The Development and the Use of a Climatic Database for CEOS using the COADS Dataset

CLAUDE ROY*

ROY MENDELSSOHN**

* Sea Fisheries Research Institute
Private Bag X2
Rogge Bay 8012
Cape Town
SOUTH AFRICA

** Pacific Fisheries Environmental Group
(PFEG)
1352 Lighthouse Avenue
Pacific Grove
CA 93950-2097
USA

ABSTRACT

The Comprehensive Ocean-Atmosphere Dataset (COADS) was selected by CEOS for use in analyzing the climatic variability of the world upwelling ecosystems during the past four decades. The COADS database summarizes over 100 million surface meteorological observations collected by ships of opportunity and other platforms over the world. This dataset has world-wide coverage and the earliest data dates back to 1854.

A preliminary investigation of the climate variability of the world coastal upwelling regions was performed using a reduced version of the COADS dataset (the 2° by 2° monthly summary files). It appeared that these files are subject to numerous biases and are not suitable for performing the retrospective analysis of the climatic variability planned by CEOS. It was necessary to set up a database using the individual observations instead of the summary files. A version (CMR-5) of the 100 million individual records available in the COADS was provided by NCAR (National Centre for Atmospheric Research - USA). The CMR-5 version of the COADS dataset was reorganized in a fashion that allowed for rapid access to the dataset using a micro-computer, and software for processing and summarizing the data was developed. The reorganized dataset was then put onto CD-Roms for distribution. The software and the five CD-Roms represent the core of the climatic information used by the CEOS program for the retrospective and comparative analyses.

Some important biases encountered when using the COADS data are reviewed. Changes through time in the measurement procedures are the most common source of systematic errors in COADS; these biases occur in particular in sea surface temperature and wind. The sudden change that occurs after the Second World War due to the use of insulated buckets and of engine-intake measurements is thought to be responsible for the abrupt change in the SST time series that occurs at the same time. Wind data reported by ships are either measured with an anemometer or are estimated from sea-state. Estimated wind data predominate before the Second World War. Today, wind data collected using measurement devices are predominant. Several examples are given in order to illustrate the potential biases that can affect the seasonal cycle and the long term trend of the wind intensity due to the gradual shift through time from estimated wind data to measured wind data. The potential biases introduced by merging buoys data with ship data are illustrated by looking at the wind data off California during the last twenty years. Finally, a discussion on the reality of the positive trend existing in the wind data off West Africa is presented.

RÉSUMÉ

La base de données COADS (Comprehensive Ocean-Atmosphere Dataset) a été sélectionnée par CEOS pour analyser la variabilité climatique des zones d'upwelling mondiales au cours des quarante dernières années. Cette base de données rassemble 100 millions d'observations météorologiques de surface collectées par les navires marchands et d'autres plates-formes sur les océans du monde entier. La couverture est mondiale et les premières données remontent à 1854.

Un premier travail sur la variabilité climatique fut réalisé en utilisant une version réduite de COADS (moyennes mensuelles par carrés de 2°par 2°). Il est apparu que les données contenues dans ces fichiers sont affectées par de nombreux biais qui rendent hasardeuse leur utilisation pour l'étude de la variabilité climatique sur le long terme. Pour pallier ces difficultés, il fut nécessaire de bâtir une base de données rassemblant les données originales au lieu de données déjà pré-traitées. Une version de la base de données COADS (CMR-5) contenant les 100 millions d'enregistrements originaux fut acquise auprès du NCAR (National Centre for Atmospheric Research - USA). Cette base de données fut restructurée afin de permettre un accès rapide à partir d'un micro-ordinateur, et un logiciel d'extraction et de traitement fut développé. L'ensemble de la base de données fut ensuite transféré sur un jeu de 5 CD-Roms.

Les principaux biais rencontrés lors de l'utilisation de COADS sont passés en revue. Les changements au cours du temps des méthodes de mesure sont à l'origine des erreurs les plus courantes, les variables concernées sont plus particulièrement la température de surface de la mer et le vent. Des modifications des modes de prélèvement et l'introduction de nouveaux instruments de mesure à la fin de la Seconde Guerre Mondiale sont à l'origine d'un changement abrupt dans les séries de température de surface de la mer à la fin des années quarante. Les données concernant le vent, récoltées par les navires marchands, sont de deux types : "estimées" visuellement à partir de l'état de la mer ou "mesurées" à l'aide d'un anémomètre. Jusqu'aux années cinquante, l'essentiel des données de vent était du type "estimées", aujourd'hui les données "mesurées" à l'aide d'un anémomètre sont prédominantes. Plusieurs exemples sont donnés afin d'illustrer les biais introduits dans les séries temporelles de vent, construites à l'aide de COADS, par l'augmentation régulière du nombre de données "mesurées" au cours des quarante dernières années. Les biais introduits par la prise en compte de données provenant de bouées météorologiques sont illustrés à partir de séries extraites devant les côtes de Californie. Les séries temporelles de vent obtenues à partir de COADS devant les côtes ouest-africaines sont comparées à d'autres séries, et une discussion concernant la réalité de la tendance observée dans les données COADS dans cette région est réalisée.

INTRODUCTION

One of the major objectives of the CEOS program is to assemble, summarize and analyze the climatic data record of the four eastern ocean boundary upwelling ecosystems during the past four decades, as well as the record for the other upwelling areas (Bakun *et al.*, 1993). For this purpose, a climatic data base covering the last fifty years with a worldwide coverage is needed. This also should be homogeneous among areas and over time and thus allow the comparative analyses planned by the CEOS program.

The Comprehensive Ocean-Atmosphere Dataset (COADS) is one of the few datasets that meet these criteria (see Slutz *et al.* (1985) and Woodruff *et al.* (1987) for an extensive description of the source data and of the procedures used to create the COADS dataset). The COADS database summarizes over 100 million surface meteorological observations collected by ships of opportunity and other platforms over the world oceans which have been quality controlled and put into a consistent format. This dataset is the most complete record of surface marine climate to date. It has a world-wide coverage and the earliest records date back to 1854. The early studies of the CEOS program have used Release 1 of the COADS for the period 1854 to 1979; and for the period from 1980 to 1990, an interim release has been used, which is compatible in format and organization to Release 1, but is constructed using simplified procedures and preliminary input data.

Besides the individual observations, which are available in several formats containing different amounts of the data, reduced or summarized versions of the COADS dataset are available in several forms, including monthly or decadal summaries average over a spatial grid of 2° by 2° for the entire ocean. Due to the massive number of individual reports (100 million), the majority of researchers have used the monthly summaries. A preliminary investigation of the climate variability of the world coastal upwellings was performed using the 2° by 2° monthly summary files. Except for a trimming procedure which removes extreme outliers, the summary files are computed using all the available data. Biases in the long term behaviour of the time series derived from these pre-processed files rapidly arose: data from meteorological buoys are merged with ship data (see Section 3); wind data are aggregated without distinction between the measurement procedure used (estimated using the Beaufort scale or measured using an anemometer, see Section 3). The fixed 2° by 2° spatial grid is also a strong limitation for coastal studies where the spatial grid has to be flexible in order to be adapted to the shape of the coastline. After this first attempt, it appeared that the 2° by 2° degree summaries files were not suitable for performing the retrospective analysis of the climatic variability planned by CEOS. To avoid some of the bias existing in the summaries files, it was necessary that the retrospective analysis be performed using the individual observations instead of the 2° by 2° summary files.

Producing useful time series from the 100 million individual records available in COADS is a much more complicated task than working with the 2° by 2° monthly summaries files; this is even more so as the CEOS program needs to run on micro-computers rather than mainframes or workstations in order to allow dissemination of the database throughout the CEOS network. At the time the project started, large capacity optical storage devices became available which allowed for the storage and processing of large amounts of data with a micro-computer. A version of the 100 million individual records available in the COADS database which uses a compressed binary format (CMR-5) was provided by NCAR (National Centre for Atmospheric Research - USA) and transferred to a set of optical disks connected to a micro-computer. The CMR-5 version of the dataset was reorganized in a fashion that allowed for rapid access to the dataset by area, and software for processing and summarizing of the data was developed. The reorganized dataset was then put onto CD-Roms for distribution. The software and the five CD-Roms represent the core of the climatic information used by the CEOS program for the retrospective and comparative analyses.

1. DATA SOURCE USED FOR THE DEVELOPMENT OF THE CEOS CLIMATIC DATABASE

The primary data source for the CEOS climatic database is the Release 1 of COADS which covers the period 1854-1979. An interim data product is used to cover the period from 1980 to 1990 (Woodruff and Lubker, 1986). During the design and development phases of the database, our main goal was, despite the massive volume of information, to preserve access to the raw individual observations and to allow access with a microcomputer. The CMR-5 format provides a good compromise between keeping access to the raw data and a significant reduction of the total volume. It consists of approximately 100 million individual observations from 1854 to 1990. It is a packed binary format containing the most frequently used information, designed as a compact alternative to the original reports. The variables included in each report under the CMR-5 format are given in Table 1. The following observed quantities are available:

- air temperature;
- sea surface temperature;
- dew point depression;
- zonal and meridional wind components;
- sea level pressure;
- cloudiness;
- present weather.

Each record in the CMR-5 format contains data on the measurement procedure (fields BI, WI and HI in Table 1), on the precision and units (fields TI and DI in Table 1), on the type of observing vessel (field ST in Table 1) and on the origin of the record (field CD in Table 1). Quality control indicators (noted as 'flags' in Table 1) for each observed quantity area also available. The COADS documentation (Slutz *et al.*, 1985) gives a detailed description of the quality control procedures. A brief summary of these procedures is given below.

A multiple step statistical method was set up during the development of COADS in order to identify outliers for six variables: sea surface temperature, air temperature, sea level pressure, zonal wind, meridional wind and humidity. The result of this process is the definition of the smoothed lower and upper median deviation (s_1, s_2) around the smoothed mean (X) for each of the six variables; these calculations are performed with a monthly time step and a 2° spatial step. These means and limits are used to create trimming bounds for the variables. The value assigned to the quality control flags for a given observation is set according to Table 2.

The entire CMR-5 format version of the COADS dataset was reorganised and sorted in order to have an homogeneous structure over the whole time period (1854-1990). 648 folders (or subdirectories) were created, each one corresponding to a 10° by 10° geographical square. A folder (or subdirectory) contains 25 data files, one for each 2° geographical square within the 10° square. This structure allows quick access and retrieval of data for a specific ocean location.

A computer program, the Comprehensive Ocean Data Extraction or CODE was developed in order to quickly access, process and summarize using a microcomputer, the COADS dataset in the CMR-5 format. The advantage of CODE over the use of the 2° by 2° monthly pre-processed COADS files is that all calculations are performed by going back to the individual observations. This is particularly important for the calculation of derived parameters such as wind stress or wind speed cubed which are provided by CODE. It also allows a flexible spatial grid (0.1° of latitude or longitude step in the data, 0.5° in the CODE program), and makes it possible to select between several types of platform (buoys, merchant or research ship, ocean station , etc.) or between the methods used to collect the data (estimated wind using the Beaufort scale or measured wind using an anemometer).

2. Bias in the COADS data

Some important biases encountered when using the COADS data are reviewed. Changes through time in the measurement procedures are the most common source of systematic errors in COADS; these biases occur in particular in Sea Surface Temperature (SST) and wind. Air temperature and dew point temperature are also affected by systematic

Field	Description	Value	Units
	Location		
BOX10 *	10° box number	1≤ 648	1
MONTH		1≤ 12	1
BOX2 *	2° box number	1≤ 16202	1
YEAR		1800 ≤ 2054	1
DAY		1≤ 31	1
HOUR		1≤ 23	1
LAT	(from BOX2 SW corner)	0 ≤ 2.0	0.1°
LONG	*(idem)*	0 ≤ 2.0	0.1°
	Temperature		
SST	Sea Surface Temperature	-5.0 ≤ 40.0	0.1°C
BI	Bucket Indicator	0 ≤ 2	1
AT	Air Temperature	-88.0 ≤ 58.0	0.1°C
DP	Dew Point Depression	0 ≤ 70.0	0.1°C
TI	Temperature Indicator	0 ≤ 5	1
	Wind		
U	Eastward component	-102.2 ≤ 102.2	0.1 m/s
V	Northward component	-102.2 ≤ 102.2	0.1 m/s
DI	Direction indicator	0 ≤ 5	1
WI	Wind speed indicator	0 ≤ 1	1
	Pressure and clouds		
SLP	Sea Level Pressure	870.0 ≤ 1074.6	0.1 mb
C	Total cloud amount (**)	0 ≤ 9	1
NH	Lower cloud amount (**)	0 ≤ 9	1
CL	Low cloud type (**)	0 ≤ 10	1
H	Cloud height (**)	0 ≤ 10	1
HI	Cloud height indicator (**)	0 ≤ 1	1
CM	Middle cloud type (**)	0 ≤ 10	1
CH	High cloud type (**)	0 ≤ 10	1
	Miscel.		
ST	Ship type	0 ≤ 7	1
PW (**)	Present weather (**)	0 ≤ 99	1
CD (**)	card deck (**)	0 ≤ 999	1
	Flags		
LF	Landlocked flag	0 ≤ 0	1
SF	SST flag	0 ≤ 2	1
AF	Air temperature flag	0 ≤ 2	1
RF	Relative humidity flag	0 ≤ 2	1
WF	Wind flag	0 ≤ 2	1
PF	Pressure flag	0 ≤ 2	1

Table 1: List of all variables available under the CMR-5 format.
(*) : this refers to the NCAR 10° and 2° box coordinate systems (see Slutz *et al.*, 1985) for detail.
(**) : parameters cannot be extracted using the CODE program.

Flag value	Trimming limits
0	$X - 2.8s_1 \leq x \geq X + 2.8s_2$
1	$X - 3.5s_1 \leq x \geq X + 3.5s_2$
2	$x < X - 3.5\,s_1$ or $x > X + 3.5s_2$
missing	missing

Table 2: Quality control flags values in the CMR-5 format; x is the value of an individual observation, X is the smoothed median and s_1 and s_2 are the smoothed lower and upper median deviation.

errors but they are not documented here (see Isemer and Hasse, 1987; Folland *et al.*, 1984; Kent *et al.*, 1983a and 1983b for details). Along the US coastline, data coming from moored meteorological buoys have been incorporated into the COADS database during the 1980s; the effect of such changes is documented in an area where buoy data have become predominant after 1980. Comparisons between atmospheric pressure data from COADS and several other independent atmospheric pressure data have been made; the results are generally good and yield some confidence in the usefulness of the COADS pressure data (Jones *et al.*, 1986; Ward and Hoskins, 1996).

2.1. Biases in SST

The method of collecting SST samples has changed through time. Before the Second World War, SST was measured on a sea water sample collected using a bucket; engine intake measurements became predominant for the later period (Jones *et al.*, 1986). Due to evaporative cooling, measurements performed using an uninsulated bucket is thought to be about 0.5°C lower than a measurement from an engine intake (Ramage, 1984). The sudden change that occurs after the Second World War due to the use of insulated buckets and of engine-intake measurements is thought to be responsible for the abrupt change in the SST time series that occurs at the same time (Folland and Parker, 1990). This sudden shift of the mean value of SST is illustrated in Figure 1.

Several models has been developed to correct bias in bucket observations prior to the Second World War (Folland *et al.*, 1984; Folland and Hsiung, 1987; Folland and Parker, 1990; Jones and Wigley, 1992), however these corrections cannot be applied to individual observation because of the unreliability of the bucket/intake indicator and of the lack of information on the kind of bucket used in COADS.

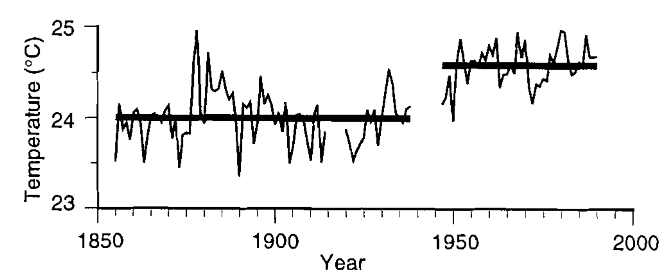

Fig. 1: Time series of Sea Surface Temperature in the tropical Atlantic (10°N-30°N, 30°W-50°W). A shift in the mean value of SST occurred in the late 1940s.

Since the beginning of the 1950s, insulated bucket or engine intake measurements are thought to be the predominant method used for measuring SST. Thus, even if some bucket measurements have been used after 1945, the biases are probably considerably less than those occurring before.

2.2. Biases in the wind data

Wind data reported by ships are either measured with an anemometer or are estimated from sea-state. Estimated wind data predominate before the Second World War. The percentage of measured wind data started to increase in the late 1940s (Fig. 2). Today, wind data collected using measurement devices is the predominant way of collecting wind data (Fig. 2). The CMR-5 format contains a flag for each wind record which indicates whether the data were estimated or measured. Several extractions were made in order to compare the characteristics of the wind signals calculated using either estimated, measured or the combination of estimated and measured wind data in several areas. These comparisons illustrate the potential biases that can affect the long term trend of the wind intensity due to the gradual shift through time from estimated wind data to measured wind data.

Measured wind data are measured in meters per second. Estimated wind data are reported using the Beaufort scale and have to be converted to an equivalent wind speed. Conversion from the Beaufort scale to wind speed expressed in m/s is done using a Beaufort equivalent scale. The scale used in COADS is the CODE 1100 (or old WMO scale). This scale is known to underestimate wind speed for Beaufort number less than 6 and to overestimate wind speed for Beaufort number greater than about 6 (WMO, 1970; Cardone, 1969; Kaufeld, 1981; Isemer and Hasse, 1991). The calculation of wind speed using measured wind data will therefore give different results than the same calculation performed using estimated wind data. Examples of the differences obtained in the determination of the mean monthly seasonal cycle of the wind speed for several areas located in different upwelling ecosystems are given in Figure 3. From those examples, it appears that the mean monthly values of wind speed calculated using measured wind data are from 0 to 0.8 m/s higher than wind speed calculated using estimated wind; the phase of the seasonal cycle does not appear to be strongly affected.

When both estimated and measured wind data are merged to calculate wind speed, the variation through time of the ratio between the numbers of estimated and measured wind observations introduces an artificial variability in the time series of wind intensity. Because estimated wind data tend to be underestimated for Beaufort scale lower than 6, wind speed will tend to be

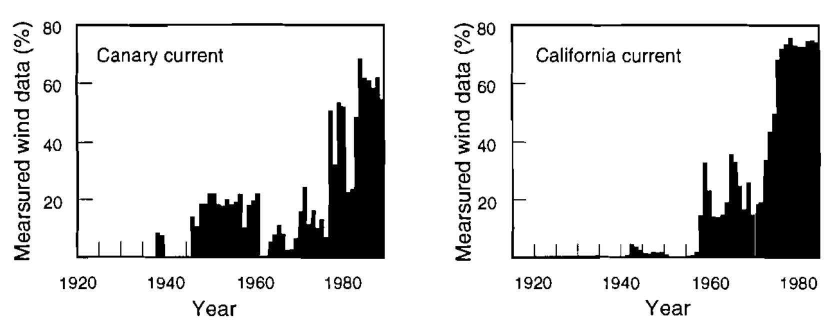

Fig. 2: Percentage of measured wind data in COADS from 1920 to 1990 in two eastern boundary current ecosystems.

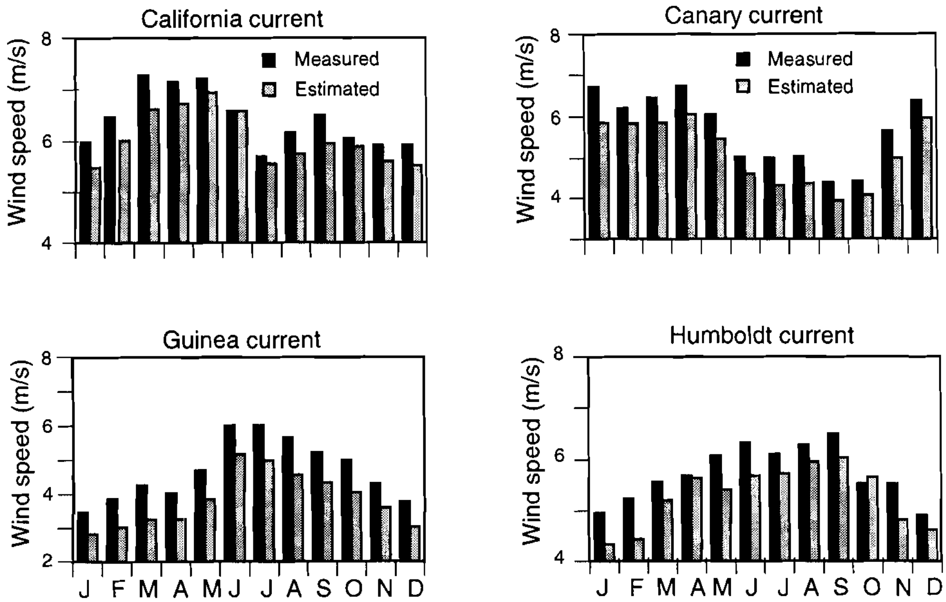

Fig. 3: Mean monthly seasonal cycle of the wind speed calculated in four areas located in the major eastern boundary current ecosystems using estimated and measured wind data (calculated for the period 1950-1990).

lower during time periods characterized by a high number of estimated wind data than during time periods when measured wind data are predominant (all other things remaining constant). The time series presented in Figure 4 illustrate the effect of the increasing percentage of measured data wind during the last 40 years on the long term variability of the wind speed. The percentage of measured wind data starts to increase significantly around the late 1960s. As a consequence, the mean annual wind speed calculated using all the available wind data (both estimated and measured, E&M time series) becomes significantly higher than the mean annual wind speed calculated using only estimated wind data (E time series).

Figure 5 shows that the difference between the annual mean of the E&M and the E wind time series is mostly explained by the interannual variability of the percentage of measured wind data. The difference reaches 0.5 m/s in the California Current and in the Guinea Current when the measured wind comprises 70% of the E&M time series.

The consequence of the increasing percentage of measured wind data in the E&M time series is the introduction of an artificial positive trend superimposed over the long term tendency given by the E wind time-series (Fig. 4). This artificial increase of the wind intensity is particularly noticeable in the California Current and the Guinea Current regions. When a linear trend is fitted from 1970 to 1990 to the time series of the California and Guinea Current wind speed, the slope of the trend of the E&M wind speed time series is 60%, respectively 80% higher than the slope of the trend given by the E wind speed time series.

A way to avoid such biases would be to use measured wind data only, but for many areas, the data density of the measured wind data in COADS is not high enough to allow the calculation of a reliable monthly time series of measured wind data before the early 1970s. Time series of wind speed based on estimated wind data can go back in some areas as far as the late nineteenth century, but the mean value of the wind speed will be affected by the bias introduce with the use of the CODE 1100 conversion table. Changes in the observation practices that were used to estimated the wind from sea-state,

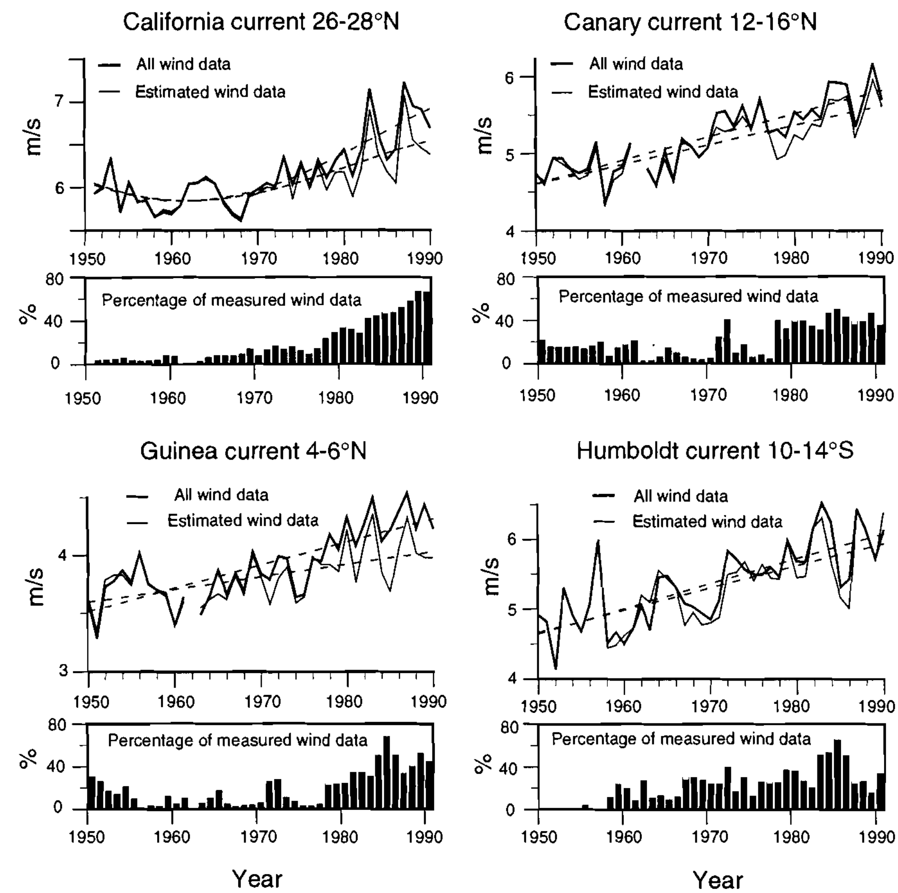

Fig. 4: Time series of wind speed calculated using all the available wind data (E&M time series) and using only estimated wind data (E time series) in four upwelling areas. For each region, the annual percentage of measured wind data is given. The long-term trend of the wind speed for each time series is also presented.

the increase of the mean size of the ships over the century and many other factors may also have introduced biases in the long term behaviour of estimated wind time series.

2.3. Meteorological buoys off the US coast

The operation of a network of moored buoys by the US National Data Buoy Centre (NDBC) started in the early 1970s. Hourly buoys data have been incorporated into the COADS interim release for the period 1980-1990 (Woodruff and Lubker, 1986). Some of the biases introduced by merging buoy and ship data are documented here using data extracted off the California coast.

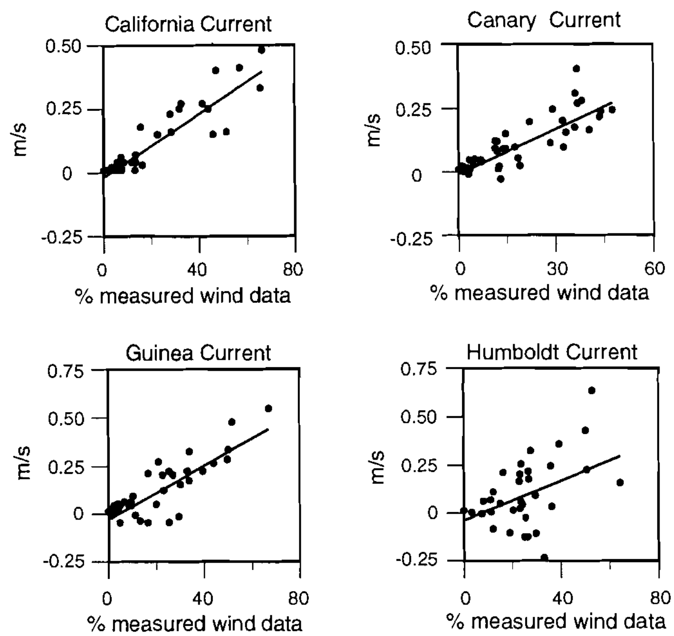

Fig. 5: Plots of the difference between the E&M and M wind speed time series against the percentage of measured wind data in four different upwelling areas.

Several buoys were installed off the California coast during the 1980s. Hourly SST, wind and atmospheric pressure data from these buoys have been incorporated into COADS. The example given in Figure 6 shows that between 36°N and 40°N the number of wind observations stays relatively constant during the 1970s (between 3000 and 4000 observations per year) and that it suddenly increases by a factor of 5 in 1981 to reach 22 920 observations. This sudden increase of the number of observations in 1981 is the consequence of the introduction of the buoy data into COADS. The reason for the decrease of the number of observations after 1984 is unknown to us.

The buoys are located in the nearshore area and the ship traffic lines are located several miles offshore. Therefore, we can suspect that the dominance of data coming from the buoys during the 1980s introduces important changes in the value of the surface atmospheric parameters in the time series. Unfortunately, the ship type indicator is missing for the interim release and therefore it is impossible to discriminate between buoy and ship observations from 1980 to 1990. The buoys are equipped with an anemometer and the buoy wind data were recorded in COADS as 'measured' data. In the COADS interim release, this is the only way to investigate the change introduced by the buoys data into COADS. The time series of the scalar wind speed calculated using 'measured' wind data in an area located off the California coast shows that a sharp decrease of the scalar wind speed occurred in 1981 when the data from the buoys are incorporated (Fig. 7). This apparent relaxation of the wind speed is likely the result of the coastal location of the buoys where the wind field presents a local minima (when compared to the offshore domain).

The time series of the scalar wind speed calculated using both measured and estimated wind is also affected by the implementation of the buoys (Fig. 7). In this series, the dominance of data coming from the buoys after 1980 results in a spurious relaxation of the wind. When wind variables are calculated using 'estimated' wind data, the buoy data are not taken into account. The time series of scalar wind speed calculated using 'estimated' wind data shows that the 1981 wind speed relaxation is an artifact resulting from the dominance of buoy data after 1980 (Fig. 7).

The COADS monthly 2°x2° summary files do not discriminate between either ship or buoy data and it is likely that wind time series off California extracted from these files will be affected by the spurious relaxation of the wind after 1980.

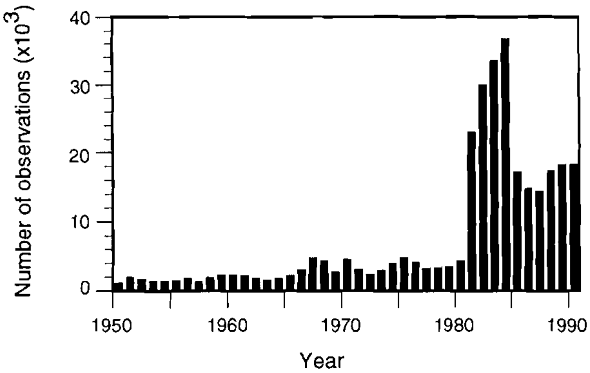

Fig. 6: Annual number of wind data available in the COADS database along the California coast between 36°N and 40°N.

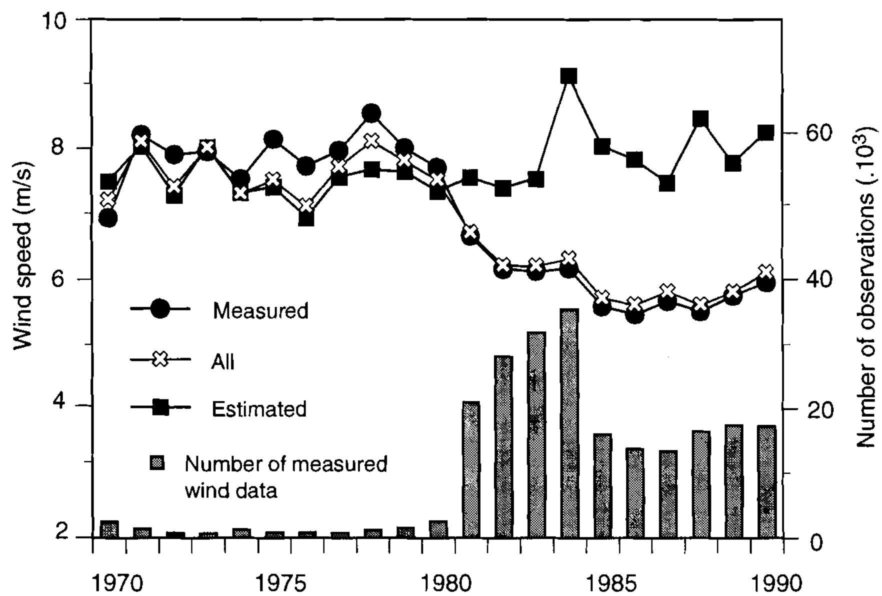

Fig. 7: Scalar wind speed along the California coast between 36°N and 40°N calculated using 1) measured wind data, 2) all wind data (estimated and measured) and 3) estimated wind data. The annual number of measured wind data in the area is also presented.

3. Trend in the wind off West Africa: reality or artifact?

A dominant feature of the time series of the COADS wind data off West Africa since the late 1940s is a continuous positive trend. Between 12°N and 22°N, wind stress (proportional to the square of the wind) increased by almost 50% from 1946 to 1990. During the upwelling season, it is expected that such an important increase of the wind stress would intensify the upwelling process and the upward flux of cold subsurface water to the surface, resulting in colder SST. Off West Africa, the magnitude of the wind stress increase from 1950 to 1990 is significant enough that drastic changes would have been introduced in the physical and biological properties of the region. However, the validity of the trend in the COADS wind data has been the subject of intense debates (Ramage, 1987; Cardone *et al.*, 1990; Isemer, 1995). In the following paragraphs, an attempt is made to examine the validity of the trend in the COADS wind off West Africa, between 12°N and 24°N.

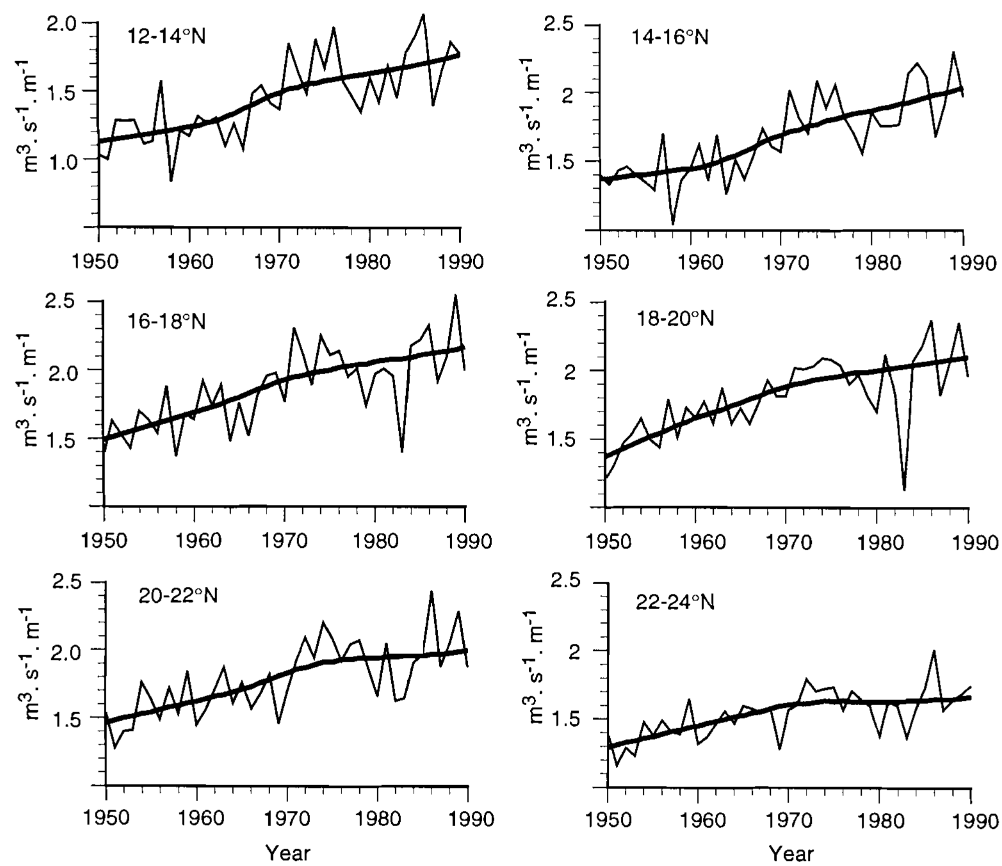

Fig. 8a: Time series of the Coastal Upwelling Index (CUI) averaged over the upwelling season (January through June) from 1950 to 1990 along the coast of West Africa between 12°N and 24°N.

Monthly time series of wind stress using the estimated wind data and of SST are extracted from the COADS database along the coast of West Africa, from 12°N to 24°N with a 2° latitudinal step. A Coastal Upwelling Indices (CUI) is calculated using the wind stress data following Bakun (1973). CUI is the offshore component of the total wind induced Ekman transport. With an upwelling favourable wind, the Ekman transport is directed toward the offshore direction and CUI is positive.

From 12°N up to 24°N, monthly CUI and SST are averaged over the upwelling season (Fig. 8a and 8b). In each area, the long term behaviour of the CUI time series is characterized by a continuous positive trend (Fig. 8a). The long term trends of the SST time series are rather stable over the whole time period with a slight cooling in the 1970s following by a warming in the 1980s (Fig. 8b). This suggests that there is no apparent relationship between the long term variability of the upwelling process and SST.

A comparison of the COADS data with other information would be extremely valuable but very few wind data exist with such a long term coverage. The wind data routinely collected every 3 hours at the Dakar-Yoff airport (14°40' N) is one of the few time series available in the region to perform such a comparison. The airport is located on the tip of the Cap-Vert peninsula

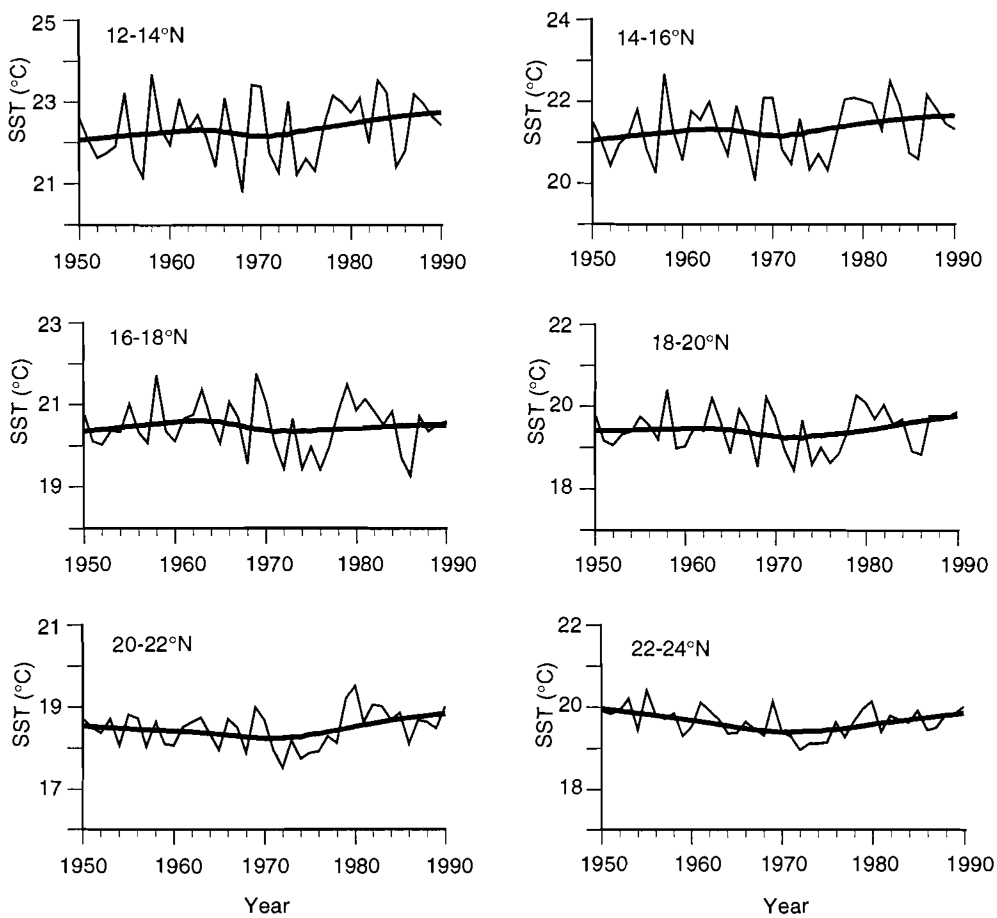

Fig. 8b: Time series of the Sea Surface Temperature (SST) averaged over the upwelling season (January through June) from 1950 to 1990 along the coast of West Africa between 12°N and 24°N.

and the wind data are thought to be representative of the wind over the offshore domain (Roy, 1989). The annual values during the upwelling seasons and the trend of the CUI derived from the Dakar-Yoff airport wind data from 1964 to 1990 are compared with the CUI time series derived from the COADS estimated wind data in the same region (14°N-16°N) (Fig. 9). The interannual variability of both time series are quite similar but there is no apparent positive trend in the CUI derived from the Dakar-Yoff wind data (Fig. 9); during the same time period, the trend in the CUI derived from the COADS wind data shows an increase of about 50% (from 1.5 to 2 $m^3.s^{-1}.m^{-1}$) of the upwelling intensity (Fig. 9). This comparison between the COADS data and the airport data gives two different pictures of the long term variability of the wind in the region.

Previous studies of the interannual variability of wind and SST in the region had shown that variability of the upwelling favourable wind accounted for a significant part of the variability of SST (Arfi, 1985; Portolano, 1986; Roy, 1989; Nikjaer and Van Camp, 1994). As an intensification of the upwelling favourable wind results in an increase of the CUI and in an intensification of the upward flux of cold subsurface water, a negative correlation is expected between CUI and SST. With the COADS data, the correlation coefficients between the mean values of CUI and SST during the upwelling seasons from 1950 to 1990 are low (between -0.22 and -0.4 except for the northern most area where it reaches 0.58) (Fig. 10). This

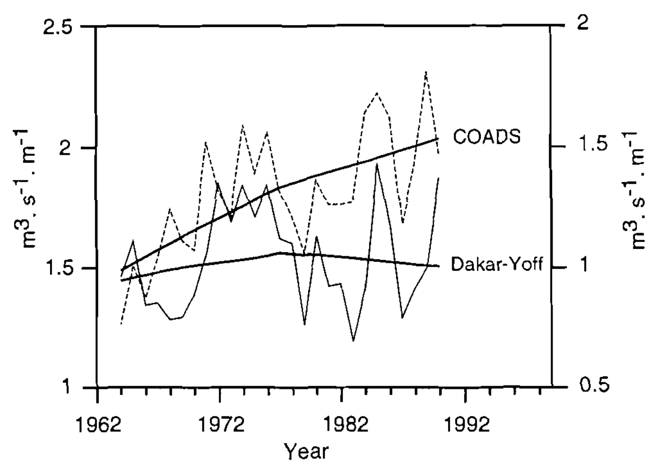

Fig. 9: Coastal Upwelling Index (CUI) averaged over the upwelling season (January through June) from 1964 to 1990 calculated using a) the wind data at the Dakar Yoff airport and b) the COADS wind data in the corresponding square (14°N-16°N, 19°W-16°W).

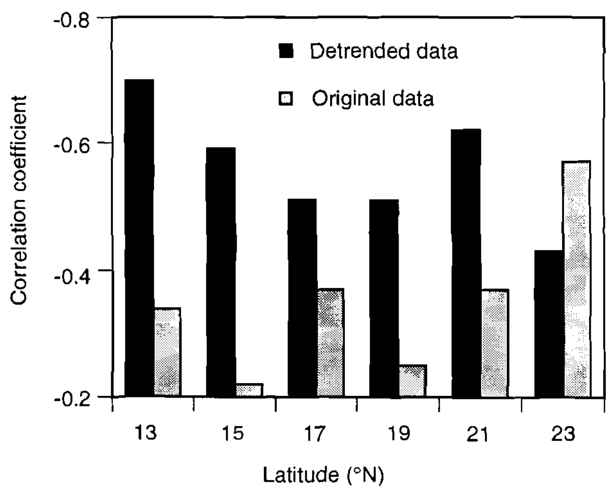

Fig. 10: Correlation coefficient between CUI and SST during the upwelling season (January through June) along the coast of West Africa between 12°N and 14°N using a) the original CUI data and b) the detrended CUI data.

suggests that the increase of the upwelling favourable wind in the COADS data over the whole time period has had little effect on the interannual variability off SST. Given the magnitude of the wind intensification (50% over 40 years) and previous studies, this is a surprising result that raises some doubt about the reality of the trend in the COADS wind data off West Africa. The use of a detrended CUI time series gives a totally different picture. The correlation coefficients between the detrended CUI and SST time series varies between -0.48 and -0.70 (Fig. 10). Except for the northern-most area, using the detrended CUI time series results in a significant increase of the correlation between the strength of the upwelling and SST.

To obtain a significant correlation between the fluctuations of upwelling and SST is in accordance with our knowledge of the dynamics of the region and with previous studies, which have shown similar results. This questions the validity of the trend in the COADS wind data off West Africa. The comparison between the wind data from the Dakar-Yoff airport and the COADS wind data also raises some doubt about the validity of this trend. If this trend is not a real phenomenon, its origin remains unknown to us. The wind data used to calculate the CUI are restricted to the estimated wind data and the origin of the trend cannot be explained by an increasing fraction of anemometer type data in the wind time series used to calculate the CUI.

CONCLUSION

The COADS database represents a very important source of information for studying the long term variability of the climate over the oceans. In many areas like in the tropics, the COADS data are the only information available regarding the variability of the environment. For fishery oceanography, it represents a very valuable tool. However this dataset is not exempt of important limitations. Time series of surface meteorological parameters constructed using the COADS data can be strongly affected by several biases and lead to erroneous interpretation.

For the wind variables, the main difficulty results from the difference between estimated wind data using the Beaufort scale and measured wind data using an anemometer. It is shown that the calculation of wind speed using the COADS dataset can be strongly biased depending on the kind of data used. Wind speeds calculated using estimated wind data will tend to be underestimated. The calculation of the mean monthly cycle of the wind speed in several areas shows that difference between estimated and measured values can be as high as 0.8 m/s but the phase of the signal does not seem to be strongly affected.

The consequence of the difference between estimated and measured wind data is that time series of wind speed can be severely biased by the progressive shift from estimated wind data to measured wind data. When both measured and estimated wind data are merged, the resulting time series is affected by an artificial positive trend. In some areas, the slope of the trend in the wind speed time series calculated using both measured and estimated wind data can be 60% to 80% higher than the slope of the trend calculated on a wind speed time series using estimated data only. Recently, Ward and Hoskins (1996) made a comparison between the reported winds in COADS (from the COADS 2°x2° monthly summaries files) and a wind derived from seasonal mean sea level pressure gradients over the world oceans. Their results show a considerable disagreement between the long term trend in the reported wind and the trend in the pressure derived wind. Globally averaged over the world oceans, there is no trend of any substance in the pressure derived wind, whereas there is an upward trend in the reported wind data of about 14% from 1949 to 1988. They conclude that the difference between the reported wind and the pressure derived wind was mainly due to the growing percentage, over time, of measured wind data in the COADS records.

For the calculation of a reliable wind time series over a long time period, it is strongly recommend to use either estimated or measured wind data and to avoid the use of time series where both wind data type are merged. This represents a strong limitation for the use of the COADS 2°x2° monthly summaries files to study long term climatic variability. In those summary files, both estimated and measured wind data are used to calculated a mean monthly wind speed.

The introduction of the buoy data along the California coast during the 1980s is another example of the bias introduced by merging data from different origins in COADS. Along the US California coast, the wind time series given by the COADS 2°x2° monthly summaries files are strongly affected by the dominance of the buoys data in the data records after 1980, resulting in an artificial relaxation of the wind. Obviously, the use of these summary files can lead to a completely erroneous interpretation of the long term variability of the wind in the region.

The positive trend in the COADS wind time series off West Africa between 12°N and 22°N is a striking feature. Using estimated wind data, the magnitude of the trend results in an increase of about 50% of the wind stress over the last 40 years. This positive trend cannot be explained by an increasing fraction of anemometer type data in the wind time series as the anemometer wind data were not included in the calculation. The comparison with an independent wind time series and the dramatic increase of the correlation between CUI and SST when the trend is removed raises some doubt as to the validity of this trend. However, one should notice that a positive trend exists in the pressure derived wind in the region (see Fig. 6 in Ward and Hoskins, 1996). The trend of the North Atlantic Oscillation (NAO), which is an indicator of the strength of the atmospheric circulation in the Atlantic Ocean related to the Acores high (Hurrell 1995, 1996), is also in accordance with the long term trend that exists in the COADS wind data in the 14°N-16°N area (Fig. 11), but with a less pronounced upward trend during the last 30 years. Given these contradictory results, the validity of the trend in the COADS wind over West Africa remains an open question.

From our experience with the use of the COADS database to study long term climatic variability over the oceans, we conclude that extreme care should be taken in the interpretation of the results. It is strongly recommended to get access to the individual records rather than using the COADS 2°x2° monthly summaries files. The five CD-Roms produced by CEOS provide an easy to use alternative to these summary files.

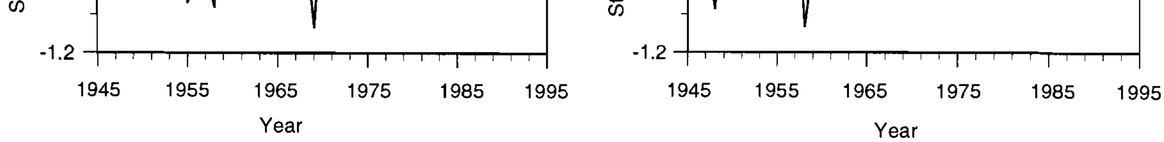

Fig. 11: Time series of the standardized anomalies of the North Atlantic Oscillation (NAO) and of the scalar wind speed during the upwelling season in an area off West Africa (14°N-16°N) from 1946 to 1990.

References cited

Arfi R. 1985. Variabilité interannuelle d'un indice d'intensité des remontées d'eau dans le secteur du cap Blanc (Mauritanie). *Can. J. Fish. Aquat. Sci.*, 42 (12): 1969-1978.

Bakun A. 1973. Daily and weekly upwelling indices, west coast of North America 1946-71. *U.S. Dep. Comm., NOAA Tech. Rep. NMFS SSRF*-671, 103p.

Bakun A., V. Christensen, C. Curtis, P. Cury, M.H. Durand, D. Husby, R. Mendelssohn, J. Mendo, R. Parrish, D. Pauly and C. Roy. 1993. *The Climate and Eastern Ocean Systems Project*. Naga, The ICLARM Querterly 15(4): 26-30.

Cardone V.J. 1969. *Specification of the wind distribution in the marine boundary layer for wave forecasting*. Report TR69-1, New York University, New York, NY, 131p.

Cardone V. J., J.G. Greenwood and M. A. Cane. 1990. On trends in historical marine wind data. *J. Climate*, 3: 113-117.

Folland C.K., D.E. Parker and F.E. Kates. 1984. Worldwide marine temperature fluctuations, 1856-1981. *Nature*, 310: 670-673.

Folland C.K. and J. Hsiung. 1987. Correction of seasonally varying biases in uninsulated bucket sea surface temperature data using a physical model. *Met. Office Synoptic Climatology Branch Memo.*, 154.

Folland C.K. and D.E. Parker. 1990. Observed variations of sea surface temperature. *In*: M.E. Schlesinger (ed.). Climate-ocean interaction, NATO workshop, Oxford, Kluwer Academic Publishers: 21-52.

Hurrell J. W. 1995. Decadal trends in the North Atlantic oscillation: regional temperatures and precipitation. *Science*, 269: 676-679.

Hurrell J. W. 1996. Influence of variations in extratropical wintertime teleconnections on northern hemisphere temperatures. *Geophys. Res. Lett.*, 23: 665-668.

Isemer H.J. 1995. Trends in the marine surface wind speed: ocean weather stations versus volontary observing ships. *In*: H.F. Diaz and H.J. Isemer (eds.). *Proceedings of the International COADS Winds Workshop*, Kiel Germany, 31 May-2 June 1994. US dept. of Commerce-Institut fur Meereskunde, Kiel: 68-84.

Isemer H.J. and L. Hasse. 1987. *The bunker atlas of the North Atlantic Ocean*. 2: Air-sea interactions. Springer-Verlag, 252p.

Isemer H. J. and L. Hasse. 1991. The scientific Beaufort equivalent scale: effect on wind statistics and climatological air-sea flux estimates in the North Atlantic Ocean. *J. Climate*, 4: 819-836.

Jones P.D., T.M.L. Wigley and P.B. Wright. 1986. Global temperature variations, 1861-1984. *Nature*, 322: 430-434.

Jones P.D. and T.M.L. Wigley. 1992. Corrections to pre-1941 SST measurements for studies of long-term changes in SSTs. *In* : H.F. Diaz, K. Wolter and S.D. Woodruff (eds.). *Proceedings of the International COADS Workshop*, Boulder, Colorado, 13-15 January 1992. US Department of Commerce, NOAA. 227-237.

Kaufeld L. 1981. The development of a new Beaufort equivalent scale. *Meteor. Rundsch.*, 34: 17-23.

Kent E.C., P.K. Taylor, B.S. Truscott and J.S. Hopkins. 1983a. The accuracy of voluntary observing ship's meteorological observations - Results of the VSOP-NA. *J. Atmos. & Oceanic Tech.*, 10: 591-608.

Kent E.C., R.J. Tiddy and P.K. Taylor. 1983b. Correction of marine daytime air temperature observations for radiation effects. *J. Atmos. & Oceanic Tech.*, 10: 900-906.

Nykjaer L. and L.V. Van Camp. 1994. Seasonal and interannual variability of coastal upwelling along northwest Africa and Portugal from 1981 to 1991. *J. Geophys. Res.*, 99 (C7): 14197-14207.

Portolano P. 1986. Analyse des séries vent-températures de la mer en surface le long des côtes sénégalaises. *Océanogr. trop.*, 21 (2): 205-227.

Ramage C. S. 1984. Can shipboard measurements reveal secular changes in tropical air-sea heat flux ? *J. Climate Appl. Meteor.*, 23: 187-193.

Ramage C.S. 1987. Secular change in reported wind speeds over the ocean. *J. Clim. Appl. Met.*, 26: 525-528.

Roy C. 1989. Fluctuations des vents et variabilité de l'upwelling devant les côtes du Sénégal. *Oceanologica Acta*, 12 (4): 361-369.

Slutz R.J., S.J. Lubker, J. D. Hiscox, S.D. Woodruff, R. L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms. 1985. *Comprehensive Ocean-atmosphere Data Set*; Release 1. NOAA Environmental Research Laboratories, Climate Research Prograam, Boulder, CO, 268p.

Ward M.N. and B.J. Hoskins. 1996. Near surface wind over the global ocean 1949-1988. *J. of Climate*, 9: 1877-1895.

Woodruff S.D. and S.J. Lubker. 1986. COADS 1980-85 update. *In*: S.D. Woodruff (ed.). *Proceedings of a COADS Workshop*, Boulder, Colorado, January 22-24, 1986. US Department of Commerce, NOAA. Technical Memorandum, ERL ESG-223: 36-48.

Woodruff S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer. 1987. A Comprehensive Ocean-atmosphere Data Set. *Bull. Amer. Meteor. Soc.*, 68: 1239-1250.

World Meteorological Organization (WMO). 1970. *Reports on marine science affairs*. 3: The Beaufort scale of wind force. WMO Geneva, Switzerland, 22p.