A time Series Analysis for Fish Production and Fish Supply in Sudan

Nidal Mohamed Mustafa Abd Elsalam

Department of Statistics & Computation, Faculty of Technological Mathematical Sciences & Statistics, Al Neelain University. E-mail: nidalmm@yahoo.com

ABSTRACT

In this work two time series, one representing Sudan annual fish production and the other the monthly fish supply in Khartoum markets, were studied. The classical decomposition method was first used, after checking for stationarity of both series, to detect and to estimate any trend or seasonal component that may be present in the fish supply series. Spectral analysis was also used to shed more light on the presence of cyclical patterns in the two series. On the time domain, an attempt was made to build mathematical models for fish production and fish monthly supply that may be used to forecast future production and supply. For this end, several models, including ARIMA and Exponential Smoothing were tried and the best among these was determined. Application of the decomposition method to the fish supply revealed an apparent seasonal effect in certain months. Spectral analysis showed that the dominance of a single cycle of length of about 7 years in the fish production series and a cycle of length 12 months in the fish supply series confirmed the decomposition method results. On the other hand, examination of the different models showed that the best forecasting model for fish product and supply series were the Quadratic Trend Model and Winter's Exponential Smoothing model, respectively. Forecasts for both series were carried out.

المستخلص

في هذه العمل تمت دراسة سلسلتان زمنيتان, واحدة تمثل الانتاج السنوي للسمك في السودان و الأخرى تمثل العرض الشهري للسمك بأسواق الخرطوم. أستخدمت طريقة التفكيك التقليدية أولا, بعد أن تم اختبار السكون في السلسلتين, حتى يتم كشف و تقدير مكوني الاتجاه العام و الموسمية الموجودين في سلسلة عرض السمك. كما أستخدم أيضا التحليل الطيفي لالقاء الضوء بصورة مكثفة على وجود الأنماط الدورية في السلسلتين. أما في مجال الزمن, فقد تمت المحاولة لبناء نماذج رياضية لاستخدامها بغرض التنبؤ المستقبلي بانتاج و عرض السمك. و تتضمن تلك النماذج الأريما و التمهيد الأسي, و ذلك كي يتم اختيار النموذج الأفضل من بينها. و قد كشفت طريقة التفكيك التقليدية التي أجريت على سلسلة عرض السمك عن تأثير موسمي واضح في بعض شهور السنة. كما أظهرالتحليل الطيفي سيطرة دورة منفردة يبلغ طولها حوالي 7 سنوات في سلسلة انتاج السمك, و أخرى بطول يبلغ 12 شهرا في سلسلة عرض السمك, وهذا يأتي تأكيدا لنتائج طريقة التفكيك التقليدية. و على صعيد اخر, أوضح اختبار جودة النماذج المختلفة أن النموذج الأفضل لسلسلتي انتاج و عرض السمك هو نموذج الاتجاه العام من الدرجة الثانية و نموذج "وينترز" للتمهيد الأسي, على التوالي. بعد ذلك تم التنبؤ بانتاج السمك السنوي العام 2030م, و عرض السمك الشهري الى شهر ديسمبر 2013م.

KEYWORDS: decomposition method, spectral analysis, Holt's Method, Winter's method and seasonal ARIMA.

INTRODUCTION

It is known that Sudanese seem to favor consuming red meats or white meats, namely fish and chickens, despite the high nutritional and health values associated with fish product (fresh or dried). No one appears to have a scientific justification for this matter, thus many questions persist like: is it because of the climate? Is it the tradition and taste? Is it the prices? or is it just a matter of lack of awareness as to which of the two types is healthier for consumption and domestic use?. These questions played a major role in motivating the Ministry of Animal and Fish Resources in Sudan [1] to plan to encourage and increase the citizen's share in white meat consumption, fish in particular, through a fish culture (Aquaculture) program. This program targets five states (Khartoum, Kassala, Nahral-Niel, Sinnar and White Nile) to establish 200 farms for fish culture with the aim to increase the fish production to 190000 tons annually instead of the current 70000 tones. Accordingly, this will lessen the burden from the natural fisheries while it is the dominant world policy where the country supports production of small fish and then the private sector will take over the whole process. Thus the need arises to study the Sudan fish production and fish supply in Khartoum state which may be affected by trend, seasonality and cyclical components.

The main objectives of this paper are; to study and analyze the fish production and supply time series in frequency domain and time domain, to explore and trace the effects of changes of trend and seasonality, and to try to find the most appropriate model representing the fish production and supply that can be used for forecasting.

Initially, the decomposition method, Makridakis [2] and Abraham [3], were applied to the series of fish supply. It seems that there is an obvious and strong seasonal component,

and no trend component is detected. In the context of frequency domain. Priestly [4] did a study which showed that, regarding fish production in Sudan from 1991 to 2010, the series is affected greatly by trend, accompanied by a slow cyclical component. For the monthly fish supply series, strong cyclical, as well as seasonal components, is dominant.

In the time domain, a number of models were tried out for both series to reach to an end about the best model for forecasting by Harvey [5] and Gaynor [6]. As a result, the Quadratic model of curve estimation procedure, and Winter's Exponential smoothing model, came out to be the best for forecasting of the fish production and fish supply, respectively.

MATERIALS and METHODS

To start with, checking for stationarity was a priority. This was done through plotting the auto-correlation functions (ACF) which can reveal if non-stationarity in mean or variance or both of them exists. If the series is stationary, the auto-correlations will drop to zero after the second or the third time lag. For non-stationary series, the ACF is significant from zero for several time lags, and this was what happened in both series. Figures (1) and (2) illustrate the presence of non-stationarity and trend effect. Secondly, the classical decomposition method was used for the series (fish supply). The first series is the Sudan fish production (fish-product) which is an annual series of twenty years long, from 1991 to 2010. The fish production is demonstrated at (000) tons. The second series is the fish supply in Khartoum fish markets which is a monthly series starting from January 2000 through December 2008 and the fish supply is given in tons. Sequence plots of the series is shown in figures (3) and (4).

Journal of Science and Technology Vol. 14
ISSN -1605 – 427X
Natural and Medical Sciences (NMS No. 1)
mena@sustech.edu

June /2013

Classical decomposition method

This method revealed that there is an apparent seasonal effect in certain months (May, June, December and January). The seasonal indices, shown in table (1), were the average deviation of each month's fish supply level from the level that was due to the other components to that month.

Period 1, which was in January, averaged about 0.82 units below the de-seasonalized fish supply. Periods 5 & 6 have the highest supply levels, while periods 1 & 12 have the lowest supply levels, leading to the fact that fish supply rises in Autumn and dips in Winter.

Table 1: Seasonal Indices of "Fish-supply"

Period	Index
1	0.824380
2	0.908099
3	0.986139
4	1.09277
5	1.11851
6	1.10465
7	1.04212
8	1.01488
9	1.03423
10	1.05433
11	0.960739
12	0.859170

The seasonal decomposition model accuracy measures as (MAPE: 10.7501, MAD: 3.2072 and MSD: 17.4934) with a trend line equation which will be needed for forecasting, as:

$$Yt = 31.2002 + 8.77E\text{-}02*t$$

Journal of Science and Technology Vol. 14
ISSN -1605 – 427X
Natural and Medical Sciences (NMS No. 1)
mena@sustech.edu

June /2013

Spectral analysis

It is widely known that spectral analysis continuous to be a valuable tool in finding out the various kinds of periodic and non-periodic behavior in a series [4]. This is through determining the magnitude and phase of periodic variation since it is looked up to the series as a result of waves of sine's and cosines. And in this case, the analysis will be model free, and is purely mathematical and is not based on any theory about a process underling the series, while its main objective will be to know the most effective waves' lengths and frequencies. Thus, a spectral analysis is carried out for both series. Considering the series of Sudan fish production, a periodogram along with spectral density are plotted separately in figures (5) and (6), respectively. In drawing the periodogram (an unsmoothed plot of spectral amplitude plotted on logarithmic scale against frequency or period), plots are produced by frequency ranging from 0 which represents the constant or mean term, to frequency 0.5 which is the term for a cycle of two observations. Then spectral density (a plot of the periodogram after it has been smoothed according to certain specifications such as width of the smoothing window i.e. span and weights applied to the neighboring observations) is plotted, and the chosen spectral window is Tukey-Hamming. This is done after trying all of the other kinds of windows such as: Parzen, Barlett and Daniell. Due to the fact that all of them came up with almost the same density, we stick to the most popular window (Tukey-Hamming). Regarding the span, a number of them, starting from 5, 7, 9, and up to 11 were tried out. This was done so as to choose the span that makes the spectral density plot easier to read, and in this case, it was span of 5. This is due to the fact that the wider the span, the more bias will be introduced through the missing of spikes corresponding to important periodic variation at narrow frequency ranges.

From figure (3) we can realize that the plot of fish product shows a strong trend effect with a slow cycle in the data as well. Moreover, the series periodic behavior has no sinusoidal shape at each important frequency. In this case, trend removal of the data may be helpful. This is because of the effects of a trend being like very low frequency variation will load most heavily on the lowest frequencies of the periodogram, and it will be reflected to a smaller degree in higher frequency terms. The trend removal attained here was through curve estimation procedure with a linear trend. The err series (err stands for error) which is created is a detruded series. Thus, a periodogram as well as spectral density was formed for forth series of fish product before and after trend removal of the detrended series. This is shown in figures (5), (6), (7) and (8). The same result was deduced from both analysis, the highest frequency was at nearly 0.15 (3/20) but with more clear spikes in figures of the detrended series. This indicates that we have only a single cycle of length of about 7 years in the 20 observations.

Moving to the series of fish supply in Khartoum Markets (fish-supply), the sequence plot of figure (2) demonstrates that the series contains 108 monthly observations, starting from January 2000 through December 2008. It seems that it was dominated by a seasonal effect but no trend effect was detected. Thus, periodogram and spectral density plots were carried out and this is shown in figures (9) and (10), respectively. Since there was no strong low frequency phenomenon dominating the periodogram, there was no need to remove seasonality which is a major concern of this study.

Figures (9) and (10) show that the highest peak was at frequency 1/12 (0.0833). This means that we have a single cycle of length 12 dominating the data. The fact that there was no trend effect is clearly demonstrated, since no peak is found to be very close to frequency 0.

Journal of Science and Technology Vol. 14
ISSN -1605 – 427X
Natural and Medical Sciences (NMS No. 1)
mena@sustech.edu

June /2013

Time domain

Time domain includes the Trend Analysis, Exponential Smoothing methods, Makridakis [2] et al, and ARIMA models, Anderson [7] and Box Jenkins [8]. But firstly, regarding series (fish-product), curve estimation method for Trend Analysis was used to build the most appropriate mathematical model. This is illustrated in table (2).

Table 2: Curve estimation of series (fish-product)

Model Type	Trend Equation	Accuracy Measures
Linear	Yt = 32.5 + 2.04286*t	MAPE: 5.53682,MAD: 2.65714 MSD:9.78643
Quadratic	Yt = 27.2588 + 3.47228*t - 6.81E-02*t**2	MAPE: 3.1181,MAD: 1.66167 MSD: 5.71937
Exponential	Yt = 34.1284*(1.04168**t)	MAPE: 7.33983,MAD: 3.67135 MSD: 17.1536

- MAPE: (stands for Mean Absolute Percentage) $MAPE = \frac{1}{n}\sum_{t=1}^{n} |\frac{et}{yt}|$
- MAD: (Mean Absolute Deviation) $MAD = \frac{1}{n}\sum_{t=1}^{n} |et|$
- MSD: (Mean Squared Deviation) $MSD = \frac{1}{n}\sum_{t=1}^{n} et^2$

Whereas; et, yt, represent the errors and the series observations respectively.

From table (2), it is obvious that the best trend equation is that of the model type (Quadratic), for its lowest accuracy measures among the other models, indicating a high level of accuracy and precision.

Moving to exponential smoothing models, Brown [12], and since the data in series (fish-product) is on annual basis, the relevant method is "Double Exponential Smoothing Method-Holts" with:

Smoothing Constants (Alpha (level): 0.552883, Gamma (trend): 0.170358)), and Accuracy Measures (MAPE: 3.54, RMSE: 3.051 and R^2 : 0.944). This can be more elaborated through introducing the Holts equations as follows:

$$S_t = 0.552883 X_t + (1 - 0.552883)(S_{t-1} + b_{t-1})$$

$$b_t = 0.170358(S_t - S_{t-1}) + (1 - 0.170358)b_{t-1}$$

$$F_{t+m} = S_t + b_t m$$

Regarding the ARIMA model, the best fitted model was ARIMA (0, 1, and 0), that can be expressed as:

$$(1 - B)X_t = e_t$$

$$\text{or} \quad \nabla X_t = e_t$$

This model indicates that the series under study was a random process, since no auto regression or moving average processes are found, while only the first difference has been taken to transfer the series into a stationary one. The model accuracy measures indicate that R-squared is 0.918, while MAPE is 3.653 and RMSE is 3.276.

Secondly, the same steps were followed in the second series (fish-supply). Starting by the exponential smoothing method, the most appropriate one was Winter's Three Parameters Method that is used specially when there is seasonality in the series. The chosen model was the Winter's Additive (where the magnitude of seasonal variation does not depend on the overall level of the series) with the three parameters' values; Alpha (level) 0.803, Gamma (trend) 0.0000268, and Delta (season) 0.001, leading to the following equations:

$$S_t = 0.803 \frac{X_t}{I_{t-L}} + (1 - 0.803)(S_{t-1} + b_{t-1})$$

$$b_t = 0.0000268(S_t - S_{t-1}) + (1 - 0.0000268)b_{t-1}$$

$$I_t = 0.001 \frac{X_t}{S_t} + (1 - 0.001)I_{t-1}$$

$$F_{t+m} = (S_t + b_t m)I_{t-L+m}$$

The accuracy measures namely were; R-squared 0.887, RMSE 2.086 and MAPE 4.005.

Moving to ARIMA modeling, the best fitted model is ARIMA$(0,1,0)(1,1,0)^{12}$ which could be expressed as:

$$(1 - \Phi_1 B^{12})(1 - B)(1 - B^{12})X_t = \varepsilon_t$$

The previous model stands for [(seasonal Auto regression of order 1), (non-seasonal difference) and (seasonal difference)]. The estimate of the Autoregressive parameter (AR)Φ was (-0.620) which is significant with standard error of 0.088. The seasonal and the non-seasonal differences were of the first order. This was associated with the model accuracy measures as R-squared 0.584, RMSE 2.422 and MAPE 4.644. There was one outlier in the observations in January 2002, with an estimate value of -6.405, which is significant with standard error (SE) of 1.427.

The best forecasting model

For the first series (fish-product), it can be noticed that the Quadratic model in the curve estimation procedure of figure (11), was the best as a model for forecasting among the Double Exponential Smoothing (Holts) and the ARIMA random process model [ARIMA (0,1,0)], since it acquires the lowest MAPE value. But if the comparison is done through the Normalized Basyian Information Criterion (BIC), which is followed by most statistical packages, ARIMA (0,1,0) will be the best model for forecasting, as suggested by Pankratz [9] and Bowerman [10, 11]. This is clear from table (3) which demonstrates the different models with their different accuracy measures.

Table 3: Models & accuracy measures for the series "fish-product"

Models	*Accuracy measures*
Curve Estimation(Quadratic Trend Model)	MSD: 5.71937,MAPE: 3.11812 , R^2:0.934
Holt's Exponential Model	MAPE: 3.54, RMSE: 3.051, R2:0.944,

Journal of Science and Technology Vol. 14
ISSN -1605 – 427X
Natural and Medical Sciences (NMS No. 1)
mena@sustech.edu

June /2013

	BIC: 2.531
ARIMA (0,1,0)	MAPE: 3.65300, RMSE:3.276, R^2: 0.918, BIC: 2.528

Moving to the series "fish-supply", the three methods' Accuracy Measures of seasonal decomposition method, Winters' exponential smoothing method and seasonal ARIMA model, are presented in table (4). Since the lowest MAPE value (4.005), highest R^2 value (0.887) and lowest Normalized BIC value (1.600) were that of Winter's Method (figure 12), it is regarded as the best model for forecasting, as indicated by Bowerman [10,11].

Table 4: Models &accuracy measures for the series "fish-supply"

Models	*Accuracy measures*
Seasonal Decomposition	MAPE: 10.7501
Winter's Exponential Model	MAPE: 4.005, R^2: 0.887,BIC:1.600
ARIMA $(0,1,0)(1,1,0)^{12}$	MAPE: 5.042, R^2: 0.422,BIC:2.135

RESULTS

In this study, the classical decomposition method was applied to detect any trend or seasonal component in the series of fish supply. Subsequently, spectral analysis was carried out to explore and examine the cyclical component of the two series "fishproduct" and "fishsupply". Later on, exponential smoothing methods and ARIMA models were also carried out, among which comparisons were made to reach to the best model for forecasting. For more elaboration, the basic findings of the study are presented in:

• The classical decomposition method for the series of fish supply revealed an apparent seasonal effect. Fish supply increase in Autumn and decrease in Winter.

• Periodograms and spectral densities for the annual series of fish product, before and after trend removal, and for the monthly series of fish-supply, showed the dominance of a single cycle of about 7 years length, and another one of 12 months length, respectively. On the other hand, in the time domain context, the best forecasting model for the first annual series was the Quadratic model of the curve estimation procedure. While the best forecasting model for the second monthly series was Winter's exponential Smoothing Method model.

Forecasts for Sudan's fish production and fish supply in Khartoum's fish markets up to year 2030 and December 2013, respectively, had been carried out. This is clearly seen from figures (11 & 13). In figure (11), fish production forecasting shows an apparent increase up to year 2020, and then it starts to

Journal of Science and Technology Vol. 14
ISSN -1605 – 427X
Natural and Medical Sciences (NMS No. 1)
mena@sustech.edu

June /2013

decline up to the end of the period. Taking into consideration figure (13), it can be realized that almost the same forecasts are carried out throughout the period of forecasting, with a very slight increase in fish supply. This is due to the fact that Winter's Method, as an exponential smoothing method, is usually used to forecast only one period ahead.
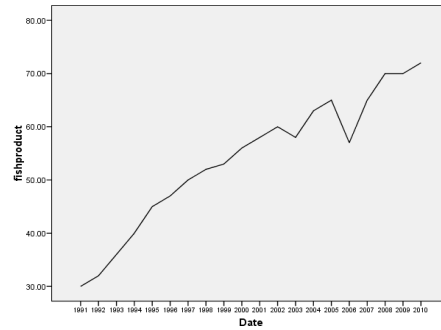

Figure 3: Sudan Annual Fish Production from 1991 to 2010
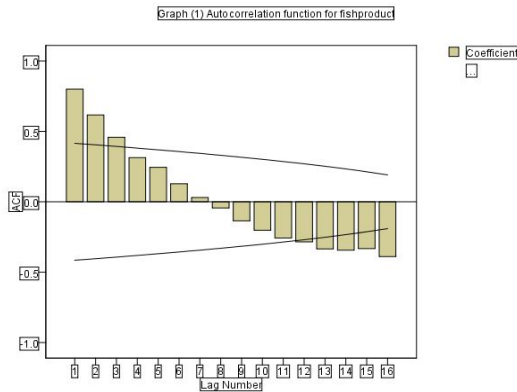
Figure 1: Autocorrelation function of fish-product
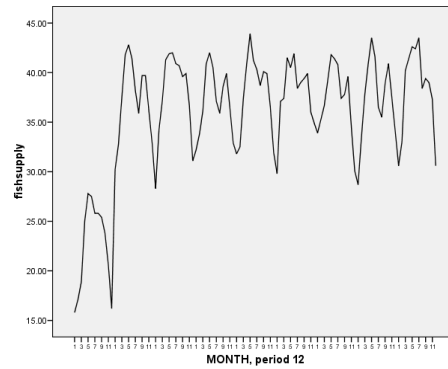
Figure 4: Monthly Fish Supply in Khartoum's Fish Markets
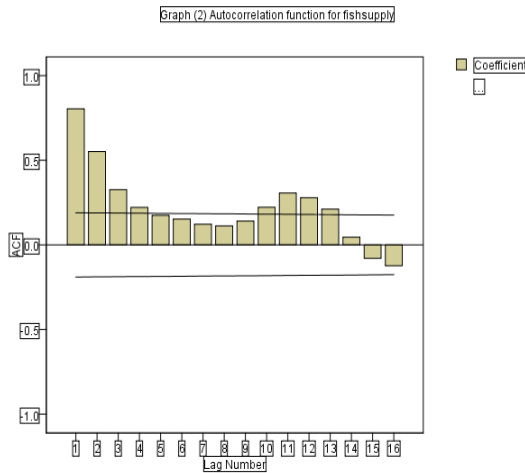
Figure 2: Autocorrelation function of fish-supply
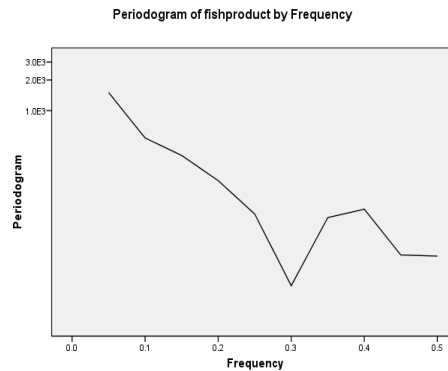
Figure 5: Periodogram of fish-product by frequency

51
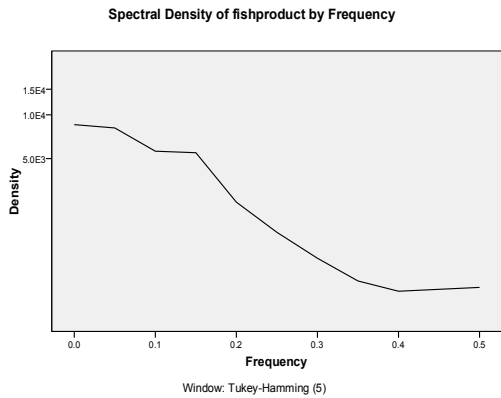
Figure 6: Spectral density of fish-product by frequency
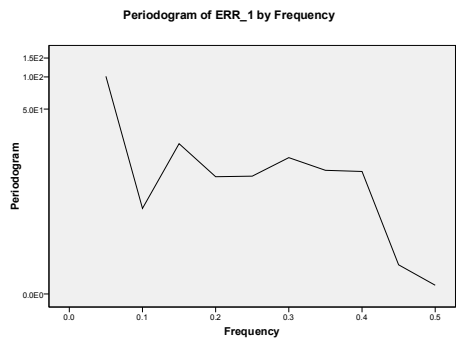
Figure 7: Petrodogram of err-1 by frequency
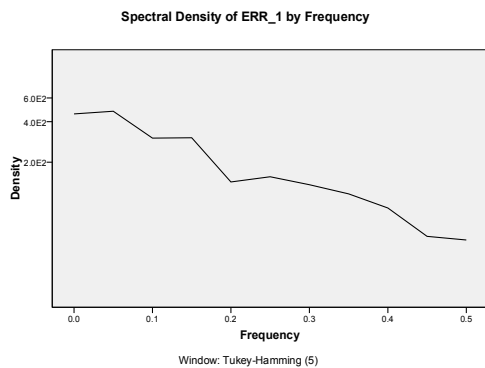

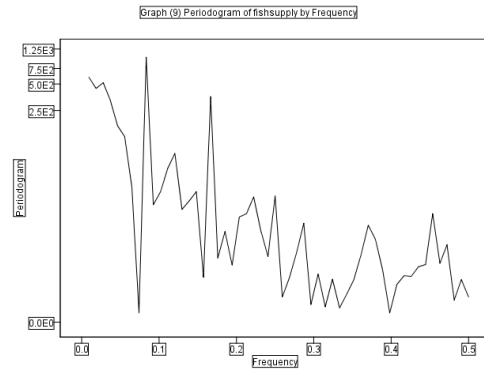*Figure 8: Petrodogram of err-1 by frequency*
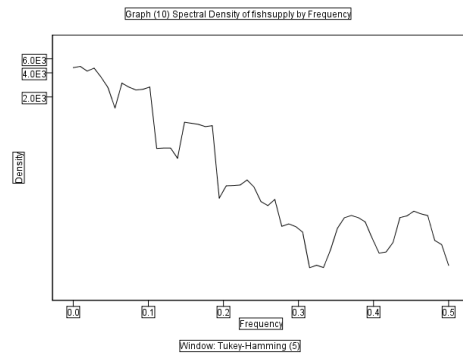

Figure 9: Petrodogram of fish-supply by frequency

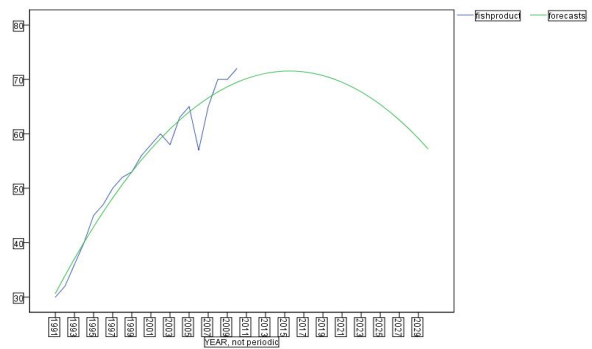*Figure 10: Spectral analysis of fish-supply by frequency*

Figure 11: Forecasts for the fish-production series up to 2030 using the Quadratic Model

DISCUSSION

An attempt has been made to study Sudan's fish production and fish supply in Khartoum's fish markets. This was done in both the frequency domain and the time domain. Checking for stationary of both series, results in the fact that they are not stationary in mean and variance. Spectral analysis indicated that the annual fish product is strongly dominated by trend effect which is accompanied by a low cyclical behavior. While, the monthly fish supply reflected a strong cyclical and an apparent seasonal variation. In the time domain, Holt's exponential smoothing model was not appropriate as a forecasting model of "fisproduct" series even though the series at the beginning exhibited a slow changing pattern after which it started to change rapidly, from the year 2002. The ARIMA (0, 1, and 0) model was not suitable either since only a random process took place instead of auto regression or moving averages processes. For the series "fish-supply", the seasonal decomposition method was neither the best for forecasting, nor the ARIMA model [ARIMA (0,1,0) (1,1,0)] [12], since cyclical and seasonal components were dominant. Moreover, the presence of an outlier at the beginning of the series complicated the situation. It is hoped that the results will help to throw more light on the issues concerning fish production and consumption.
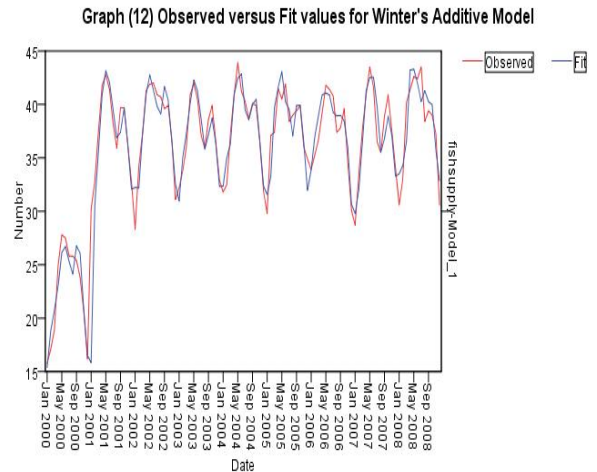

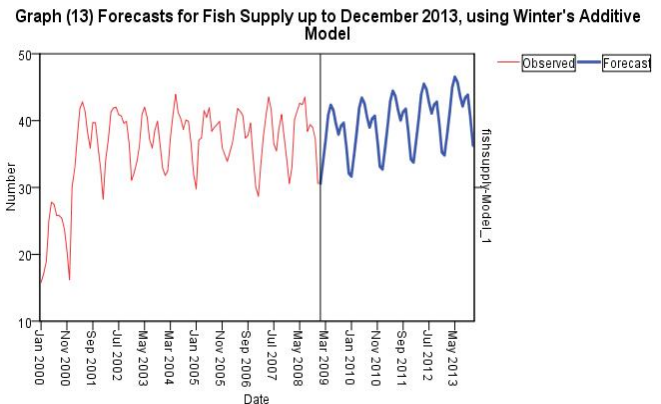*Figure 12: Fit values of Winter's additive model*

Figure 13: Forecasts for fish-supply up to December 2013 using Winters model

REFERENCES

1.http://www.marf.gove.sd/statistics-mag.pdf (retrieved 17 March 2013).

2. Makridakis S, Wheelwright SC, and Hyndman RJ. (1998). *Forecasting Methods and Applications,* 3rdedition. John Wiley, New York, USA.

3. Abraham B, and Ledolter J. (1983). *Statistical Methods for Forecasting.* John Wiley & sons, New York, USA.

4. Priestly MB. (1981). *Spectral Analysis & Time Series.* Academic Press, New York, USA.

5. Harvey AC. (1981). *Time Series Models.* John Wiley, New York, USA.

6. Gaynor E, and Kirkpatrick C. (1994). *Introduction to Time Series Modeling& Forecasting in Business & Economics.* McGrow-Hill, New York, USA.

7. Anderson TW. (1971). *The Statistical Analysis of Time Series.* John Wiley & Sons, New York, USA.

8. Box GP, Jenkins GM, and Reinsel GC. (1994). *Time Series Analysis: Forecasting & Control.* Holden Day, San Francisco, USA.

9. Pankratz A. (1983). *Forecasting With Univariate Box-Jenkins Model: Concepts & Cases.* John Wiley, New York, USA.

10. Bowerman BH, and O'Connell RT. (1979). *Forecasting & Time Series: An Applied Approach,* 3rdedition. Duxbury Classic Series, Boston, USA.

11. Bowerman BH, and O'Connell RT. (1987). *Time Series Forecasting.* Duxbury Press, Boston, USA.

12. Brown RG. (1962). *Smoothing, Forecasting and Prediction of Discrete Time Series.* Prentice Hall, New Jersey, USA.