

Analysis of Some Linear Dynamic Systems with Bivariate Wavelets**Prof. Dr. Taha Hussein Ali**drtahaalaa1970@yahoo.com**Mardin Samir Ali**mardins.ali@gmail.com**Abstract**

There are many statistical methods related to the forecasting of time series without any input variables such as autoregressive integrated moving average (ARIMA models). In this research, some linear dynamic systems, represented by ARIMA with exogenous input variables (ARIMAX models) were used to forecast crude oil prices (considered as output variable) for OPEC organization with the help of crude oil production (considered as input variable) depending on the data starting from the period of 1973 until 2018. Using traditional ARIMAX method and proposed method (Bivariate Wavelet Filtering) for the time series data in order to select one of them for forecasting through comparing some measures of accuracy, such as MSE, FPE, and AIC. Then, applying crude oil prices for OPEC using the traditional ARIMAX models and ARIMAX models with applying the *bivariate wavelet* filtering, especially bivariate Haar wavelet. The main conclusions of the research were that the success of bivariate wavelet filtering in forecasting of crude oil prices using proposed model was more appropriate than traditional models, and the forecasting of crude oil prices using proposed method in 2020 will be fairly less than 2019.

Keywords: Time Series, linear dynamic systems ARIMAX models,

Wavelet transformation, Bivariate Wavelet Filtering

This is an open access article under the CC BY 4.0 license
<http://creativecommons.org/licenses/by/4.0/>.

تحليل بعض النماذج الديناميكية الخطية مع الموجات الثنائية

الملخص

هناك العديد من الطرائق الإحصائية المتعلقة في التنبؤ بالسلسلة الزمنية دون أية متغيرات داخلية مثل أنموذج الانحدار الذاتي المتكامل للأوساط المتحركة نماذج (ARIMA). وفي هذا البحث، تم استخدام بعض الأنظمة الديناميكية الخطية المتمثلة بأنموذج ARIMA مع متغيرات المدخلات الخارجية (نماذج ARIMAX)، فقد استخدمت للتنبؤ بأسعار النفط الخام (بعد كمتغير تابع) لمنظمة أوبك بمساعدة إنتاج النفط الخام (بعد كمتغير مستقل) واعتماداً على البيانات التي تبدأ من الفترة 1973 حتى 2018. باستخدام الطريقة التقليدية ARIMAX والأسلوب المقترح (مرشح الموجات الثنائية) لبيانات السلسلة الزمنية من أجل اختيار إحدهما للتنبؤ من خلال مقارنة بعض مقاييس الدقة مثل MSE ، FPE ، AIC .

* Professor/ Statistics Dept./ Adm. & Eco. College/ Salahaddin University – Erbil. ** MSc Student/ Statistics

**Dept./ Adm. & Eco. College/ Salahaddin University – Erbil

ثم تطبيق اسعار النفط الخام لأوبك باستخدام نماذج ARIMAX التقليدية ونماذج ARIMAX مع تطبيق مرشح الموجات ثنائية المتغير، وخصوصاً موجة هار الثنائية (Haar). ومن أهم الاستنتاجات التي توصلت إليها الدراسة أن نجاح مرشح الموجات ثنائية المتغير في التنبؤ بأسعار النفط الخام باستخدام النموذج المقترح كانت أكثر ملائمة من النماذج التقليدية، وسيكون التنبؤ بأسعار النفط الخام باستخدام الطريقة المقترحة في سنة 2020 أقل نوعاً ما من 2019.

الكلمات الرئيسية: السلاسل الزمنية ، الأنظمة الديناميكية الخطية، نماذج ARIMAX ، تحويل الموجات ، مرشح الموجة الثنائي.

1. Introduction

An important field of Statistics is Time Series Analysis which could be used in many other scientific fields in order to make future planning and management for governmental and non-governmental institutions. It deals with the methods and theory involved in analyzing datasets which are collected over time and helps to understand the past behavior of phenomena. It can be used for comparison between two or more times series concerning the type of growth, for instance growth in crude oil prices, consumption of a product, etc.

Forecasting of Linearly time series analysis is a famous developed method that researchers use in their forecast, but in the actual world most data systems are nonlinear, which demonstrates itself in a much larger extents of possible dynamical behaviors (Brockwell and Davis, 2016). Also, it is clear that models based on the time series only, without taking any related input variable into consideration, could not be accurate. Therefore, more accurate models are taken by integrating exogenous input variables into the model which is called Linear Dynamic Systems (Nelles, 2001).

The linear dynamic systems (LDS) model is the most commonly used time series model for economic, engineering, financial, and medical applications because of its respective simplicity, mathematically predictable behavior, inference, and predictions could be done effectively for the models. It is possible to identify LDS models using *parametric* and *nonparametric* models, and both are divided into *time domain* models and *frequency domain* models. Time domain models consist of *equation error* models and *output error* models. Equation error models are the study's goal. Thus, the equation error models is divided into ARX, ARMAX, and ARARX (Nelles, 2001; Saaed , 2015; 2017 (أبوليدة،)).

The ARIMAX models can be utilized to enhance the forecasting of output series (Y), through applying the past data of both the Y series and the exogenous input series (X). That is especially right if the input series is an important indicator (Wei, 2006).

In human life, data are used to deal with daily work in any field and very important to organize life, but analyzing and forecasting of these data need some processing such as denoising in order to reduce contamination through using many filtering methods, such as wavelet filtering transform. The wavelet transform is a new mathematical tool developed mainly since the mid-1980's (Sheng, 2000). The first such use of wavelets was in geophysics for analyzing seismic surveys that are used in oil and mineral exploration in order to get pictures of layering in subsurface rock. Geophysicists rediscovered them; mathematicians had developed them to solve abstract problems some 30 years earlier, but had not anticipated their applications in signal processing (Boggess and Narcowich, 2009).

There are many types of wavelet transforms such as Continuous Wavelet Transform (CWT), Discrete Wavelet Transform (DWT), Fast Wavelet Transform, etc. The DWT has many attractive qualities that make it a good method for time series, displaying features that change in both time and frequency. Haar Wavelet is one of the DWT that was used first by its creator Alfred Haar in 1910.

In the practical part of this study, applying two forecasting methods represented by traditional and researcher's proposed methods on real data for forecasting yearly crude oil prices with helping of input exogenous time series variable represented by crude oil production for the Organization of the Petroleum Exporting Countries (OPEC).

The main objectives of this research are denoising contaminated data using bivariate wavelet filtering in estimation of some linear dynamic systems and forecasting them, yearly crude oil prices forecasting for OPEC Organization using traditional and proposed methods, and comparison between these two methods in order to obtain best model using statistical measurements.

The great importance of this research is that many institutions and organizations headed by OPEC Organization, governmental bodies and private bodies, ministry of planning, ministry of finance and economy can benefit from it, and use it as the basis of future strategic plans.

2. Methods and Materials

In this section, time series analysis will be discussed with dynamic system models including ARIMAX models and then bivariate wavelet also will be discussed. So, this section describes theories and models that will be applied in practical section of the research.

2.1 Time Series Analysis and Linear Dynamic Systems

One of the most important fields of future estimation is time series, It is a data collection recorded over a period of time such as weekly, monthly,

quarterly, or yearly. An analysis of history can be used by administration to make present decisions and plans depending on long-term forecasting. Usually a researcher can assume that past patterns will continue into the future (Cryer and Chan, 2008).

Time series is a set of repeated observations of the same variable over time (Cochrane, 2005). Time series, as a stochastic process, is an ordered sequence of observations made sequentially in time. The most important feature of such data is the likely lack of independence between successive observations in time. Time series data can be univariate (contains one random variables) or multivariate (more than one random variable) (Akpanta and Okorie, 2014).

2.1.1 Linear Dynamic Systems (LDS)

These types of models are the most common, it is also known as Black Box models. In general, the models of LDS are an expansion of the time series models by adding the input variable (x) to them, as shown in Figure 1, because the time series models are usually not accurate, and divided LDS models into two parts: Equation error models and Output error models (Andersen, 1997; Isermann and Münchhof, 2010; Nelles, 2001; الشرايبي, 2017).

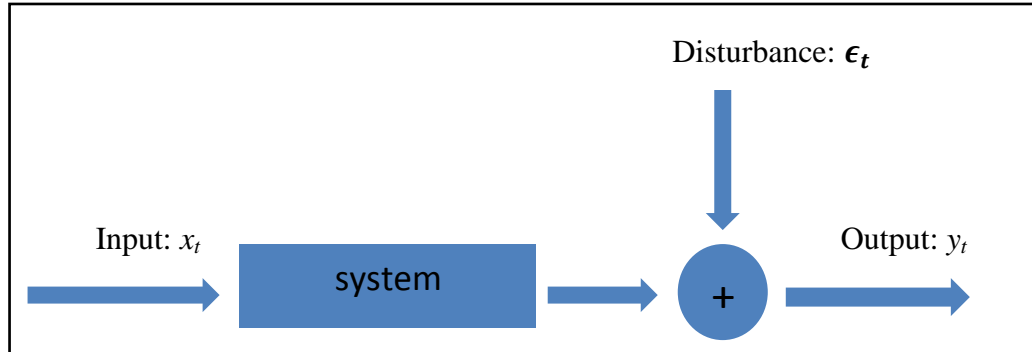


Figure (1): Dynamic System with Input x_t , Output y_t and Disturbance ϵ_t

2.1.1.1 Equation Error Models

This type includes ARX and ARMAX models, and is characterized by the existence of a polynomial order $\Phi(B)$ as the dynamic denominator of the input transform function and the noise transform function. This is in line with the fact that noise does not affect directly on the output of the model, but instead enters the model before the $\frac{1}{\Phi(B)}$ filter, where $\Phi(B)$ is presented in Equation 1 (Ljung, 1999).

$$y_t = \frac{1}{\Phi(B)} \epsilon_t \dots (1)$$

where: $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$

a) Autoregressive with exogenous input (ARX) model

This model represents the simple relationship between the input and output, (see Figure 2) provided by the linear difference Equation 2, which is also sometimes called the equation error model due to its description of white noise as a discrete error (Ljung, 1999; Edrees, 2015; شعيث وجمعة ,2017):

$$y_t = \frac{\Psi(B)}{\Phi(B)} x_t + \frac{1}{\Phi(B)} \epsilon_t \dots (2)$$

where $\Psi(B)$ and $\Phi(B)$ are the transfer functions of the deterministic part and the stochastic part, respectively and x_t is the input signal, y_t is the output signal, and ϵ_t is the white noise (equation error).

Thus, Equation 2 could be rewritten as shown below:

$$\Phi(B)y_t = \Psi(B)x_t + \epsilon_t \dots (3)$$

$$\Rightarrow y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_{na} y_{t-na} = \psi_1 x_{t-nk} + \psi_2 x_{t-nk-1} + \dots + \psi_{nb} x_{t-nk-nb+1} + \epsilon_t \dots (4)$$

The orders of the above model are **na** and **nb**, whereas **nk** is the model time delay.

By introducing B-notation, which is widely used in the system identification literature for the sake of compatibility with the traditional definition of the **B** transform (Backshift Operator), the numerator and denominator of Equation 2 can be defined as the following polynomials (Ljung, 1999):

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{na} B^{na}$$

and

$$\Psi(B) = \psi_1 B^{nk} + \psi_2 B^{nk+1} + \dots + \psi_{nb} B^{nk+nb-1}$$

It is possible to forecast the output variable using the its previous observations, as presented in the following formula:

$$\hat{y}_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_{na} y_{t-na} + \psi_1 x_{t-nk} + \psi_2 x_{t-nk-1} + \dots + \psi_{nb} x_{t-nk-nb+1} + \epsilon_t \dots (5)$$

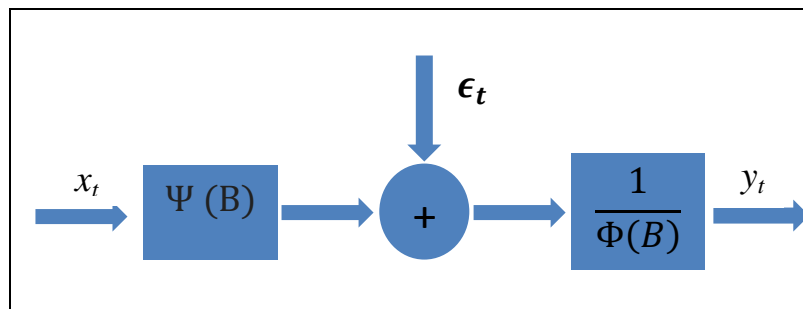


Figure (2): ARX Model Structure (Nelles, 2001; Ljung, 2001)

It should be noted that the 'AR' part of 'ARX' denotes the autoregressive part $\Phi(B)y_t$ while the letter 'X' denotes the extra input $\Psi(B)x_t$.

b) Autoregressive Moving Average with Exogenous input model (ARMAX)

The main problem with the ARX models is the limited scope for defining of the disturbance term. ARMAX models overcome this problem by defining the equation error as a Moving Average (MA) of the white noise in the ARX models (Roubal, 2009; Nelles, 2013). The ARMAX model is presented in Equation 6 and its flowchart is depicted in Figure 3.

$$y_t = \frac{\Psi(B)}{\Phi(B)} x_t + \frac{\Theta(B)}{\Phi(B)} \epsilon_t \dots (6)$$

where $\Psi(B)$, $\Phi(B)$ and $\Theta(B)$ are the transfer functions of the deterministic part and the stochastic part respectively, while x_t is the input variable, y_t is the output variable, and ϵ_t is the white noise (Saaed, 2015).

The numerator and denominator of Equation 6 could be defined as the following polynomials (Wei, 2006; Lee, 2010):

$$\begin{aligned}\Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{na} B^{na} \\ \Psi(B) &= \psi_1 B^{nk} + \psi_2 B^{nk+1} + \dots + \psi_{nb} B^{nk+nb-1} \\ \Theta(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_{nc} B^{nc}\end{aligned}$$

where **na**, **nb** and **nc**, are the orders of model, whereas **nk** is the model time delay.

It is possible to rewrite Equation 6 in another way, and it is as follows (Ljung, 2001):

$$\Phi(B)y_t = \Psi(B)x_t + \Theta(B)\epsilon_t \dots (7)$$

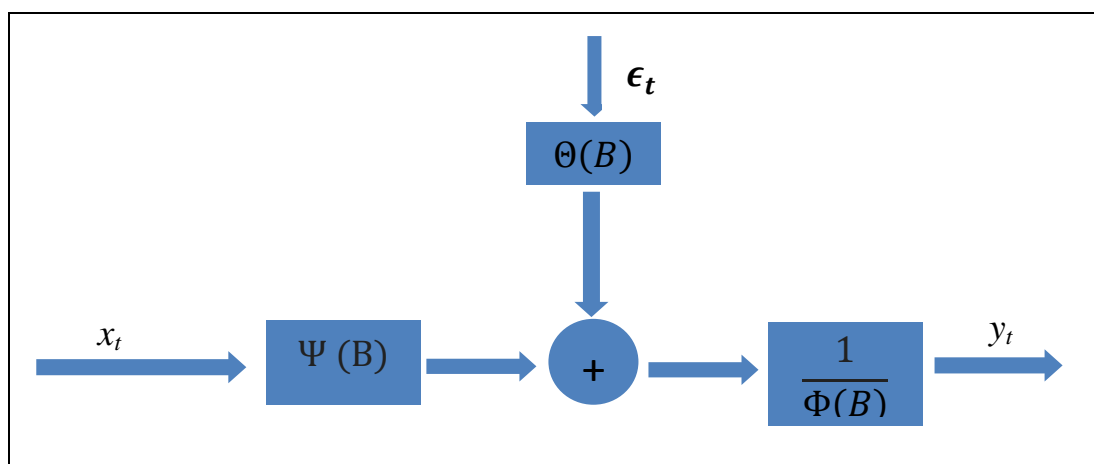


Figure (3): ARMAX Model Structure (Nelles, 2001; Ljung, 2001)

The above ARMAX model is the most general among the models, because it has some special case (Nelles, 2001; الشرايبي, 2017) :

- 1) When $\Psi(B) = 0$ and $\Theta(B) = 1$, AR model will be obtained.
- 2) When $\Psi(B) = 0$ and $\Phi(B) = 1$, MA model will be obtained.
- 3) When $\Phi(B) = 1$ and $\Theta(B) = 1$, FIR¹ model will be obtained.
- 4) When $\Theta(B) = 0$, ARX model will be obtained.
- 5) When $\Psi(B) = 0$, MA model will be obtained.

It is possible to model the equation error as an autoregressive in which case the gotten model will be called an ARARX model if the equation error was modeled using the AR approach (instead of an additional MA filter $\Theta(B)$, it possesses an additional AR filter $\frac{1}{\Phi^*(B)}$ of the disturbance), or an ARARMAX model if an ARMA approach was used (Nelles, 2001).

It is possible to forecast the output variable using its previous observations, as presented in the following formula:

$$\hat{y}_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_{na} y_{t-na} + \psi_1 y_{t-nk} + \psi_2 y_{t-nk+1} + \dots + \psi_{nb} y_{t-nk-nb+1} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_{nc} \epsilon_{t-nc} \quad \dots (8)$$

2.2 Wavelet Analysis

Wavelets become very useful when applied to many problems, including analysis of time series (in acoustics, geology, filtration and forecasting in meteorology and economics), effective data storage, especially images (computer graphics, image animation in movie industry). Lately, a very fast development of wavelet based data mining techniques may be observed (Percival and Walden, 2006). One of the duties of knowledge discovery preprocessing is noise reduction. The goal of this sub-process is to separate the noise from the signal and then to reduce or remove the noise. The definition of a noise stays unclear because for some physical processes, it is difficult to clearly define it. In this research, it is assumed that the goal of filtering is to change the input signal in such a manner that the values of a time series are changed but the characteristic of signal stays the way it was. Most commonly used methods consider statistical approach or involve Fourier transform (Kozłowski, 2004; Yu et al., 2001).

There are two functions that play a primary role in wavelet analysis, they are the wavelet function ψ and the scaling function ϕ . These two functions generate a family of functions that can be used to reconstruct a signal. To emphasize the "marriage" involved in building this "family", ϕ is sometimes called the "father wavelet", and ψ is indicated to the "mother wavelet" (Boggess and Narcowich, 2009).

¹ FIR: Finite Impulse Response.

2.2.1 Wavelet Properties

Wavelet is a small wave that grows and decays in a limited time period instead of being a wave that goes on forever, such as sinusoidal waves which are determined on a whole time domain $(-\infty, \infty)$ as seen in Figure 4 (Percival and Walden, 2000; Kozłowski, 2005).

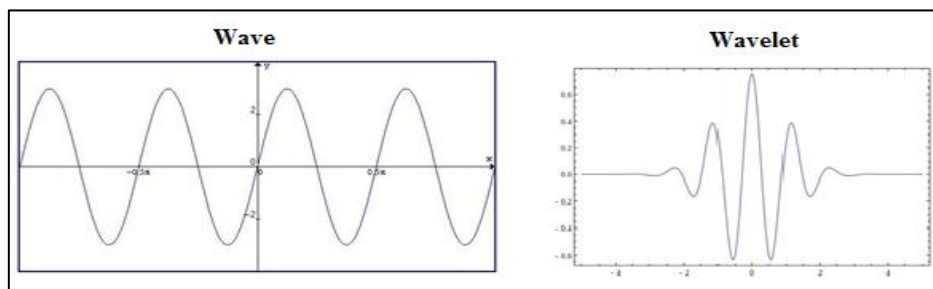


Figure (4): Difference between Wave and Wavelet (Chan, 1995)

Let ψ be a real function of a real variable ω which fulfills the following conditions (Kozłowski, 2005):

$$\int_{-\infty}^{\infty} \psi(\omega) d\omega = 0 \quad \dots (9)$$

and

$$\int_{-\infty}^{\infty} \psi^2(\omega) d\omega = 1 \quad \dots (10)$$

Condition 10 means that for any $0 < \varepsilon < 1$, there is an interval $(-T, T)$ such that (الزبيدي، 2009):

$$\int_{-T}^T \psi^2(\omega) d\omega > 1 - \varepsilon \quad \dots (11)$$

Equation 11 is meaning the following also:

$$\int_{-\infty}^{-T} \psi^2(\omega) d\omega + \int_T^{\infty} \psi^2(\omega) d\omega < \varepsilon \dots (12)$$

If ε is close to 0, it may be seen that only in an interval $(-T, T)$ corresponding to this ε values $\psi(\omega)$ are different than 0. Outside of this interval they must be equal to 0. Interval $(-T, T)$ is small compared to an interval $(-\infty, \infty)$, on which a whole function is determined. Equation 9 implies that if $\psi(\omega)$ has some positive values, it also has to have some negative ones (a function “waves”). Therefore Equations 9 and 10 introduce a concept of a wavelet (small wave).

There is another property of wavelets which is the admissibility condition and it is as follow (Sheng, 2000):

$$\int \frac{|\psi(\omega)|^2}{|\omega|} d\omega < +\infty \dots (13)$$

This can be used to first analyze and then reconstruct a signal without loss of information. In (13) $\psi(\omega)$ stands for the Fourier transform of $\psi(t)$.

The admissibility condition implies that the Fourier transform of $\psi(t)$ vanishes at the zero frequency, i.e. $|\psi(\omega)|^2|_{\omega=0} = 0$

There are different Wavelet types of transformations that are used to suit different practical applications, such as (Hubbard, 1996; عبدالقادر, 2011):

- Continuous Wavelet Transform (CWT)
- Discrete Wavelet Transform (DWT)
- Fast Wavelet Transform
- Wavelet Packet Transform
- Stationary Wavelet Transform

2.2.2 Discrete Wavelet Transform (DWT)

One of the most important transforms that is used in wavelets is DWT. This kind of transform and its variants are applied in a wide variety of disciplines, such as geology, turbulence, atmospheric science, applied mathematics, etc. (Gencay et al., 2002). Therefore, the researcher tried to provide a broad idea about this type of wavelet transforms as an introduction of the practical see then chapter of this research and using it with time series analysis because it is a basic tool required for time series studying through wavelets.

The difference between CWT and DWT is that the DWT is obtained by using a finite number of scales instead of making transform for all scales. Discrete wavelets are again formed from a mother wavelet, but with scale and shift in discrete steps. The DWT makes the connection between wavelets in the continuous time domain and “filter banks” in the discrete time domain in a multiresolution analysis framework. Mathematically, the difference between CWT and DWT is in the mother wavelets (ψ), where t and λ will be transformed into $t = 2^j k$ and $\lambda = 2^j$ (Chui, 1992; Vidakovic, 1999):

$$\psi_{\lambda,t}(\omega) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{\omega - 2^j t}{2^j}\right) \quad \dots (14)$$

where:

j: Scale level parameter.

k: Shift parameter.

There are many types of wavelets, such as Haar, Daubechies, Meyer, Koiflets, Morlet, Shannon, Franklin, etc. (Daubechies, 1992, Grossmann and Morlet, 1984). In the next section, the researcher will discuss only one kind of these wavelets which is called Haar wavelet, because of its simplicity and easiness by most researchers and compatibility with the data that is used in practical part.

2.2.3 HAAR Wavelet

First wavelet filter was introduced by the Hungarian mathematician Alfred Haar in 1909 – 1910 (Haar, 1910). The Haar wavelet is one of the earliest examples of what is known now as a compact, dyadic, and orthonormal wavelet transform. The Haar wavelet, being an odd rectangular pulse pair, is the simplest and oldest orthonormal wavelet with compact support (Stankovic and Falkowski, 2003). There are two functions that have great roles in this wavelet, they are Scale function (ϕ) and wavelet function or mother wavelet (ψ). The scale function could be defined as follows (Burrus et al., 1998; Cheng, 2011; Morettin et al., 2017; Debnath and Shah, 2017):

$$\phi(\omega) = \begin{cases} 1 & 0 < \omega < 1 \\ 0 & \text{otherwise} \end{cases} \quad \dots (15)$$

whereas, the mother wavelet could be written as follows:

$$\psi^{(H)}(\omega) = \begin{cases} 1 & 0 < \omega < \frac{1}{2} \\ -1 & \frac{1}{2} < \omega < 1 \\ 0 & \text{otherwise} \end{cases} \quad \dots (16)$$

The following figure shows the Scale function and wavelet function (mother wavelet) (Burrus et al., 1998).

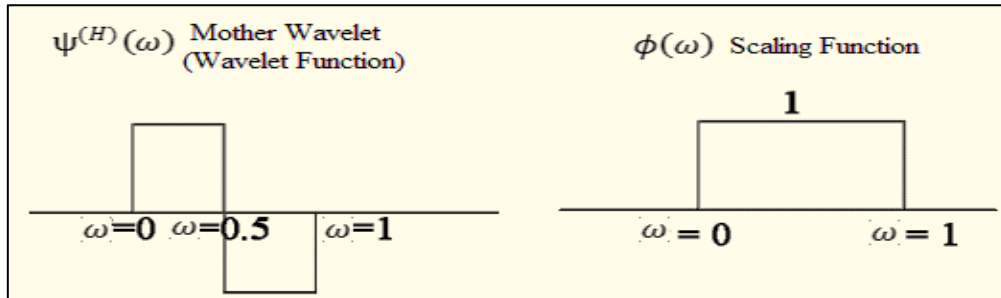


Figure (5): Haar Wavelet Type Describing Scaling Function and Mother Wavelet

The Haar wavelet has the following characteristics:

1. Compact Support
2. Orthogonal
3. Biorthogonal
4. Symmetric

As a result of wavelet transform, it could be obtaining a set of wavelet coefficients calculated at different levels (scales) and in a wide range of observation points. There are many ways of doing this, two most often applied are orthonormal DWT and its slightly modified version which

preserves scales but calculates wavelets in more densely chosen observation points (Kozłowski, 2005).

2.2.4 Wavelet Shrinkage

One of the most explored signal smoothening or cutting method utilizing wavelets is Wave Shrink method. Most of the applications of wavelet shrinkage deal with a single run of a time series expressed in a regression fashion: data = signal + noise (Donoho and Johnstone, 1995; Morettin et al., 2017). The method is composed of three main steps (Raimondo, 2002; Kozłowski, 2005):

- 1) Apply the discrete wavelet transform DWT: At the beginning the observed time series is transformed into the wavelet space by DWT.
- 2) Shrink the wavelet coefficients towards zero: In step two wavelet coefficients are modified, reduced according to the selected shrinkage function and a given threshold value. To accomplish this, one of three shrinkage functions may usually be used to establish how to modify wavelet time series coefficient.
- 3) Inverse DWT is applied on wavelet coefficients and as a result smoothened original signal (with reduced noises) is derived.

Shrinkage functions are formulas that define a correction coefficient $\delta_\lambda(x)$, which is subtracted from the corresponding wavelet coefficient. Calculated correction is relevant (different from 0) for those wavelet values, which exceed a given threshold parameter λ (Kozłowski, 2005).

2.2.5 Threshold Level

Thresholding allows the data itself to decide which wavelet coefficients are significant. There are some types of thresholding, such as hard thresholding is a 'keep' or 'kill' rule that has discontinuous function, soft thresholding is a 'shrink' or 'kill' rule that has a continuous function. In addition to these two types, there are many others, such as semi-soft shrinkage (firm shrinkage), mid shrinkage, non-negative Garrote, etc. and as long as the shrinkage function preserves the sign ($\text{sign}(\delta_\lambda(x)) = \text{sign}(x)$), and shrinks the magnitude ($|\delta_\lambda(x)| \leq |x|$) one can expect a denoising effect (Antoniadis, 2007).

Mathematically wavelet coefficients are estimated using the following thresholding levels (Kozłowski, 2005):

2.2.5.1 Hard Shrinkage Function:

$$\delta_{\lambda}^H(x) = \begin{cases} 0 & |x| \leq \lambda \\ x & |x| > \lambda \end{cases} \dots (17)$$

Subtracting this correction reduces those wavelet coefficients of the wavelet time series, which exceed threshold value, to zero.

2.2.5.2 Soft Shrinkage Function:

$$\delta_{\lambda}^S(x) = \begin{cases} 0 & |x| \leq \lambda \\ x - \lambda & x > \lambda \\ \lambda - x & x < -\lambda \end{cases} \dots (18)$$

By subtracting δ_{λ}^S correction considered wavelet time series coefficients are reduced to λ for positive coefficients and to $-\lambda$ for negative ones.

3. Main Results and Discussions

Nowadays, fluctuations of crude oil prices have great impacts on the budgets of OPEC exporting countries. Therefore, it is a very important to find the best forecasting model to determine future prices for crude oil in order to create optimum plans for these countries to increase their economic growth while on the other hand, attempting to determine the best forecasting ARIMAX models with and without bivariate wavelet filtering for time series data.

This section will be divided into sub-sections that are related to data analysis for the comparison of the effects of bivariate wavelets on time series forecasting of the yearly crude oil prices for OPEC organization. Depending on the objectives that were stated in this research, linear dynamic systems have been used in this section to make the forecasts before and after treating of data contamination problem. The obtained data were processed and analyzed through some packages, such as MATLAB & SIMULINK R2018b, SPSS v25, EasyFit v5.4, and Microsoft Excel 2016.

3.1 Data collection and Information about OPEC

The yearly crude oil prices and production of OPEC organization countries have been extracted from the U.S. Energy Information Administration (EIA) organization that is working as an international agency for collecting, analyzing, and publishing worldwide energy information (EIA, 2019). The yearly time series for the prices and production starts from the period of 1973 to 2018. The yearly crude oil price is in US dollar per barrel (Dollar/d) and the yearly production is in million barrels per day (Mb/d) for all OPEC members together.

Crude Oil is one of the most important raw materials in the Middle East. Everyday people are using hundreds of things that are made from oil or gas, such as petrol for automobiles, gas for cooking and heating, electric power supply, street pavement, clothes, etc. Therefore, the production and prices of crude oil products are considered as essential economic and political information needed for governments and citizens nowadays in order to organize their plans and intentions for the future.

One of the most important organizations in the field of crude oil production and exporting is the OPEC. This organization was created in Baghdad on September, 1960 with a total of 5 founding members (Iraq, Iran, Kuwait, Saudi Arabia, and Venezuela). It is a permanent, intergovernmental Organization consisting of the following countries²: Iraq, Iran, Kuwait, Saudi Arabia, Venezuela, Libya, United Arab Emirates, Algeria, Nigeria, Ecuador, Angola, Gabon, Equatorial Guinea, and Congo. The Purpose of establishing this organization is to merge oil policies among the members of these countries, so as to keep prices stable and steady supply of crude oil to importing countries (OPEC Organization, 2019).

3.2 Data Presentation and Processing

Presentation of any time series data is very important at the beginning of time series analysis in order to discover its properties, such as trend, seasonality and cyclical fluctuations. Therefore, it is clear from the time series plot that the crude oil price and production data are not stationary, i.e. the data have trend and an overall fluctuation. As seen in Figure 6, ADF test will be used to identify the stationary of these time series data.

² Qatar joined OPEC in 1961, but cancelled its membership in January 2019. Indonesia joined OPEC in 1962, deactivated its membership in 2009, reactivated it in 2016, and again decided to deactivate it in 2016.

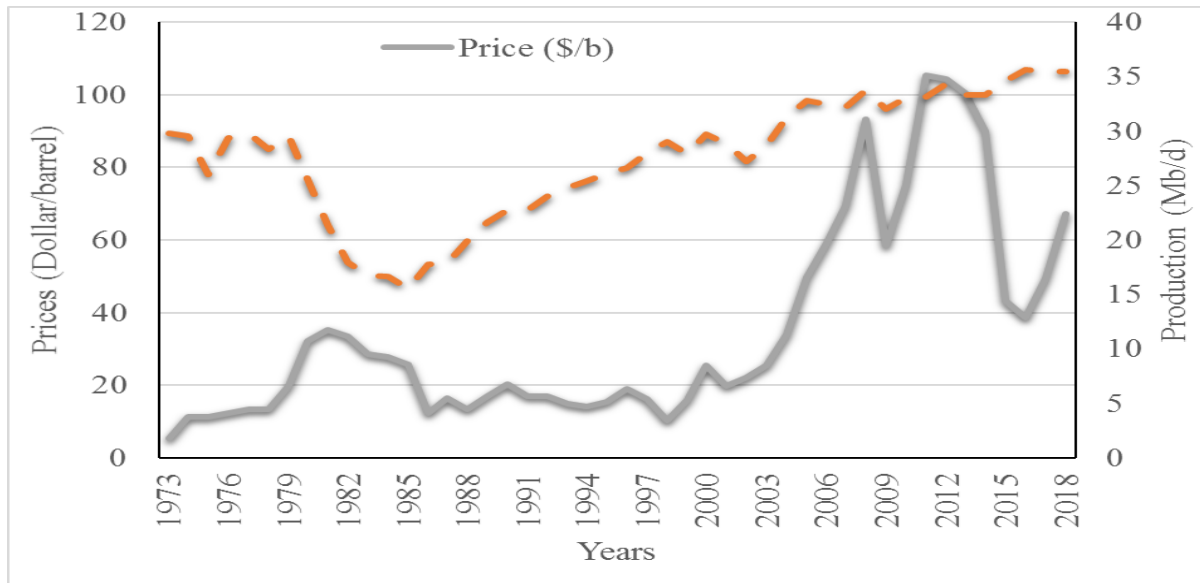


Figure (6): Yearly Crude Oil Prices and Production of OPEC Organization (1973 – 2018)

After observing the time series plot of the data, it is clear that they are not stationary. In order to be confident, it is better to use ADF test for determining whether the time series data is stationary or not, that is, the test of the following hypothesis:

H_0 : The time series has a unit root (Non-stationary)

H_1 : The time series does not have a unit root (Stationary)

Table 1 shows the ADF test results where the p-value of both variables is greater than 0.05, i.e. not rejecting of null hypothesis (None of the two variables are significant). Thus, the two- time series are not stationary and should make transformation to convert them to stationary through differencing, detrending, or any other method.

Different kinds of transformation are used and, at the end the best transformation method for the two time series to be stationary was detrending method. The ADF test was reused and reveals that the two variables are stationary because the p-values were less than 0.05, i.e. rejection of unit root hypothesis, as presented in Table 1.

Table (1): ADF test for Crude Oil Price and Production data of OPEC (before and after Detrending)

Time series	ADF test	
	P-value before Detrending	P-value after Detrending
Crude Oil Production (X)	0.785	0.0326
Crude Oil Prices (Y)	0.525	0.0300

It is obvious from figure 7 that the 2 variables are stationary in spite of cyclic variations.

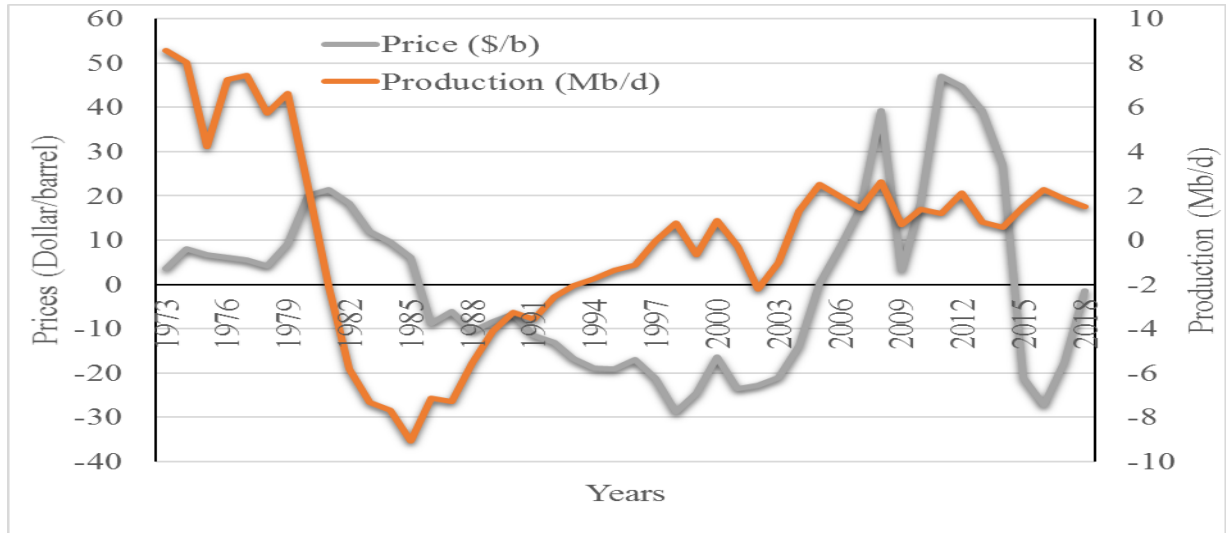


Figure (7): Detrended Yearly Crude Oil Prices and Production of OPEC Organization (1973 – 2018)

3.3 Model Identification and Fitting

In this section, model identification through System Identification Toolbox (SIT) will be used in order to forecast the time series data using traditional and proposed methods.

3.3.1 ARMAX Model Identification without Bivariate Wavelet Filtering

In this stage of LDS, the researcher will try to identify the best model and fit it to the processed data using SIT toolbox in MATLAB. The following steps are taken into consideration to find the best model:

- 1) Normality testing of model's variables (Input and output variables) using EasyFit program.
- 2) The processed input and output data are entered to the SIT;
- 3) Selecting time domain data type for model identification;
- 4) Estimation of the parameters of polynomial model;
- 5) SIT will automatically select the order of the model (na, nb, nk) for ARX models;
- 6) Choosing the best model through the accuracy criteria, such as MSE, AIC, MAPE, etc.;

- 7) Finally, drawing ACF for model's errors and CCF for input variable and model's errors in order to validate the adequacy of the model and randomness of errors.
- 8) At the end, normality testing of model's errors using EasyFit program.

The above steps will be applied for ARX models without undertaking bivariate wavelets filtering of the datasets. In case of undertaking the data filtering, it will be conducted after the first step.

The normality test for input variable (Crude oil production) using 3 goodness of fit tests: Kolmogorov-Smirnov test, Anderson-Darling test, and Chi-Squared test. It appears from all these tests that the input variable is distributed normally at $\alpha = 0.01$, i.e. could not reject null hypothesis. Concerning the normality test for output variable (Crude oil price), it is also normally distributed at $\alpha = 0.01$, meaning not rejecting null hypothesis.

Based on ARX model parameters and MATLAB's SIT³, from 1000 models, the system selected best-fit models according to the accuracy criteria and number of parameters (MATLAB will present best-fit models that the number of their parameters do not exceed 20). The name of the parametric model contains the model type, poles numbers, zeros, and time delays. For instance, the ARX10106 model is an ARX model with $n_a = 10$, $n_b = 10$, and a delay of 6 samples.

Based on the model identification results, seen in Table 2, best ARX model is ARX (10,10,6) with a fitting percentage criterion equal to 70.22%, MSE – 34.32, FPE – 163, and AIC – 353.18. It appears from FPE column in Table 2 that the best ARX model is not ARX (10,10,6) but ARX (10,2,9).

Table (2): Best 10 ARX Models According to Accuracy Criteria

Models	AIC	FPE	MSE	Fitting
ARX(10, 10, 6)	353.18	163.00	34.32	70.22
ARX(10, 9, 6)	353.93	160.73	36.43	69.32
ARX(9, 9, 6)	362.13	182.48	47.49	64.97
ARX(10, 3, 8)	358.57	156.94	52.31	63.24
ARX(10, 2, 9)	356.91	149.31	52.70	63.10
ARX(9, 7, 6)	365.46	188.31	55.70	62.07
ARX(9, 8, 7)	368.88	206.81	57.45	61.48
ARX(9, 5, 6)	365.10	180.86	60.29	60.54
ARX(9, 6, 7)	369.75	203.20	63.86	59.38
ARX(9, 2, 9)	364.90	173.61	68.39	56.97

³ For more information about SIT, see (Ljung, 2016).

Therefore, after estimating the parameters of the ARX(10, 10, 6) model through MATLAB, it could be written as presented in the following formula:

$$\phi(B)y_t = \psi(B)x_t + \epsilon_t$$

where: $\phi(B) = 1 - 0.6312B + 0.2592B^2 - 0.0778B^3 + 0.4692B^4 - 0.3267B^5 - 0.5895B^6 + 0.3106B^7 - 0.0112B^8 + 0.4886B^9 + 0.4225B^{10}$

and

$$\begin{aligned}\psi(B) = & 2.469B^6 + 0.7093B^7 - 0.7222B^8 - 4.703B^9 + 3.554B^{10} \\ & + 0.1538B^{11} + 1.073B^{12} - 1.051B^{13} + 1.204B^{14} \\ & + 1.052B^{15}\end{aligned}$$

The above selected ARX model has random residuals and uncorrelated residual values with the exogenous input variable, as presented in Figures 8 and 9. It is clear from Figure 8 that the autocorrelation values within the confidence intervals, and also Figure 9 describes the cross-correlation of these residuals with the input variable. Both correlations are within acceptable intervals.

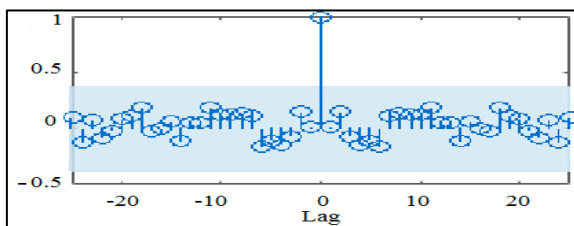


Figure (8): Residual's Autocorrelation for ARX(10, 10, 6) Model

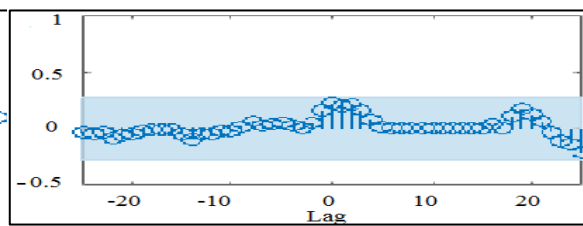


Figure (9): Cross-Correlation Residuals for ARX(10,10,6) Model with of Input Variable

Now, it is possible to test the error term for normality test. It appears that the normality test for it using 3 goodness of fit tests: Kolmogorov-Smirnov test, Anderson-Darling test, and Chi-Squared test is distributed normally at $\alpha = 0.01$.

3.3.2 ARX Model Identification with Bivariate Wavelet Filtering

After finding an appropriate ARX model for the time series data which is related to crude oil price for OPEC with the help of exogenous input variable (Crude oil production), the researcher will try to use a proposed method for cleaning contaminated data (Denoising data) through appropriate bivariate wavelet filters for the time series data using MATLAB, and then forecasting it.

Many bivariate wavelet filters are used for the stationary two datasets such as Haar wavelet, Discrete Meyer wavelet (demy), Daubechies

wavelet, and Coiflets (coif) wavelet. At the same time, many Threshold levels are used with these wavelet methods, such as Fixed Form threshold, SURE Threshold, Cross-Validation, and Minimax threshold. Also, for each threshold level there are soft and hard thresholding rules. That means that more than 30 different combinations of bivariate wavelet filters are used on the two datasets, and then using ARX models on the filtered datasets (transformed data).

The most appropriate bivariate wavelet method is Haar wavelet method with fixed form thresholding level and soft thresholding rules. In order to start applying ARX models, the stationary test is used to be confident that the 2 datasets are still stationary. The ADF test was used and reveals that the two datasets are stationary because the p-values were less than 0.05, i.e. rejection of unit root hypothesis.

The normality test for input variable (Crude oil production) after undertaking wavelet filtering for it using 3 goodness of fit tests: Kolmogorov-Smirnov test, Anderson-Darling test, and Chi-Squared test, it appears from all these tests that the input variable is distributed normally at $\alpha = 0.01$. Concerning the normality test for the output variable (Crude oil price), it appears that it is normally distributed at $\alpha = 0.01$, meaning do not rejecting null hypothesis.

As presented in Figure 10, the Bar that is highlighted on the plot in red colour indicates a type of best-fit criteria that minimizes the sum of the squares of the difference between the validation data output and the model output.

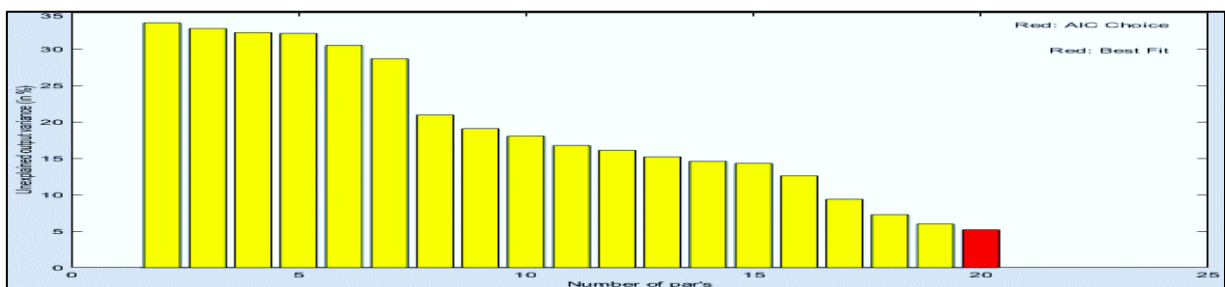


Figure (10): Best ARX Models according to Accuracy Criteria and Number of Parameters after Bivariate Wavelet filtering

Based on the model identification results, the best ARX model is ARX (10,10,8) with a fitting percentage criterion equals to 77.39%, MSE – 11.03, FPE – 52.39, and AIC – 300.97. Therefore, after estimating the parameters of the ARX(10, 10, 8) model through MATLAB, it could be written as presented in the following formula:

$$\phi(B)y_t = \psi(B)x_t + \epsilon_t$$

where: $\phi(B) = 1 - 0.6288B - 0.3186B^2 + 0.278B^3 + 0.3661B^4 - 0.3176B^5 - 0.7228B^6 + 0.5012B^7 + 1.058B^8 - 0.7302B^9 + 0.3623B^{10}$

and

$$\psi(B) = 3.123B^8 - 1.773B^9 - 1.183B^{10} + 0.6994B^{11} + 0.7898B^{12} - 0.6128B^{13} - 0.4608B^{14} + 0.438B^{15} + 3.285B^{16} - 1.955B^{17}$$

Among more than **30,000** different ARX models with bivariate wavelets filtering, the researcher arrived to the above selected model which is the best model. The selected ARX model has a random residuals and uncorrelated residual values with the exogenous input variable, as presented in Figures 11 and 12. It is clear from Figure 11 that the autocorrelation values are within the confidence intervals, and also Figure 12 describes the cross-correlation of these residuals with the input variable. Both correlations are within acceptable intervals.

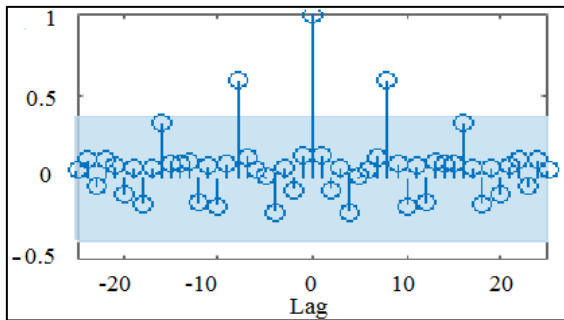


Figure (11): Residual's Autocorrelation for ARX(10, 10, 8) Model

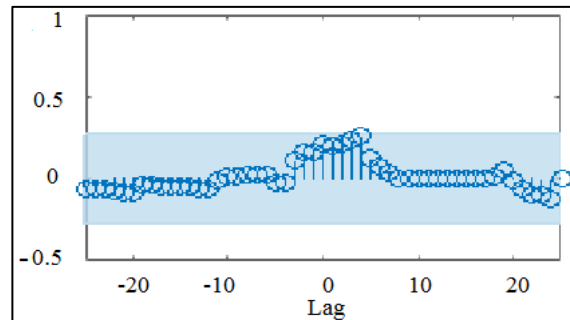


Figure (12): Cross-Correlation Residuals for ARX(10, 10, 8) Model with of Input Variable

Now, it is possible to test the error term with wavelet filtering for normality test. The normality test for it using 3 goodness of fit tests: Kolmogorov-Smirnov test, Anderson-Darling test, and Chi-Squared test, is distributed normally at $\alpha = 0.01$.

4. Forecasting Performance

All the previous sections have been done in order to discover an appropriate ARX model for the time series data. But the process is not finished at this point but the last step remains which is related to forecasting of output data. After model identification and validation is executed through accuracy criteria, the forecast for crude oil prices with the help of production input series are presented in Table 3. It shows the forecasting of the OPEC prices through two identified ARX models with

and without bivariate models. The forecasts for the ARX(10, 10, 6) are very high in 2019, then will decrease in 2020. While in the proposed ARX (10, 10, 8) model the forecasts in 2019 are moderate compared to the prices in world markets, where the forecast prices are between 59 and 69 USD.

Table (3): Forecasting of Crude Oil prices of OPEC Organization through Determined Dynamic Models

Year	ARX(10, 10, 6) without Wavelet Filtering	ARX(10, 10, 8) with Haar Wavelet Filtering
2019	92.77	68.37
2020	68.01	59.20

From comparison of the above 2 models and looking at the accuracy criteria in Table 4, it is concluded that the proposed new forecasting ARX model with bivariate wavelet filtering is better than traditional ARX model and even the forecasting values in 2019 and 2020 in the proposed model are better than traditional ARX model.

Table (4): Accuracy Criteria Comparison between Traditional and Proposed Dynamic models (with and without Bivariate Wavelet Filtering)

Models	Models	AIC	FPE	MSE	Fitting
Traditional ARX Model	ARX(10, 10, 6)	353.18	163.00	34.32	70.22
Proposed New Model	ARX(10, 10, 8)	300.97	52.39	11.03	77.39

Table 4 shows that the values of accuracy criteria for the proposed new model ARX(10, 10, 8) are lower than the traditional ARX(10, 10, 6) model and even the fitting value is more than it by 7 percent. Therefore, the proposed new model is reliable and can be used for planning purposes by OPEC countries or any other entities.

5. Conclusions and Recommendations

Depending on the results of the analyses in section four, the most important conclusions are as follows:

1. The ARX model successfully benefited from bivariate wavelet filtering, thus, the proposed ARX model is more powerful than the traditional ARX model.
2. The forecasts of OPEC crude oil prices in the next two years using proposed model were more appropriate than traditional models.
3. ARX models have more forecasting parameter options than ARIMA models that allow the selection of more accurate data models, as well as the impact of the independent variable on the dependent variable.
4. It appeared that detrending was better than differencing to convert applied time series data to stationary when using bivariate wavelet filtering.
5. It is concluded from the estimated model that the average crude oil prices in 2020 will be fairly less than 2019.

Based on the conclusions given in previous section, the following major recommendations are drawn:

1. Strongly recommending OPEC administration and related institutes to use the proposed forecasting model for crude oil prices.
2. Applying other input exogenous variables that have effects on crude oil prices of OPEC Organization rather than crude oil production.
3. Applying other kinds of dynamic systems in forecasting multiple time series data, such as ARIMAX, BJ, and Output Error methods, and comparing them with filtering models using Haar wavelets.
4. Applying other kinds of linear and nonlinear dynamic systems in forecasting multiple time series data.
5. Since the proposed method was better than traditional methods, it is recommended to apply multivariate dynamic systems with multivariate wavelet filtering.

References:

1. أبولبدة، محمد عبدالرحمن جادالله ، 2017. إستخدام نماذج ARIMAX في التنبؤ بالسلاسل الزمنية. غزة: جامعة الأزهر (رسالة ماجستير) .
2. الزبيدي، طه حسين علي ، 2009. إستخدام مرشح الموجة الصغيرة المتقطعة في تحليل السلسلة الزمنية (AR(1) ومقارنته مع مرشحات أخرى. بغداد: جامعة بغداد (أطروحة دكتوراه غير منشورة).
3. الشرايبي، نجلاء سعد إبراهيم ، 2017. دراسة لأساليب تقطيع البيانات والتقدير المتكرر في تشخيص النظم الخطية مع التطبيق. نينوى: جامعة موصل (أطروحة دكتوراه غير منشورة).

جمعة، أحلام أحمد؛ شعيث، أنمار رسول ، 2017. استخدام أنموذجي ARX و ARMAX للتنبؤ بالسلاسل الزمنية مع تطبيق عملي - دراسة مقارنة. مجلة كلية الرافدين الجامعة للعلوم، 40(1681-6870)، pp. 64 - 93.

4. عبدالقادر، قيس مصطفى، 2011. مقارنة نماذج بوكس - جينكنز قبل وبعد الترشيح المويجي من حيث تخفيض الرتب مع التطبيقات. أربيل: جامعة صلاح الدين - أربيل (رسالة ماجستير غير منشورة).

5. Akpanta, A. C. & Okorie, I. E., 2014. Application of Box-Jenkins Techniques in Modelling and Forecasting Nigeria Crude Oil Prices. International Journal of Statistics and Applications, 4(6).

6. Boggess, A. & Narcowich, F. J., 2009. A first course in wavelets with fourier analysis. 2nd ed. s.l.: John Wiley & Sons.

7. Brockwell, P. J. & Davis, R. A., 2016. Introduction to Time Series and Forecasting. 3rd ed. s.l.: Springer.

8. Burrus, C. S., Gopinath, R. A. & Guo, H., 1998. Introduction to Wavelets and Wavelet Transforms: A Primer. New Jersey: Prentice Hall .

9. Chan, Y. T., 1995. Wavelet Basics. s.l.: Springer Science & Business Media.

10. Cheng, D.-C., 2011. Haar Wavelet Analysis. [Online] Available at: <https://www.cyut.edu.tw/teacher/ft00002/Wavelet1.pdf> [Accessed 14 3 2019].

11. Chui, C. K., 1992. An Introduction to Wavelets. London: Academic Press.

12. Cochrane, J. H., 2005. Time Series for Macroeconomics and Finance. Chicago: University of Chicago.

13. Cryer, J. D. & Chan, K.-S., 2008. Time Series Analysis (With Applications in R). 2nd ed. New York: Springer Science.

14. Debnath, L. & Shah, F. A., 2017. Lecture Notes on Wavelet. Cham: Springer International Publishing.

15. Gencay, R., Selcuk, F. & Whitcher, B., 2002. An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. London: Academic Press.

16. Gutiérrez, E., 2005. Time series analysis of European Grid Blackouts, s.l.: European Laboratory for Structural Assessment (ELSA).

17. Hamilton, J. D., 1994. Time Series Analysis. New Jersey: Princeton University Press.

18. Hyndman, R. J. & Athanasopoulos, G., 2014. Forecasting: Principles and Practice. s.l.: Otexts.com.

- 19.Kaiser, G., 2011. A friendly guide to wavelets. s.l.:Springer Science & Business Media.
- 20.Kozłowski, B., 2005. Time series denoising with wavelet transform. journal of telecommunication and information technology, Volume 3, pp. 91-95.
- 21.Lee, M.-T., 2010. Short-Term Freeway Traffic Flow Forecasting With ARIMAX Modeling. s.l.:University Carolo-Wilhelmina (MSc Thesis).
- 22.Ljung, L., 1999. System Identification: Theory for the User. 2nd ed. New Jersey: Prentice Hall PTR.
- 23.Ljung, L., 2001. Black-box Models from Input-output Measurements. Budapest, Hungary, EEE Instrumentation and Measurement Technology Conference, pp. 138-146.
- 24.Mills, T. C., 2011. The Foundations of Modern Time Series Analysis. s.l.:Palgrave and Macmillan.
- 25.Montgomery, D. C., Jennings, C. L. & Kulahci, M., 2008. Introduction to Time Series Analysis and Forecasting. New Jersey: JOHN WILEY & SONS.
- 26.Morettin, P. A., Pinheiro, A. & Vidakovic, B., 2017. Wavelets in Functional Data Analysis. s.l.:Springer.
- 27.Nelles, O., 2001. Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models. New York: Springer-Verlag Berlin Heidelberg.
- 28.Pankratz, A., 1983. Forecasting with Univariate Box-Jenkins Models: Concepts and Cases. s.l.:Wiley series.
- 29.Percival, D. B. & Walden, A. T., 2000. Wavelet methods for time series analysis. s.l.:Cambridge university Press.
- 30.Raimondo, M., 2002. WAVELET SHRINKAGE VIA PEAKS OVER THRESHOLD. [Online]
Available at:
<https://pdfs.semanticscholar.org/885a/5dc1cda0be465e8d91f619f6acdce5b2fc1a.pdf> [Accessed 2 4 2019].
- 31.Saaed, T. E., 2015. Structural Control and Identification of Civil Engineering Structures. s.l.:(PhD Thesis) Luleå University of Technology.
- Sheng, Y., 2000. The Transforms and Applications Handbook: Second Edition: Wavelet Transform. Boca Raton: CRC Press LLC.
- 32.Shumway, R. H. & Stoffer, D. S., 2011. Time Series Analysis and Its Applications. 3rd ed. Springer.
- 33.Spyros G. Makridakis, S. C. W. R. J. H., 1998. Forecasting: Methods and Applications. s.l.:s.n.

34. Stankovic, R. S. & Falkowski, B. J., 2003. The Haar wavelet transform: its status and achievements. Computers and Electrical Engineering , Volume 29, p. 25–44.
35. UNDP/IRAQ, n.d. Reconstruction of the Kurdistan Region's Electricity Sector. [Online]
Available at:
http://www.iq.undp.org/content/iraq/en/home/operations/projects/poverty_reduction/kr-electricity-reconstruction.html [Accessed 29 11 2017].
36. Vidakovic, B., 1999. Statistical Modeling by Wavelets. New York: John Wiley & Sons.
37. Wei, W. W. S., 2006. Time Series Analysis: Univariate and Multivariate Methods. 2nd ed. s.l.: Pearson Education.
38. Zoubir, L., 2017. A brief history of time series analysis. [Online]
Available at: <https://www.statistics.su.se/english/research/time-series-analysis/a-brief-history-of-time-series-analysis-1.259451> [Accessed 24 1 2019].