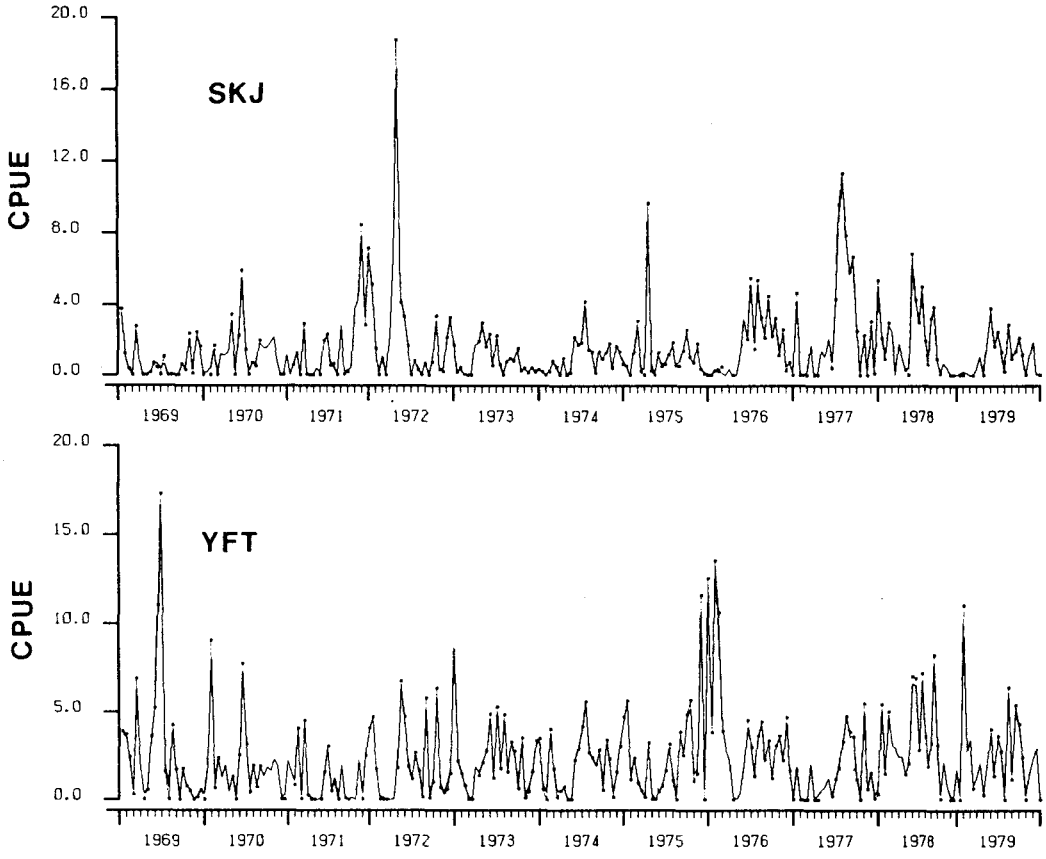

Figure 4. Catch per unit effort (CPUE) for skipjack (SKJ) and yellowfin (YFT) in area 5: observed CPUE solid points and faint line; CPUE predicted by the model, dark line.

distort as little as possible the statistical properties of existing data.

When the numbers of observations of CPUE are very low, such as in areas 1, 2, and 11 (Figs. 5, 6, and 7), it is doubtful that the predicted values are quantitatively accurate. However, the predicted data may capture known qualitative properties of the fishery that would give us some reason to believe that the models reflect the actual dynamics. In area 1 (Fig. 5), the model predicts a significantly positive CPUE for yellowfin, but not for skipjack, for the last four months of 1969, 1976, 1977 and for the last two months of 1974. Examining the environmental conditions during this period, we find values slightly lower than normal, but not out of line with conditions in which tuna can survive. In 1970, 1971, 1975 and 1979 there is positive CPUE during the same part of the year and for similar environmental conditions. The data for 1980 and 1981, not used in estimating the model, show positive CPUE for both species starting in September of the year and ending around March the next year. Thus the predicted pattern of CPUE is consistent with the observed tolerances of the fish.

In area 2 (Fig. 6), the predicted values show positive CPUE for both skipjack and yellowfin between June and the first three months of the next year, with the values for skipjack being more constant. The first three months of this period correspond to the main upwelling season, when the thermocline is relatively shallow and when high concentrations of nutrients are found near the surface. Later, water masses upwelling in the eastern part of the Gulf of Guinea are drawn into this area through advection by the South Equatorial Current. Thus the environmental conditions in area 2 during this time of year do appear to be favorable for tuna. The extension of the fleet into this area in 1980 and 1981 resulted in significant CPUE for skipjack in March and April, and again in the last two months of the year. Areas 1 and 2 are also considered to be reproductive areas for skipjack (Pianet 1983; Cayré and Farrugio this volume) where this fish

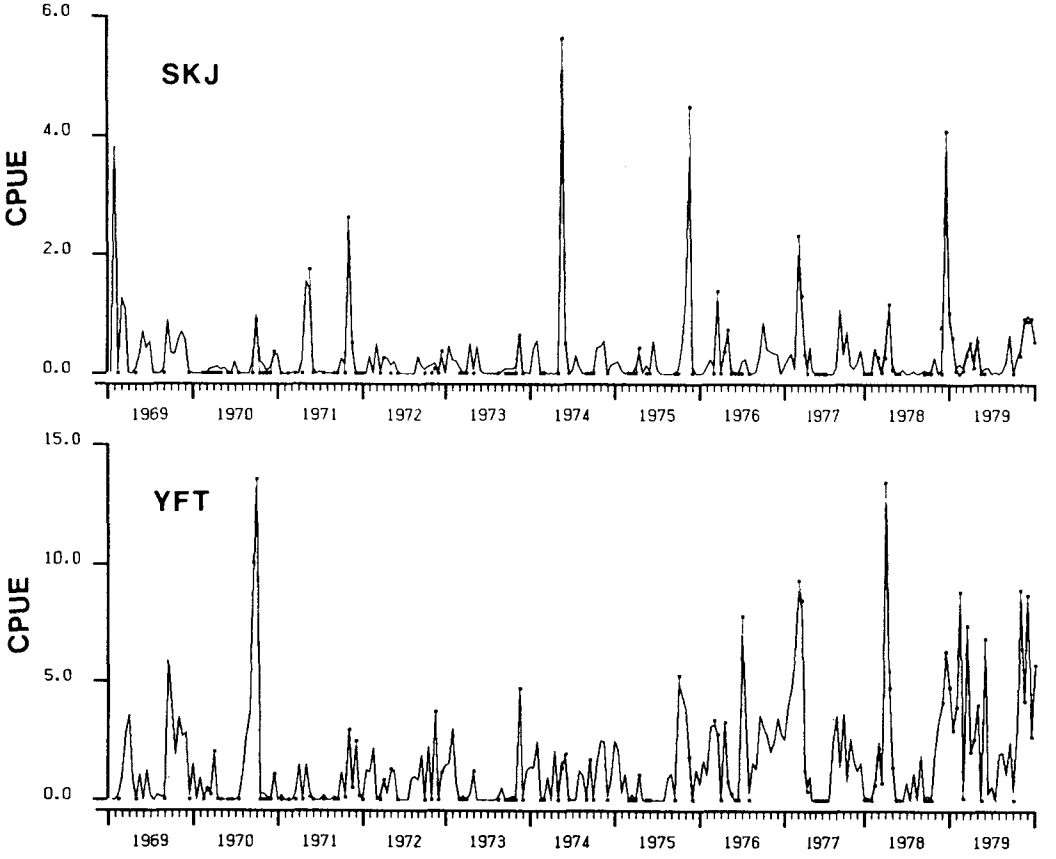

Figure 5. Catch per unit effort (CPUE) for skipjack (SKJ) and yellowfin (YFT) in area 1: observed CPUE solid points and faint line; CPUE predicted by the model, dark line.

migrates from the eastern part of the Gulf of Guinea and from the northeast region (Senegal and Mauritania) during the period from July to March of the following year.

In area 11 the fleet fishes only during the last seven months of the year, while the model predicts that there should be positive CPUE for both yellowfin and skipjack during the first five months also (Fig. 7). However, for the first four months of the year the SST values are generally below 20°C, (i.e., at temperatures where tuna are not generally encountered in this region). When fishing occurs during the last seven months of the year, the SST is at higher values, between 20°C and 28°C. As the estimated models are linear models, they will linearly extrapolate the relationship at the higher temperatures to the relationships at lower temperatures where fishing has not occurred. However, the relationship between CPUE and SST apears to have a threshold effect, and is therefore nonlinear. This is clearly a limit of our model, and cautions that the model will be most accurate when there has been at least some fishing over the entire range of the environmental variables.

Another way to test the validity of the estimation of CPUE by the AR(2) model is to remove a part of the CPUE data, to estimate the AR(2) model on the new reduced series and then to compare the estimated CPUE with the observed data not used in the estimation. This was done for CPUE data for both species in area 5. We removed the data starting from the second fortnight of May 1974 to the first fortnight of February 1975. For yellowfin the CPUE values estimated by the model were very close to the observed values (Fig. 8), showing the same three peaks as in the observed data. For skipjack, the estimated CPUE values were higher than the observed values (Fig. 8); however, if the predicted values are decreased by a constant amount the fit is quite good. This suggests that the local model is mispredicting the mean effect of the environment on CPUE for skipjack but describing in a reasonable manner the relative fluctuations. This may be because the bulk of skipjack catch is by

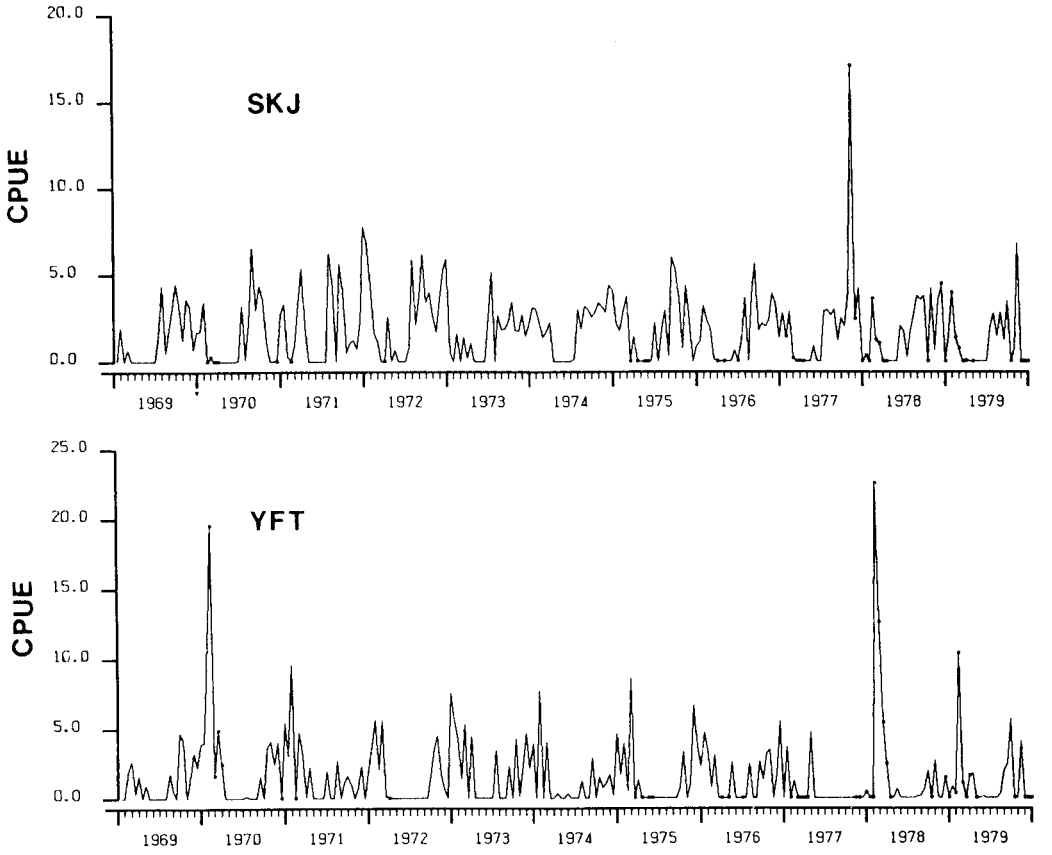

Figure 6. Catch per unit effort (CPUE) for skipjack (SKJ) and yellowfin (YFT) in area 2: observed CPUE solid points and faint line; CPUE predicted by the model, dark line.

the pole-and-line fishery. Catches of skipjack by the seiners probably depends on subsurface conditions that we only indirectly describe in our models. However, the estimated values for skipjack CPUE do appear to be qualitatively reasonable.

Our ability to predict the artificially removed yellowfin CPUE data could be due to three factors: strong seasonality in the data (i.e., we are only estimating the seasonal cycle); the fact that fishing conditions tend to persist over a number of time periods; or a strong relationship between the yellowfin CPUE and the other data in the model, that is the environmental data.

The estimated spectrum for the yellowfin CPUE data in this area (not shown) essentially contains no seasonal cycle. Moreover, the ten month stretch of data have high frequency variations that are also well fitted by the data estimated by the model. Thus it seems unlikely that we are only estimating the seasonal cycle. As ten months of data have been removed, again it follows that persistence is not the explanation for the close agreement.

Thus, logically we conclude that the close agreement between the artificially removed yellowfin CPUE data and the corresponding estimated data is due to a strong relationship between the environmental variables in the model (or more likely, between oceanographic processes they represent) and the yellowfin CPUE data. This fact is confirmed when we examine the formulas and parameter values used to estimate the missing data.

5. Interpretation of the Local Models

In this section, we examine the local models used to fill in the missing data to see if we can discern a pattern that would suggest what oceanographic processes might be associated with successful fishing. This is done by examining the relative importance of each of the five variables in predicting yellowfin and

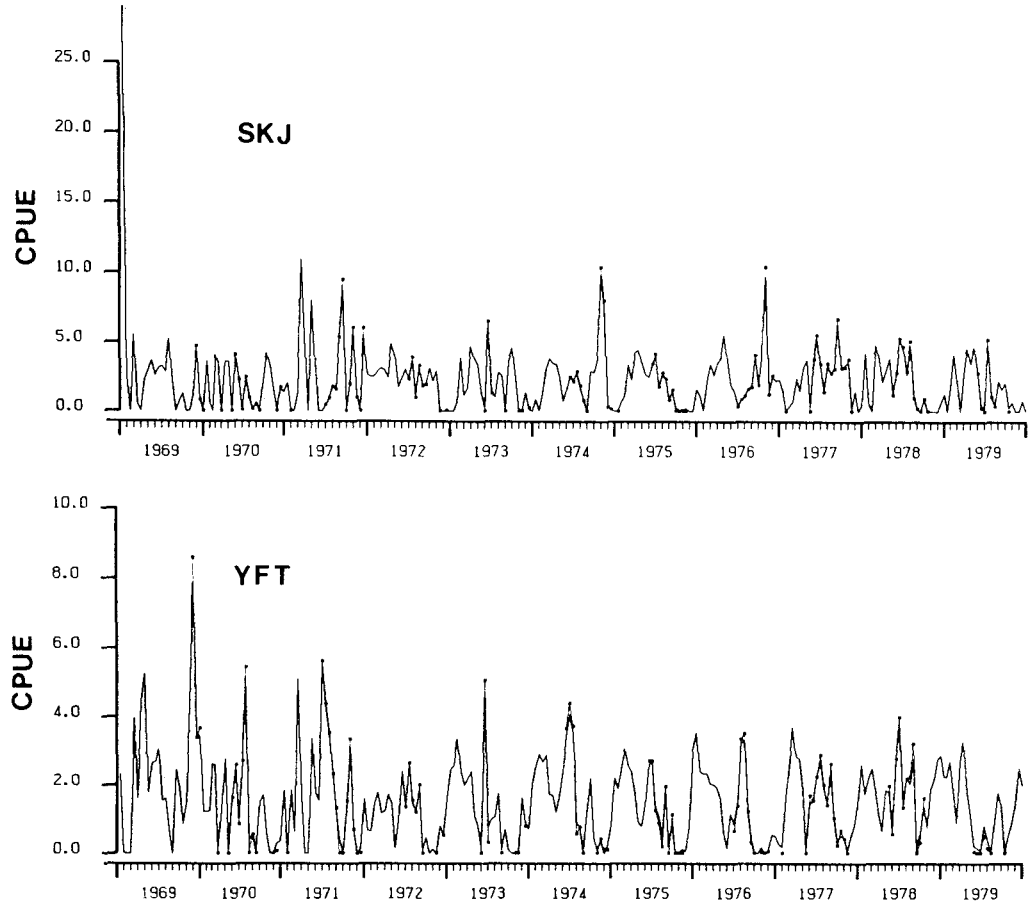

Figure 7. Catch per unit effort (CPUE) for skipjack (SKJ) and yellowfin (YFT) in area 11: observed CPUE solid points and faint line; CPUE predicted by the model, dark line.

skipjack CPUE. The most accurate way to determine the relative importance of each of the model parameters in predicting CPUE would be to standardize the coefficients by their standard errors. As discussed in the appendix, the computational burden involved precluded calculation of the standard errors. Instead we estimated the relative importance of each variable for predicting CPUE by looking at its average contribution to the predicted value, that is the relevant coefficient of Aij at time t–1 and of Bij at time t–2 times the mean of the variable in question. (For skipjack CPUE in area 1, for example, the average value for SST is 240 and the parameter at time t–1 has a value of .05 so the average contribution is .05 × 240 = 12; the north-south component of wind speed has an average value of 81 and the parameter at time t–1 has a value of –.09 for an average contribution of –.09 × 81 = –7.3. SST thus has a slightly larger average contribution.)

The estimated values of CPUE, SST and wind for all areas are given in Table 1. Figure 9 synthesizes results from the model for yellowfin and skipjack CPUE. The "SST" or "Wind" displayed in each area in the figures shows the variable(s) that were determined to be the most important, on average, for predicting the CPUE in that area (it should be noted, however, that at any particular time period a different variable may have been the most important in determining the level of CPUE for that fortnight). Also, Figure 9 shows what relative values of SST would give the largest predicted CPUE in that area. Thus, for yellowfin CPUE in area 8, the estimated value for SST one fortnight previous is 0.263 and for two fortnights previous is –0.268 (Table 1). Therefore, the model predicts a relatively larger CPUE when SST is relatively warmer one fortnight ago (to magnify the effect of the positive parameter value), and relatively cooler two periods ago (to minimize the effect of the

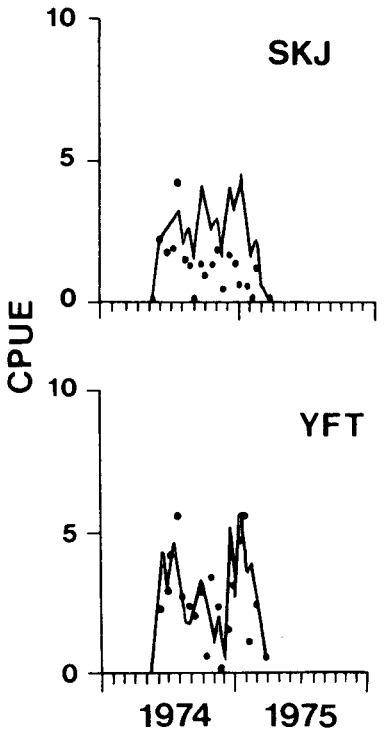

Figure 8. CPUE for skipjack (SKJ) and yellowfin (YFT) estimated by the model from the second fortnight in May 1974 to the first fortnight in February 1975 where observed data for the period was removed before fitting the model. The actual data points are indicated by the circles.

negative parameter value). This is indicated for area 8 in Figure 9 (yellowfin) by the letters SST, and W(warm) associated with t–1, and C(cold) with t–2.

The most obvious aspect of these data is that for both species, SST on the whole is the most important variable for predicting CPUE (Fig. 9). This is not surprising as many other workers have noted the association of tuna with particular water temperatures. However, having the desirable range of SST while fishing does not guarantee good catches. For yellowfin (Fig. 9), SST is the dominant or codominant factor in predicting CPUE everywhere except in areas 1, 5, and 7. Areas 5 and 7 are characterized by strong thermal fronts (at Cap Lopez and Cap Three Points) and area 1 is associated with the Guinea dome. Wind is codominant with SST in predicting CPUE in areas 9 and 10, which also are characterized by the occurrence of strong thermal fronts. For predicting skipjack CPUE (Fig. 9), wind is the dominant factor in two of the areas which are also characterized by thermal fronts (areas 6 and 8).

Another striking feature is that CPUE for yellowfin and skipjack one fortnight (t–1) and two fortnights (t–2) prior to the period under consideration is not an important contributor to the predictions. As the estimated models are simultaneous models (section 4.1), persistence of CPUE by itself is clearly not adequate to explain the observed pattern.

Another interesting pattern is observed when one looks at the coefficients for SST for each area (Table 1). The coefficients always change sign from period t–1 to period t–2, while the magnitude of the coefficients are similar at t–1 and t–2. One major grouping of ares (4, 8, 9, 10 and 11) has a positive SST coefficient one fortnight earlier and a negative SST coefficient two fortnights earlier (Table 1). According to this model, in these areas optimal conditions for high CPUE would occur if there are relatively cooler waters a month (two fortnights) ago followed by relatively warmer waters a fortnight ago. This is consistent with a physical process of colder waters richer in nutrients coming to the surface, followed and sustained by warmer waters at a sufficient

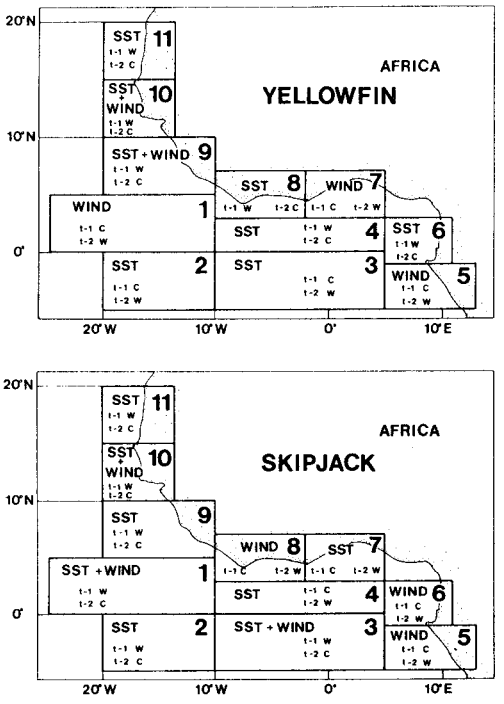

Figure 9. Summary of Table 1 for yellowfin (YFT) and skipjack (SKJ). In the upper corner of each area is the environmental parameter which contributes most to the predictive power of the model in that area. C (colder water) and W (relatively warmer water) in the lower corner indicate the time periods t–1 and t–2 that yield optimum CPUE.

time lag to allow the biological processes to develop, and for the occurrence of water temperatures more favorable to the immediate presence of tuna. Thus, as stated previously, the occurrence of favorable water temperatures is not sufficient to insure high abundance of tuna. The same pattern of relatively colder SST followed by relatively warmer SST is used in a model developed by Stretta and Slepoukha (1983) that determines fishing probabilities in the eastern part of the tropical Atlantic. While our models have not been tested as to their ability to forecast future levels of CPUE, we feel they are close to achieving this goal. The most difficult part, surprisingly, would be to forecast accurately the environmental conditions in a particular location. If these were known, there is some reason to believe accurate area by area forecasts of CPUE could be developed with lead times of two weeks to one month.

6. Space and Time Relationships

6.1 Single Variable Analysis

The results of step 2 of the analysis is described in this section. The data are analyzed using a technique called "Principal Components in the Frequency Domain" (PCFD). The idea is to find at each fre-

Table 1. Average contribution of five parameters to predicted value of CPUE and values of coefficients A_{ij} at time t−1, and B_{ij} at time t−2 for yellowfin (YF) and skipjack (SJ), as predicted by the model.

		PARAMETERS				
		CPUE YF	CPUE SJ	SST	WIND NS	WIND EW
		Av. contrib. A_{ij} B_{ij}	Av. contrib. A_{ij} B_{ij}	Av. contrib. A_{ij} B_{ij}	Av. contrib. A_{ij} B_{ij}	Av. contrib. A_{ij} B_{ij}
Area	Sp.	(t−1) (t−2)	(t−1) (t−2)	(t−1) (t−2)	(t−1) (t−2)	(t−1) (t−2)
1		−.99	−12.68	244.10	81.66	87.64
	YF	.3992 .1377	−.3655 .3128	−.0866 .1029	−.7300 .8771	.9694 −.6667
	SJ	.0273 .0371	.0824 −.1445	.0558 −.0483	−.0325 .0634	−.0974 .2862
2		13.96	4.54	234.94	57.52	32.55
	YF	.2563 .1971	.3164 .0187	−.3308 .2021	−.2007 1.3779	.5713 −.1400
	SJ	−.1157 −.0183	.3226 .2992	.0854 −.3080	−.9623 1.5594	−1.1304 −.0010
3		7.56	−3.93	219.80	27.65	20.34
	YF	.2243 .1762	.2227 −.3026	−.5507 .3473	.7931 .8018	.2496 −.0106
	SJ	−.0559 .0421	.2943 −.0276	.0447 −.1451	−.2410 .7337	−.7090 .3203
4		12.30	−3.52	239.32	52.76	64.00
	YF	.1001 −.0435	.3183 .0668	.3665 −.3806	.5093 −.1116	.1680 .1299
	SJ	−.0257 .0005	.3472 .0838	−.1575 .1556	.0199 .0909	−.0262 .1768
5		20.99	29.69	226.29	76.74	39.12
	YF	.1918 .1601	.1613 −.1354	−.0018 .0149	.0467 .2730	.0701 −.5037
	SJ	−.1040 −.0812	.4399 .1030	−.1367 .1390	.4005 −.1027	.1210 −.1145
6		33.81	28.95	247.29	48.04	56.69
	YF	.4177 −.1178	−.0490 .1385	.0974 −.0584	−.1118 .1840	−.0388 −.0026
	SJ	.0133 .0284	.3157 −.0216	−.0097 .0083	.0370 .1702	.0482 .1158
7		25.91	6.26	274.36	87.40	46.89
	YF	.2493 −.0805	−.0678 .0800	−.0038 .0034	.1818 .0644	.1586 .0575
	SJ	.1999 −.1732	.2483 −.0458	−.1217 .1444	.0289 .0090	.1607 −.1864
8		69.09	−1.79	203.00	94.16	53.48
	YF	.3208 −.0788	.2805 −.0324	.2630 −.2676	.2786 −.0269	−.0088 .1598
	SJ	.0013 .0087	.5067 −.0519	−.0087 .0065	.0687 .0170	−.0669 .1352
9		−23.99	21.18	265.80	.98	16.73
	YF	.2497 −.0314	.0088 .0761	.0469 −.0377	−.1131 −.0879	−.3317 1.2914
	SJ	.0371 −.0743	.1700 .0946	.2309 −.2123	.0323 −.0331	−.7241 .7539
10		17.70	15.97	235.96	36.16	23.78
	YF	−.0048 .0553	.2269 −.0552	.0660 −.0679	−.5808 .4520	.4996 −.0383
	SJ	−.0933 −.0695	.4325 −.0400	.0390 −.0088	.0389 .0032	.1395 .0591
11		−3.62	−1.67	203.21	21.89	56.42
	YF	.2937 .0845	−.0799 −.2876	.1215 −.1195	−.2366 −.0595	.0970 −.1214
	SJ	−.7382 .4539	.4420 −.0344	.9800 −.8695	.3106 −.0679	1.3241 −1.9779

quency the dominant mode of variation in the data. That is, at each time period we form a one-dimensional time series by taking a linear (weighted) combination through time of the relevant data series for the eleven areas that have the property of containing the most spectral variance at all frequencies compared to any other such combination. Thus the new series at any time period t is a linear combination of the eleven areas at period t, plus another linear combination of the eleven areas at time t–1, and so forth.

The new component series is an artifice, and acts as a reference point. As we calculate the relevant weights in the frequency domain, we can calculate the coherence (the spectral equivalence of correlation) and phase between any of the original series and the component series at any frequency. The coherence tells us how much of the spectral variance of any original series (at a given frequency) is contained in the component series, that is how well the component series reflects the dynamics of a given area.

The phase tells us by how much any original series lags behind the component series (at a given frequency). As the component series is an artifice, we have in effect calculated the relative lags between all eleven areas (at a given frequency). As the eleven series are distributed in space, and the lags are in time, this gives us a description of the dominant space-time dynamics (at a given frequency). Thus, the results to be presented show the dominant space-time dynamics of a variable at a given frequency. At the annual frequency, this would be like showing the dominant "average" dynamic that takes one year to occur, rather than the dynamics of any particular year, just as a regression line describes "expected" or "average" behavior.

6.1.1 Analysis of the SST data

In the tropical Atlantic ocean, the amplitude of the seasonal cycle of SST is several times greater than that of the interannual variability (Merle et al. 1980). A fourier decomposition of both the SST data and of thermal structure data has shown that the seasonal cycle is mainly composed of the first two harmonics, annual and semi-annual (Merle and Le Floch 1978).

For this reason, we describe our results for frequencies centered at periods of twenty-two and eleven fortnights. These correspond to the annual and semi-annual frequencies in our data. Only the first, or dominant *eigenvector* is discussed for each frequency, because at each frequency this component contains the predominance of the spectral variance (90%). As described in the appendix the phases are calculated by arbitrarily giving the dominant component a phase of zero with area 4.

For a period of twenty-two fortnights (Fig. 10), the coherence between the dominant principal component and each of the areas is very high except in area 9, off Sierra Leone. Area 9 is the meteorological equator (or atmospheric thermal frontier between the two hemispheres) and is affected by the bi-annual oscillations of the Inter-Tropical Convergence Zone, inducing a strong bi-annual oscillation in the SST data. When we look at the phase relations of the SST at this frequency, we see a decomposition of this part of the Atlantic into two parts; a northern part (area 10, 11) separated from the southern part by area 9. The phases in these two areas differ by roughly eleven fortnights (six months), which corresponds to the delay between the seasonal cycles of the two hemispheres.

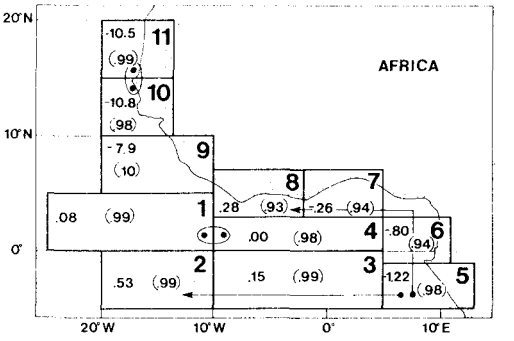

Figure 10. Phase and coherence (in parentheses) of SST between the eleven areas at twenty-two fortnights and for the first component of the *eigenanalysis*. The phases are in units of fortnights. The arrows represent our interpretation of the phase lag progression. Areas in phase are noted by encircled pairs of dots.

The southern part shows an interesting spatial pattern in the phase relationships at this frequency. Our interpretation of this pattern is based on a theory of remote forcing in the Gulf of Guinea (O'Brien et al. 1978). In this theory, the main upwelling in the Gulf of Guinea is believed to be caused by a Kelvin wave generated by an increase of the zonal wind stress in the western part of the Atlantic basin. This wave then propagates along the equator to the eastern part of the basin, and then poleward along the northern and southern coasts of west Africa, inducing a movement of the thermocline towards the surface. In Figure 10, we see that the equatorial areas north of the equator are in phase. Along the Gulf coast, from areas 6 to 8, the phase lag between areas is relatively constant (approximately 0.4 fortnight) and is consistent with the estimated speed of a Kelvin wave (1½–2 m/sec compared to a theoretical 3 m/sec). The phase lag between area 5 and area 6, however, shows movement from south to north that is not consistent with a Kelvin wave. To completely resolve this question, we

need to extend the study area farther south. However, a more detailed analysis of the SST data, to be described elsewhere, lends further support to the Kelvin wave theory. The two oceanic areas south of the equator have phase lags which show a westward movement of water apparently related to advection by the South Equatorial Current.

At the higher semi-annual frequency (Fig. 11) all the areas are highly coherent with the dominant principal component, except for area 11. The two equatorial areas north of the equator are still approximately in phase. As for the annual frequency, we again find phase lags which suggest a movement of water from area 6 to 8, but at these frequencies the movement extends to area 10. This extension at the semi-annual periods is consistent with our previous discussion about area 9 being the meteorological equator dominated by semi-annual frequencies. Also the oceanic areas south of the equator exhibit a westward movement of water similar to that at the annual frequency. Significant six month signals have been found in the subsurface thermal structure (Merle and Le Floch 1978), in sea level data (Verstraete et al. 1980), and in the wind and coastal station data (Picaut et al. 1978). Therefore, there is good reason to believe that the observed behavior at a period of six months is not just due to a simple harmonic of the annual cycle. The second upwelling event, which occurs south of the equator in December and north of the equator in January, also suggests that the semi-annual behavior is not simply a harmonic. The same kind of propagation of the SST signal has been demonstrated to occur along the northern and southern coasts during the main upwelling period by Picaut (1982), and along the coasts of Ghana and the Ivory Coast during the second upwelling period by Roy (1981).

6.1.2 Analysis of CPUE data

The dominant mode of variation displayed by the CPUE data for both yellowfin and skipjack at the periods of interest (eleven and twenty-two fortnights) does not present as clear a picture as does the SST data. At many frequencies it is not clear that the variations in CPUE represent anything more than fleet dynamics. For example, at a period of twenty-two fortnights (Fig. 13) the dominant pattern of CPUE for yellowfin appears to be a movement in the CPUE from area 8 down to area 1 and 2, and from area 10 to area 11 and 9, which is consistent with fishing behavior of the fleet from one of the ports. Skipjack CPUE especially shows little pattern, with the highest coherences at most frequencies occurring in the areas adjacent to the major ports.

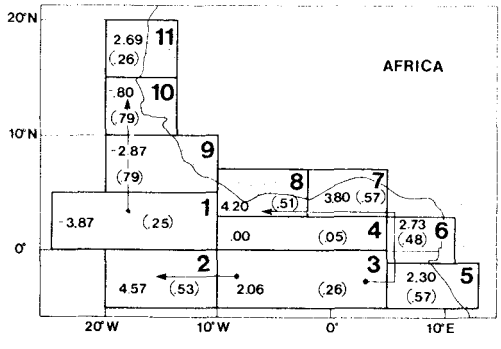

Figure 12. Phase and coherence (in parentheses) of CPUE for yellowfin between the eleven areas at eleven fortnights and for the first component of the *eigenanalysis*. The phases are in units of fortnights. The arrows represent our interpretation of the phase lag progression.

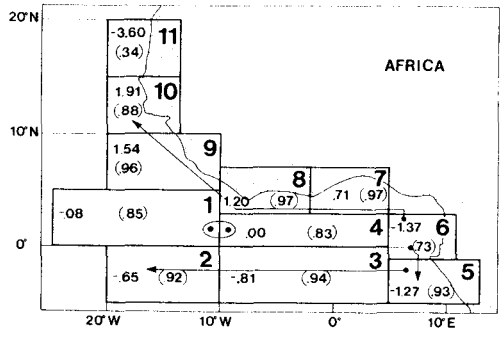

Figure 11. Phase and coherence (in parentheses) of SST between the eleven areas at eleven fortnights and for the first component of the *eigenanalysis*. The phases are in units of fortnights. The arrows represent our interpretation of the phase lag progression. Areas in phase are noted by encircled pairs of dots.

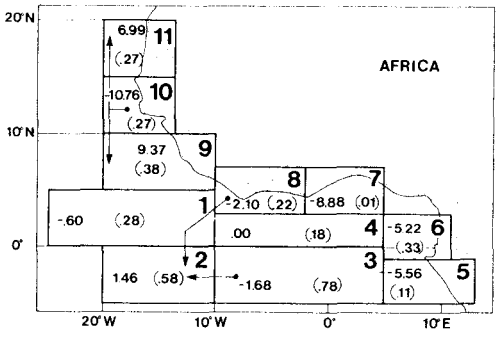

Figure 13. Phase and coherence (in parentheses) of CPUE for yellowfin between the eleven areas at twenty-two fortnights and for the first component of the *eigenanalysis*. The phases are in units of fortnights. The arrows represent our interpretation of the phase lag progression.

6.1.2.1 Yellowfin CPUE: At a period of eleven fortnights (Fig. 12) the yellowfin CPUE shows a fairly regular propagation northward along the coast from area 5 to area 6 to area 7 to area 8. This is consistent with one part of the dominant pattern in the SST's, and suggests the association of the arrival of tuna with the proper water temperatures (a question to be explored more fully in the next section). Also, at a period of 14.67 fortnights, there appears to be a steady progression northward from area 9 to area 10 and to area 11.

6.1.2.2 Skipjack CPUE: There appears to be little pattern in the skipjack CPUE except for close association with the major ports. Only at a period of eleven fortnights (Fig. 14) is there any pattern. Here we see a fairly regular propagation from area 1 to area 9, from area 11 to area 10 and to area 9, and from area 4 to 6, to 5 and to 3. The last feature suggests that fishing occurs along entrained upwelling waters, as discussed previously. The southwest propagations from area 11 to 9 could be related to the development of the thermal front associated with upwelling.

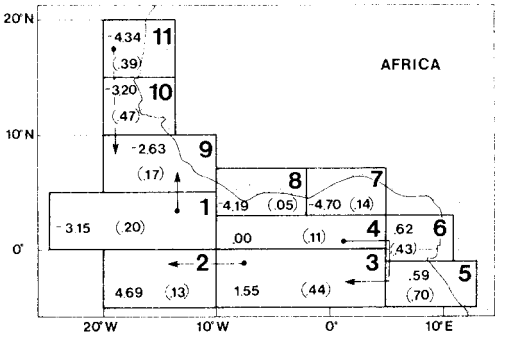

Figure 14. Phase and coherence (in parentheses) of CPUE for skipjack between the eleven areas at eleven fortnights and for the first component of the *eigenanalysis*. The phases are in units of fortnights. The arrows represent our interpretation of the phase lag progression.

What is most notable is that there appears to be relatively little understandable pattern to the dominant modes of variation in the CPUE data. As we will see in the next section, the situation changes dramatically when CPUE is analyzed in conjunction with SST. Thus this lack of pattern in the CPUE data is significant for two reasons: it underscores the importance of the environmental features, and it reinforces the need to look at each variable individually before examining a variable in conjunction with a second variable.

6.2 Analysis of Two Variables Over Space and Time

In this section, we examine the results for step 3 of our analysis. As in step 2, we work in the frequency domain, and as in step 2 we will form a one-dimensional component series that is a linear combination through time of the relevant variable measured at each of the eleven areas. However, the criterion for selecting the linear combinations is different. As we want to examine the influence in space and time of one variable on a second variable, we will be looking for the linear combinations through time of each variable that are the *most coherent* at all frequencies, rather than the linear combination that explains the most spectral variance at all frequencies.

In section 5, we presented evidence that CPUE for tuna in the Gulf of Guinea is strongly influenced by the presence or absence of particular environmental conditions. However, that analysis does not tell us if the relationship is mainly a local one, in which case there is little added predictive capability from knowing what is happening in the entire Gulf, or if the strong relationship is due to a broadscale process, in which case spatial patterns will be as important as temporal patterns in making predictions.

This analysis was performed for yellowfin and skipjack CPUE with SST as the second variable. The one-dimensional series formed by the analysis are called the *canonical series*. As before, we can calculate the phase and coherence between any original series and the corresponding canonical series (at a given frequency).

The most obvious feature of this analysis is that the pattern of propagation shown by SST is very similar to the pattern when SST is analyzed alone (compare e.g. Figs. 12 and 15). At periods of both eleven and twenty-two fortnights, we see the fairly regular propagation of SST along the coast (Figs. 15 and 16; 17 and 18), as well as a westward propagation of SST in the oceanic areas south of the equator. At a period of twenty-two fortnights (Figs. 16 and 18), we again see

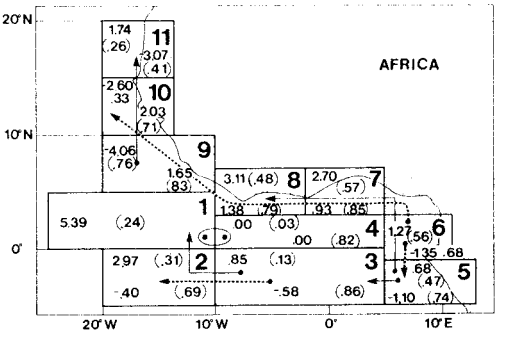

Figure 15. Phase and coherence (in parentheses) of CPUE for yellowfin (upper corner) and SST (lower corner) between the eleven areas at eleven fortnights and for the fixed components of the canonical analysis. The phases are in units of fortnight. The arrows represent our interpretation of the phase lag progression for yellowfin (solid line) and for SST (dashed line). The areas in phase are denoted by encircled pairs of dots.

for the two species that the northern regions are roughly six months out of phase with the southeastern regions. The importance of this observation is that this is the part of the SST data at these frequencies that is most coherent (coherences are all 0.9 or greater) with the CPUE data. This is very strong evidence that the evolution of CPUE for both skipjack and yellowfin in the Gulf of Guinea is closely related to the evolution of SST at the same frequencies. If we examine the phase relationships at a period of twenty-two fortnights in the yellowfin CPUE canonical series (Fig. 16), it is clear that except for area 7, where the coherence is low (0.04) the yellowfin CPUE data show the same spatial evolution in time along the eastern coast as does the SST canonical series. For the skipjack CPUE data (Fig. 18), the pattern is not as clear at this frequency.

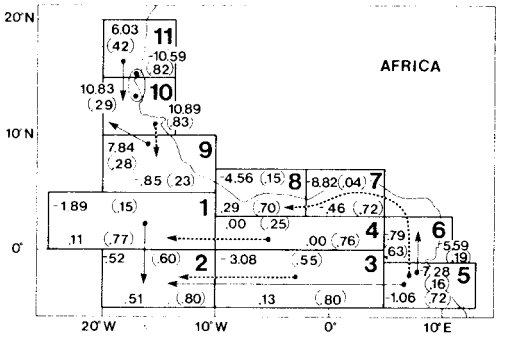

Figure 16. Phase and coherence (in parentheses) of CPUE for yellowfin (upper corner) and SST (lower corner) between the eleven areas at twenty-two fortnights and for the fixed components of the canonical analysis. The phases are in units of fortnight. The arrows represent our interpretation of the phase lag progression for yellowfin (solid line) and for SST (dashed line). The areas in phase are denoted by encircled pairs of dots.

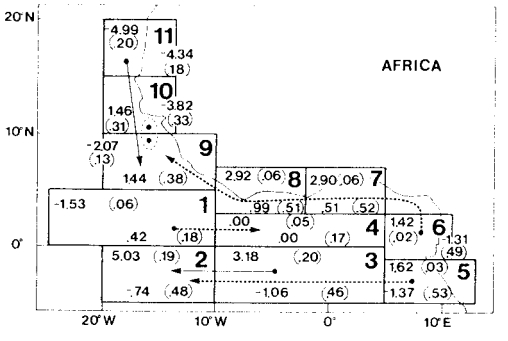

Figure 17. Phase and coherence (in parentheses) of CPUE for skipjack (upper corner) and SST (lower corner) between the eleven areas at eleven fortnights and for the fixed components of the canonical analysis. The phases are in units of fortnight. The arrows represent our interpretation of the phase lag progression for skipjack (solid line) and for SST (dashed line). The areas in phase are denoted by encircled pairs of dots.

At a period of eleven fortnights, again yellowfin CPUE and SST have remarkably similar patterns and speed of movement (Fig. 15). For skipjack CPUE and SST at this frequency, the coherence is low along the eastern coast and for the two oceanic areas north of the equator. Along the northern coast (area 11, 10, 9) the CPUE shows a regular movement from area 11 to 10 and to 9.

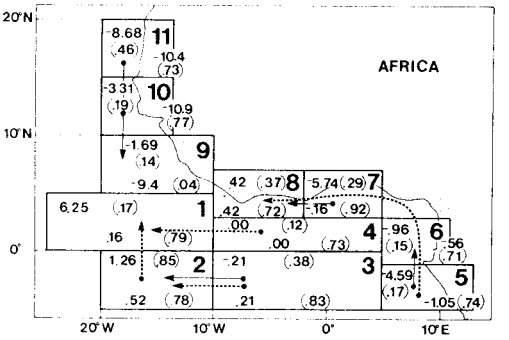

Figure 18. Phase and coherence (in parentheses) of CPUE for skipjack (upper corner) and SST (lower corner) between the eleven areas at twenty-two fortnights and for the fixed components of the canonical analysis. The phases are in units of fortnight. The arrows represent our interpretation of the phase lag progression for skipjack (solid line) and for SST (dashed line). The areas in phase are denoted by encircled pairs of dots.

As mentioned in an earlier section, the phase relationships only have meaning for a single variable across areas. A more precise discussion of these results would require calculating the phase lag between the two canonical series, so that we can better determine the timing between the propagation of SST and of CPUE. However, what is remarkable is that while the dominant modes of variation for CPUE alone show no consistent pattern, when analyzed simultaneously with SST data, the most predictable parts of the CPUE data exhibit very clear patterns of movement. The propagation of SST, coherent with the CPUE, appears to be related to the occurrence of upwelling along the coast in the Gulf of Guinea, and with advective processes in the oceanic and northern parts of the Gulf. Thus from a physical viewpoint, it is not surprising that the CPUE data should show an evolution similar to the SST propagation, as the regions with upwelling waters have an enrichment of food that is favorable for the aggregation of tuna.

7. Summary and Conclusions

In this paper we have studied both the local and the broadscale dynamic relationships between CPUE for yellowfin and skipjack tuna in the Gulf of Guinea with SST and with wind speed. It is not our purpose to review past work on environmental influences on

tuna catches. Good reviews can be found in Blackburn (1965) and Sund et al. (1981). What separates our work from any of the papers referenced in those two works is that first, we truly analyze the dynamics of the relationships, rather than just the mean habitat where high levels of CPUE occur. This allows us, for example, to explain why a particular temperature during one time period will lead to high CPUE while the same temperature conditions will lead to low CPUE at other times. Moreover, our analysis is not restricted to where fishing has actually occurred; if we are ever to forecast where good fishing will occur the only information available will be present environmental conditions and past fishing success. The forecast must be able to distinguish, for example, fronts that will have tuna aggregated and fronts that won't, *before fishing occurs*. To our knowledge our models are the first to directly tackle this problem.

In more detail, an examination of our local models shows that CPUE for both species of tuna is primarily estimated by the evolution of the environmental conditions, rather than from either persistence of fishing conditions or by the present environmental conditions. The "scenario" used by the models is consistent with known direct observations at sea. This "scenario" predicts higher concentrations of tuna when there has been upwelling one month previous to fishing followed by a relative warming of the waters two weeks previous to the fishing. This suggests that the predicted relationship is only a surrogate for a more complicated evolution of biological and physical processes which results in areas favorable for the aggregation of tuna.

The local models suggest that there are broad areas with similar behavior, and that these areas correspond to areas with similar oceanographic dynamics. To examine this further, we used a technique that allows us to study the dominant dynamics of one or two variables in both space and time and at a given frequency. This analysis reveals little pattern when CPUE is analysed by itself, but when CPUE is examined with SST we find the two variables moving in a similar pattern with a similar speed of movement. The pattern found is identical to the pattern found for SST when it is analyzed by itself, and is consistent with a theory of remote forcing in the region. This suggests that fishing conditions also are associated with the remote forcing, and that forecasting of CPUE could be enhanced by using both broadscale and local environmental conditions. The relationship to remote forcing is an entirely new result, and emphasizes the importance of not limiting studies of the relationships between CPUE and the environment to locations and times where fishing occurs.

In this study, we only used surface environmental parameters. It would be interesting and instructive to introduce some properties of the vertical structure as well. For example, the depth of the thermocline would be an important parameter to consider, because the variation of the vertical habitat of the fish and the vulnerability of the fish to surface fishing gear is related to the vertical temperature structure. However, subsurface data tends to be much sparser than surface data.

This study examines the influence of the environment only after the recruitment phase; at this stage of the life-history of pelagic fish environmental effects are not a dominant factor for the survival of the fish, but rather tend to affect the ability of the fish to aggregate and to be available to a surface fishery. Before the recruitment phase, the environment could be the dominant factor for the survival of the stock. The success of reproduction and the survival of larvae may be controlled by advective or transport processes (Parrish et al. 1981).

A curious feature of our results is the fact that none of our models directly includes recruitment. This suggests that either we have somehow indirectly modelled recruitment (unlikely), or that recruitment during the period of study has been steady, perhaps fluctuating around a mean level, or else that the fishery during the period of study has not been limited by the population size. Studies of the relationships between the environment and the larval and juvenile stages of the tuna would be necessary to help clarify these issues.

Acknowledgements

Support for one of the authors (Cl. R.) was provided through a cooperative agreement between NMFS and ORSTOM during a visit to the Pacific Environmental Group, Monterey, California. We wish to express our gratitude for the many hours of assistance rendered by Craig Nelson in preparing the data for our calculations and in reviewing the manuscript. This study has also benefited from an efficient preparation of the raw data by Jean Jacques Le Chauve and Daniel Corre of ORSTOM, Brest, France, and from detailed comments of an anonymous reviewer.

Appendix

Methods and Assumptions of the Data Analysis

A.1 The Basic Problems

There are three main problems to be dealt with when trying to analyze the tuna CPUE and environmental data from the Gulf of Guinea. These are: (1) the sheer size of the problem, (2) missing data in the time series, and (3) resolving time and space relationships among several variables at once. We will briefly discuss each of these problems. First, as described in the preceding sections, we want to examine the interrelationships among five different variables at eleven locations. If each variable in each area is treated as a separate time series, then we would be directly modelling fifty-five time series. In this case, even a simple model of the form $X_t = AX_{t-1} + e_t$ would require estimating 3025 parameters in the coefficient matrix A, which clearly is impractical.

Second, only a few of the fifty-five time series have no missing data points. While the environmental data are fairly complete, the CPUE time series have many missing values, because there are no data for an area at any time during which there was no fishing in that area. This clearly is a problem if we consider CPUE as a measure of relative abundance. The absence of fishing does not necessarily mean the absence of fish, and when studying the space-time dynamics the most consistent estimate may well be that there were fish in the area and therefore, that a positive CPUE would have been possible.

The third problem is to resolve behavior in both time and space among several variables. For the problem at hand, there are good a priori reasons to believe that standard assumptions such as isotropy are not valid; indeed it is not obvious that any of the variables even satisfy the weaker assumption of spatial stationarity. Thus we must find a method of analysis that can deal with the large size of the problem without imposing on the data any assumptions of spatial stationarity.

A.2 Estimating Missing Values

To "complete" the missing data, we elected to construct local multivariate time series models in each area, simultaneously modelling the five variables of interest. We will use the following notation:

$Y(t) = (Y_1(t), Y_2(t))^T$ = column vector of yellowfin and skipjack CPUE in time period t.

$X(t) = (X_1(t), X_2(t), X_3(t))^T$ = column vector of SST, N–S, and E–W components of wind velocity.

where ^{T} denotes matrix transpose. The model estimated in each area is a second-order multivariate autoregressive model (AR(2)) of the form:

$$\begin{pmatrix} Y(t) \\ X(t) \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} Y(t-1) \\ X(t-1) \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} Y(t-2) \\ X(t-2) \end{pmatrix} + E_t$$

$$\begin{pmatrix} U(t) \\ V(t) \end{pmatrix} = \begin{pmatrix} H_1(t) \\ H_2(t) \end{pmatrix} \begin{pmatrix} Y(t) \\ X(t) \end{pmatrix} + W_t \qquad (1)$$

where A and B are 5×5 matrices of parameters to be estimated, E_t is a 5-vector of independent Gaussian random variables with mean mu and covariance matrix Σ, W_t is the vector of observation noise, a random gaussian vector of mean zero and covariance matrix R, H(t) is a matrix of known parameters, and U(t) and W(t) are the noisily observed data series. The matrix H(t) can be used to signify that a series has a missing data point in time period t.

There are several reasons for choosing a model of the form (1). First, as most of our analysis will consist of spectral methods, the major goal in completing the time series is to cause minimal distortion of the covariance structure of the observed data. AR(2) models offer a flexible class of models that can closely approximate many different types of spectra. Typical AR(2) spectrums exhibit a small peak in the low frequency band. As most of our data are dominated by a seasonal cycle with a period of twenty-four fortnights, we would expect the actual spectrum to have a similar peak. Thus the data estimated from the AR(2) models should provide reasonable spectral estimates, but somewhat smoother (particularly at higher frequencies) than if the real data were available.

Ideally, both for more accurate estimates and for better forecasting models, some rigorous order determination method would be used to select the "best" AR model for each area, and all nonsignificant parameters would be set to zero. This approach was precluded by the already considerable computational burden (fifty iterations of the estimation routine require one hour of CPU time on a CDC 6500) and the fact that the information matrix, necessary for assessing the precision of the parameter estimates, can only be estimated numerically and requires substantially more computational time. Thus the AR(2) models would seem to give a reasonable tradeoff between computational burden and distortion of the covariance structure of the data.

Major reasons for choosing AR models are: (1) they are biologically reasonable, (2) they can be interpreted easily, and (3) there exists an algorithm (Shumway and Stoffer 1982) for exact maximum likelihood estimation of AR models when there are missing data, using Kalman filtering and the E–M

algorithm (Dempster et al. 1977). This is an iterative algorithm that starts with initial guesses for the parameters, calculates minimum mean square error estimates of the missing data, and then calculates maximum likelihood estimates of the parameters using the completed data set. The next iteration then begins by computing new minimum mean square error estimates for the missing data, using the latest parameter estimates. The algorithm converges when the completed data set produces maximum likelihood estimates of the parameters sufficiently close to those found in the previous iteration. Thus at convergence, the completed data sets yield maximum likelihood estimates of the particular parameters, and the estimated parameters give minimum mean square error of the missing data. Details of the algorithm can be found in the paper by Shumway and Stoffer (1982).

Models that also include information from neighboring areas probably would produce more accurate estimates of the missing data. We did not choose to do this for two reasons. First, the resulting models would have a large number of parameters to estimate compared with the number of time periods in our data. Second, including neighboring areas in the fill-in models could bias our findings on space-time behavior. By restricting ourselves to local models to fill-in the missing data, we feel more confident that any spatial relationships found are not artifacts. *A posteriori*, this appears to have been a wise decision, as the local models have many desirable properties and capture many known qualitative features of the eleven areas. Hence, it is doubtful that there was much to be gained from estimating the missing data using much larger models.

Spectral density matrices were calculated for the completed time series for frequency bands centered at intervals of 0.02273 cycles per fortnight. The spectral densities were estimated by Fourier transformation of the completed data series and by direct smoothing of the periodogram, as in Brillinger (1981, chapter 7). A bandwidth of 0.02273 gives thirteen degrees of freedom. While a larger bandwidth would give more precise estimates, it would no longer be possible to resolve interannual frequencies. Also, graphs of some of the series suggest there may be cyclic behavior at 2⅓ to 3 years. However, smaller bandwidths were not used to resolve these frequencies because the resulting degrees of freedom were too few. The spectral density matrix for the CPUE at frequency will be denoted by $fyy(\lambda)$ (or fyy where the frequency is either obvious or irrelevant) and the spectral density for the environmental series by $fxx(\lambda)$ (or) (fxx). The cross-spectral densities are denoted by $fyx(\lambda)$ (or) (fyx) and $fxy(\lambda)$ (or) (fxy) where fyx = fxy* and * denotes conjugate transpose.

A.3 Analyzing a Single Variable in Time and Space

In order to be able to understand and interpret the space-time variations of two variables together, it is necessary to understand the dominant mode of variation in each series separately. This gives us a clearer picture of the dynamics in each variable, and how they change when analyzed simultaneously with a second variable.

Suppose we have an r-vectored valued second-order stationary series X(t) with mean zero and covariance function $Cxx(\mu)$. Further, suppose we wish to find a q-vector valued series Y(t) with $q<r$ such that:

$$Y(t) = \sum_{\mu} b(\mu)\, X(t-\mu) \tag{2}$$

i.e., Y(t) is a filtered version of X(t) such that Y(t) is close to X(t) in the sense that we minimize:

$$E\left\{[X(t) - \sum_{\mu} c(t-\mu)\, Y(t)]^{T}[X(t) - \sum_{u} c(t-\mu)\, Y(t)]\right\} \tag{3}$$

where $c(\mu)$ is the inverse filter.

It is proven in Brillinger (1981, chapter 9) that the solution to this problem is given by the filter:

$$\begin{aligned} b(u) &= 2\pi \int_{-\pi}^{\pi} B(\alpha) \exp(i\alpha u) d\alpha \\ c(u) &= 2\pi \int_{-\pi}^{\pi} D(\alpha) \exp(i\alpha u) d\alpha \end{aligned} \tag{4}$$

where,

$$B(\lambda) = \begin{bmatrix} V_1^* \\ \cdot \\ \cdot \\ \cdot \\ V_q^* \end{bmatrix} \quad \text{and } D(\lambda) = B(\lambda)^*$$

and the vector $V_k(\lambda)$ is the k-th eigenvector of the matrix $fxx(\lambda)$ associated with the eigenvalue $m_k(\lambda)$. In practical terms, the solution to (3) is found by calculating the principal components of the complex matrices $fxx(\lambda)$ at all frequencies. It is not necessary to calculate the inverse transforms (4), as it is possible to calculate the gain, coherence, and phase between the original series and each principal component series (Brillinger 1981; Michaelson 1982a). At any given frequency, the coherence tells us how much of the spectral variance of the original series at this frequency is contained in the component series, and as the component series can be arbitrarily set to have zero phase with any area, they act as fixed reference points to determine the time relationships between the

part of each original series contained in the component. Principal components in the frequency domain have been used extensively by Michaelson (1982a, 1982b) and earlier by Wallace (1972) and Wallace and Dickinson (1972).

Principal component analysis has been used extensively in oceanography under the name "Empirical Orthogonal Function" (EOF) analysis. The difference between EOF's and the present analysis is that, in comparison to (2), EOF's are restricted to a transform of the form $b_0X(t)$. Equivalently, if we were to smooth our spectral estimates over all frequencies, then the two analyses would be equivalent (Brillinger 1981, chapter 9). Also, principal components analysis in the frequency domain produces q series that are independent at all frequencies, and hence at all lags, which is not true of EOF's. Further, the phase relationships at each frequency are readily available from this method, while EOF's would require extensive further calculations to obtain any kind of comparable results.

A.4 Analyzing Two Variables in Time and Space

At first glance, a method for dealing with two sets of variables would be to extend the previous analysis to the matrix

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

However, this is not appropriate for several reasons. First, principal component analysis is scale sensitive, and the two sets of variables may be measured in units that greatly differ in magnitude. Second, if we were to look at the linear combination of the first set of variables on the dominant component and the linear combination of the second set of variables also on the dominant component, even though this component contains the greatest total variance, it is not necessarily true that these two linear combinations will be very coherent. Since in dealing with two sets of variables we are interested in discovering related dynamics, rather than pulling out the pieces that contain the most variance, we would rather pull out the pieces that are the most coherent.

More formally, suppose we want to find the series

$$\zeta_j(t) = \int_0^{2\pi} A_j^*(\alpha) \exp(i\alpha t) dZ_x(\alpha)$$

and,

$$\eta_j(t) = \int_0^{2\pi} B_j^*(\alpha) \exp(i\alpha t) dZ_y(\alpha)$$

(with the standardizations

$$A_j^T(\lambda)\, A_j^*(\lambda) = 1,\ B_j^T(\lambda)\, B_j^*(\lambda) = 1$$

that have maximum coherence compared to any other such series that are independent at all frequencies to the series ζk, and ηk, with $k<j$. This problem is the standard canonical correlation problem, but now in the frequency domain. The solution to this problem (Brillinger 1981, chapter 10) is found by solving, at each frequency, the pair of eigenvalue problems:

$$f_{xx}^{-1}(\lambda)\, f_{xy}(\lambda)\, f_{yy}^{-1}(\lambda)\, f_{yx}(\lambda)\, A_j(\lambda) = \mu_j\, A_j(\lambda)$$

$$f_{yy}^{-1}(\lambda)\, f_{yx}(\lambda)\, f_{xx}^{-1}(\lambda)\, f_{xy}(\lambda)\, B_j(\lambda) = \mu_j\, A_j(\lambda)$$

The resulting series, $\zeta j(t)$ and $\eta j(t)$, are called the jth canonical series, and their coherence, $u_j(\lambda)$ is called the jth canonical coherence.

It is computationally more stable to solve:

$$f_{yy}^{-1/2}(\lambda)\, f_{yx}(\lambda)\, f_{xx}^{-1}(\lambda)\, f_{xy}(\lambda)\, f_{yy}^{-1/2}(\lambda\ V_j(\lambda) = \mu_j\, V_j(\lambda)$$

Then $B_j(\lambda)$ is proportional to $fyy^{-1/2}\, V_j(\lambda)$ and $A_j(\lambda)$ is proportional to $fxx^{-1}(\lambda)\, fxy(\lambda)\, fyy^{-1/2}(\lambda)V_j(\lambda)$. It follows from Brillinger (1981, chapter 10) that the jth canonical series has a spectral density at frequency:

$$A_j^*(\lambda)\, f_{xx}(\lambda)\, A_j(\lambda)\ [\text{resp. } B_j^*(\lambda)\, f_{yy}(\lambda)\, B_j(\lambda)]$$

and the cross-spectrum between the jth canonical series and the original series given by:

$$f_{xx}(\lambda)\, A_j(\lambda) \text{ and } f_{yy}(\lambda)\, B_j(\lambda)$$

From these relationships we can calculate the gain, coherence, and phase between the original series and each of the canonical series. The phase relationships are calculated so that the jth canonical series for each variable has phase zero in area 4. Thus the phases as presented have meaning only in terms of the same variable at a different location. For example, because in area 9 yellowfin CPUE may have a phase of 7.84 fortnights and SST a phase of −8.5 fortnights, it does not mean that SST has preceded CPUE by 16.34 fortnights. Rather, for example, if the CPUE canonical series has a phase of −8.00 with the SST canonical series, then SST would precede CPUE by only 8.3 fortnights. Thus, results from the canonical analysis must be interpreted with care.

When actually calculating the canonical analysis, we found that the matrix:

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

was never of full rank (i.e. 22); in fact it usually had rank of only 10 or 11. As the analysis is based only

on the submatrices of this matrix, we overcame this problem by calculating the singular value decompositions of fxx and fyy:

$$f_{xx} = Q_{ox} D_x Q_{ox}^*; f_{yy} = Q_{oy} D_y Q_{oy}^*;$$

where the Q's are orthogonal matrices and the D's are diagonal matrices. Then the six smallest diagonal elements of Dx and Dy were forced to zero, and the matrices fxx and fyy were reconstructed. This removed the singularity in each case by eliminating from each set of variables the six linear combinations of the area spectral density at that frequency. Usually only 5% or less of the total spectral power was removed, so that very little information was lost by this process.

The bulk of our analysis depends on spectral and other least squares techniques. These techniques generally "smooth" or average over different occurrences, so that the results display the "normal" pattern of relationships. However, for many fisheries problems, particularly in relation to the environment, the abnormal happening may be of equal interest (such as El Niño in the Pacific). Further refinements of our techniques are needed to examine this type of situation.