Use Box-Jenkins Model in domestic product (GDP) at Current Price in KSA Electricity, Gas and Water

D. Abuzar Yousef Ali Ahmed , 00966546530264
Assistant Prof ,Department of Basic Science, Faculty of preparatory Years King Saud University, Saudi Arabia, abuzarjeha@gmail.com

الملخص:

في هذا البحث تم استخدم نماذج السلاسل الزمنية باستخدام منهجية بوكس – جنكنز لبيانات الناتج المحلي الإجمالي بالسعر الحالي للكهرباء والغاز والمياه في المملكة العربية السعودية .

أظهرت النتائج أن النموذج هو النموذج المناسب لسلسلة نموذج أريما هو: ARIMA (٤،٠،٠)

وفقًا لنتائج التقدير لهذا النموذج ، نلاحظ التوافق بين القيم الحقيقية والمقدرة مما يشير إلى قوة النموذج والقدرة على التنبؤ.

حيث تمت مقارنة النموذج مع العديد من نماذج السلاسل الزمنية ، وحقق كل الشروط الخاصة بالسلاسل الزمنية ، حيث اثبت قدرته العالية على التنبؤ ، وان قيمه التنبؤية مشابهة ومقاربة للقيم الأصلية ويعتبر الأفضل من بين كل النماذج المختارة .

بالنسبة للإحصاءات الوصفية للنموذج ، يمثل R-squared معامل التحديد ٠.٩٣ وهذا يعني أن النموذج يمثل البيانات تمامًا (نموذج جيد).

ومن تقديرًا لمعاملات النموذج ، من النموذج نلاحظ أن مستوى الأهمية Sig = 0.00. أقل من ٠.٠٥ ، مما يشير إلى أن المعاملات ذات دلالة إحصائية ، وفعالة أيضًا ويمكن التنبؤ بها..

Abstract:
In this research, use the time series models to Create Box-Jenkins Model Gross domestic product (GDP) at current price in KSA Electricity, Gas and Water.

The results showed that the model is the appropriate model for the series of Arima Model Gross domestic product (GDP) at current price in KSA Electricity, Gas and Water

is :ARIMA (4,0,0)

According to the estimation results of this model, we observe the compatibility between observed and estimated values as these values are consistent with those in the original time series, indicating the strength of the model and predictability.

We see the agreement between the real and estimated values in light of the model's estimation findings, which highlights its predictive strength.

The model is regarded as the best among all the selected models since it outperformed all the requirements for time series, had a high level of predictive ability, and has predicted values that are comparable to and close to the original values.

For the descriptive statistics of the model, R-squared represents the coefficient of good fit if the value is greater = 0.93 more than 0.05 this mean the model represent data exactly (good model) .

This table provides an estimate of the coefficients of the model, from the model we note that the level of significance Sig= 0.00. Less than 0.05, which indicates that the coefficients are statistically significant, also effective and predictable .

Key words domestic product; Box-Jenkins ; Time Series Modeler ; stationary; identification ; estimation

.

1. Introduction

The Autoregressive Integrated Moving Average (ARIMA) model, which is fitted utilizing a method created by George Box and Gwilym Jenkins, is one form of model that does account for autocorrelation (1970). Although ARIMA models have a wide range of applications in the health sector, they have been widely employed for I infectious disease outbreak identification and (ii) the evaluation of population-level health interventions using interrupted time series analysis. Both of these approaches call for the formal characterization of the underlying pattern in a time series and the use of this pattern to predict the time series' future behavior. We predict the 95% confidence interval for a time series and the deviation of the actual time series for outbreak detection.

A signal would be data that fall within the 95% confidence interval. A public health intervention is thought to have a causal effect when actual values differ from anticipated values in an interrupted time series, where the time series is forecast into the future.

Nota Bene: • The causal framework for the ARIMA model differs slightly from the epidemiology frame and is more consistent with the Granger definition of a cause from economics. • ARIMA models do NOT predict rare "black swan" events, as there is no pattern in the time series to suggest a future event of this type.

Arima models

ARIMA stands for Autoregressive Integrated Moving Average. It is a type of time series forecasting model that uses past data points to make predictions about future values. ARIMA models are used to analyze and forecast time series data such as sales, stock prices, and other economic indicators. ARIMA models are based on the assumption that the underlying process generating the data is stationary, meaning that the mean and variance of the data remain constant over time. The model uses three parameters to capture

The following data is needed to fit an ARIMA model:

• A countable or continuous univariate time series with at least 50–100 observations

In the event that the time series contains count data, If the time series is made up of continuous data, the interval between measurements must also remain constant across time. the interval during which the count is taken must remain the same.

• A vertical vector must be used to convey the data (column of data)

1.1 What is can be forecast?

Forecasting is the process of making predictions about the future based on past and present data. Forecasts can be made for a variety of topics, including weather, economic trends, stock market performance, consumer behavior, and political events.

Many different circumstances call for forecasting, including selecting whether to construct a new power plant in the next five years, arranging workers for a call center the following week, and stocking an inventory. Predictions may be necessary months or even years in advance (for capital projects), or even only a few minutes before (for telecommunication routing). Forecasting is a crucial tool for effective and efficient planning, regardless of the situations or time frames involved.

Predicting certain things is simpler than others. Next morning's sunrise timing can be predicted with accuracy. The lottery numbers for tomorrow, however, cannot be predicted with any degree of accuracy. The likelihood of an

1.2 1.2 forecasting, planning and gools setting

Forecasting is the process of making predictions about the future based on past and present data. Forecasting can be used to help businesses plan for the future by providing an estimate of future demand for products and services. Planning is the process of setting goals and objectives, and then developing strategies to achieve those goals. Planning helps businesses to identify potential risks and opportunities, and to develop strategies to take advantage of them. Goals setting is the process of defining what a business

In business, forecasting is a common statistical task that aids in decision-making regarding the scheduling of production, transportation, and staff as well as serving as a roadmap for long-term strategic planning. Yet, corporate forecasting is typically performed ineffectively and is frequently mistaken for planning and objectives. They each represent a distinct entity.

Forecasting

is about making the most accurate projections possible given all the information at our disposal, including past data and knowledge of any potential future events.

Your desired outcomes are your goals. Objectives and forecasts should be connected, but this doesn't always happen. Too often, objectives are established without a strategy for achieving them and without forecasts for whether they are realistic.

Planning

is a reaction to projections and objectives. Planning entails choosing the best course of action to take in order to make your forecasts and goals align.

Since forecasting can be useful in many aspects of a company, it should be a fundamental component of management's decision-making processes. Depending on the individual application, modern businesses need forecasts for the short term, medium term, and long term.

For the scheduling of staff, manufacturing, and transportation, short-term projections are required. Forecasts of demand are frequently needed as part of the scheduling process.

For the purpose of purchasing raw materials, hiring workers, or purchasing machinery and equipment, medium-term predictions are necessary to assess the future resource requirements.

Strategic planning makes use of long-term projections. Market prospects, environmental factors, and other considerations must be taken into

1.3 determining what to forecast

When determining what to forecast, it is important to consider the purpose of the forecast. Different forecasting techniques are better suited for different types of data and different goals. For example, if the goal is to predict future sales, then a time series forecasting technique may be more appropriate than a regression analysis. Additionally, it is important to consider the availability of data and the accuracy of the forecast. If there is limited data available or if the forecast needs to be highly accurate, then more sophisticated forecasting techniques

Making decisions regarding what should be forecasted is necessary in the early phases of a forecasting endeavor. If predictions are necessary, for instance, for products in a manufacturing setting, it is vital

determine whether forecasts are required for:

1. Is it for each product line or for categories of products?

2. For each and every sales outlet, for outlets categorized by region, or just for overall sales?

3. Are the data weekly, monthly, or yearly?

Moreover, the predicting horizon must be taken into account. Will forecasts be needed for the next month, the next six months, or the next 10 years? Several model types will be required, depending on which forecast horizon is most crucial.

How often are forecasts necessary? necessary forecasts

It is preferable to use an automated system than to use labor-intensive manual procedures for items that need to be produced regularly.

Before putting a lot of effort into creating the predictions, it is important to take the time to speak with the people who will use them to make sure you understand their needs and how the forecasts will be used.

Finding or gathering the data on which the forecasts will be based is then necessary after it has been determined what forecasts are necessary. It's possible that the forecasting data already exist. Nowadays, there are many data records, therefore the forecaster's responsibility is frequently to determine where and how the necessary data are stored. The information could consist of a company's sales records, the historical Purchasing power for a good or a region's unemployment rate. Before creating appropriate forecasting methods, a forecaster may spend a significant amount of time locating and gathering the available data.

1.4 forecasting data and method

Forecasting data is data that is used to predict future events or trends. This data can come from a variety of sources, including economic indicators, consumer surveys, and historical data. Forecasting methods are the techniques used to analyze and interpret the data in order to make predictions. Common forecasting methods include time series analysis, regression analysis, and Monte Carlo simulations.

The right forecasting techniques mostly depend on the data at hand.

Qualitative forecasting techniques must be employed if there are no data available or if the data that are available are not pertinent to the forecasts. There are well-developed organized procedures to obtaining accurate projections without using past data, thus these methods are not just educated guesswork. Chapter 4 discusses these techniques.

When both of the following conditions are met, quantitative forecasting can be used:

There is numerical data about the past that is available, and it is plausible to predict that some patterns from the past will persist into the future.

There are many different quantitative forecasting techniques, many of which were developed inside certain disciplines for particular objectives. Each approach has unique characteristics, accuracies, and expenses that should be taken into account while selecting a certain strategy.

Most quantitative prediction problems employ cross-sectional data or time series data, which are gathered at regular periods over time (collected at a single point in time). We focus on the time series domain in this book because we are interested in predicting future data.

1.5 time series forecasting

Time series forecasting is the process of using historical data to predict future values of a time series. It is a type of predictive analytics that uses time-series data to forecast future trends and patterns. Time series forecasting can be used in a variety of applications, such as predicting stock prices, sales, and customer demand. It can also be used to forecast weather patterns, economic trends, and other types of data. Time series forecasting is an important tool for businesses to make informed decisions about the future.

A statistical technique called time series analysis is used to examine and model the trends and patterns in a collection of data points that have been gathered over time. It is frequently used to analyze data trends and forecast future outcomes in disciplines including economics, finance, and environmental science.

In time series analysis, a number of approaches are utilized, such as:

Decomposition: This entails dissecting the time series into
such as trend, seasonality, and residuals, which are its constituent pieces.

To better visualize trends and patterns, smoothing includes employing math ematical techniques to remove irregularities from time series data.
To forecast future events, ARIMA modeling requires fitting an autoregressi ve integrated moving average (ARIMA) model to time series data.
A prominent technique for predicting future values in a time series by utilizi ng a weighted average of previous observations is exponential smoothing.
The Fourier analysis includes converting a time series into the frequency do main in order to pinpoint the frequencies of various data components.
Although time series analysis is an effective technique for identifying trend s in data and making predictions, it can also be difficult and complex.
It's crucial to A method should be carefully chosen, and the conclusions should be interpreted in light of the chosen approach's constraints and underlying assumptions.
Yes, that is true! Time series analysis is an effective tool, but in order to properly interpret the results, it's also critical to comprehend the assumptions and constraints of each technique.
Making ensuring that the data being studied is stationary—that is, that its mean and variance remain consistent throughout time—is also crucial.
Using methods like differences or transformations, such the log transformation, it is possible to make non-stationary data stationary.
Checking for outliers or abnormalities in the data is a crucial part of time series analysis because they can significantly affect
the outcomes. Statistical methods like the Z-score or the median absolute deviation can be used to identify outliers.

In conclusion, time series analysis is a useful technique for identifying trends in data over time, but it's also crucial to be aware of its constraints and underlying presuppositions, as well as to pre-process the data before beginning the analysis.
Time series data examples include: daily IBM stock prices; monthly rainfall; quarterly Amazon sales; and annual Google profits.
A time series is anything that has been observed progressively over time.
Only time series that are observed at regular periods of time will be taken into consideration (e.g., hourly, daily, weekly, monthly, quarterly, annually). Time series with irregular spacing can also happen.
Forecasting time series data requires: The goal is to predict how the observations will proceed in the future. Figure 1.1 depicts Australian beer production on a quarterly basis from 1992 to the second quarter of 2010.

Forecasts for the following two years are shown by the blue lines. See how the seasonal pattern shown in the historical data has been duplicated for the following two years by the projections. The 80% prediction ranges are displayed in the dark colored area. This means that there is an 80% chance that each future number will be found in the region of darkness. 95% prediction intervals are displayed in the area that is lightly shaded. These prediction intervals are a practical tool for showing forecast uncertainty. Because it is anticipated that the projections will be correct in this situation, the prediction intervals are fairly small.
The most basic time series forecasting techniques don't try to identify the variables that influence a variable's behavior; instead, they just use data on the variable to be forecast.
As a result, they will extrapolate trends and seasonal patterns while ignoring all other data, including marketing campaigns, competition activity, changes in economic conditions, and so forth.
Decomposition models, exponential smoothing methods, and ARIMA models are some of the time series models used for forecasting.
Forecasting using time series and predictor variables
In time series forecasting, predictor variables are frequently beneficial. For instance, let's say we want to predict the summertime hourly electricity demand (ED) for a hot region. An example of a model with predictor variables is ED=f (current temperature, strength of economy, population, time of day, day of week, error). ED=f(current temperature, strength of economy, population, time of day, day of week, error) (current temperature, strength of economy, population, time of day, day of week, error). There will always be variations in electricity demand that cannot be explained by the predictor variables, therefore the relationship is not accurate. The "error" term on the right accounts for chance variations and the impacts of pertinent factors that are not taken into account by the model. Because it explains what, we refer to this as an explanatory model.

1.6 A the basic steps in forecasting

1. Define the forecasting problem: Identify the purpose of the forecast, the time frame, and the data that will be used.

2. Select a forecasting method: Choose a forecasting technique that best fits the problem.

3. Collect data: Gather historical data relevant to the problem.

4. Analyze data: Examine the data to identify trends, seasonality, and other patterns.

5. Develop a forecast: Use the chosen forecasting method to generate a forecast.

6. Evaluate the forecast: Compare the forecast to actual results and assess its accuracy.

7. Implement and monitor the forecast: Put the forecast into action and track its performance over time.

1.7 the statistical forecasting perspective

Statistical forecasting is a method of predicting future events or trends based on past data. It involves the use of mathematical models and algorithms to analyze historical data and make predictions about future outcomes. Statistical forecasting can be used to predict a variety of outcomes, including sales, customer demand, economic trends, and more. Statistical forecasting is often used in business and economics to help make decisions about investments, production levels, and other important decisions.

1.8 BOX JENKINS MODEL

The Jenkins model is a software development model that is based on the principles of continuous integration and continuous delivery. It is an open-source automation server that can be used to automate the building, testing, and deployment of software applications. The model is based on the idea that software should be developed in small increments, with each increment tested and deployed quickly. This allows for faster feedback and more frequent releases. The model also encourages collaboration between developers, testers, and operations teams to ensure that the software is of high quality and meets customer requirements.

Fundamentals To anticipate or predict future value over a period of time (for example, stock price) is to use simple language. There are various methods for predicting the value; for instance, let's take the case of a firm XYZ that tracks website traffic hourly and wishes to anticipate how much traffic would be there in the upcoming hour. What method will you use to predict the traffic for the future hour, may I ask?
Finding the average of all observations, taking the mean of the most recent two observations, giving more weight to the current observation and less to the past, or using interpolation are just a few examples of how different perspectives can differ. There are various techniques for predicting the values.
when Predicting time series values, 3 crucial terms need to be taken care of and the major goal of time series forecasting is to forecast these three terms.

The main task of time series forecasting is to forecast these three terms:

1) Seasonality
Seasonality is a simple term that means while predicting a time series data there are some months in a particular domain where the output value is at a peak as compared to other months. for example if you observe the data of tours and travels companies of past 3 years then you can see that in November and December the distribution will be very high due to holiday season and festival season. So while forecasting time series data we need to capture this seasonality.

2) Trend
The trend is also one of the important factors which describe that there is certainly increasing or decreasing trend time series, which actually means the value of organization or sales over a period of time and seasonality is increasing or decreasing.

3) Unexpected Events
Unexpected events mean some dynamic changes occur in an organization, or in the market which cannot be captured. for example a current pandemic we are suffering from, and if you observe the Sensex or nifty chart there is a huge decrease in stock price which is an unexpected event that occurs in the surrounding.

1.9. Model Type. The following options are available:

- All models. The Expert Modeler considers both ARIMA and exponential smoothing models.
- Exponential smoothing models only. The Expert Modeler only considers exponential smoothing models.
- ARIMA models only. The Expert Modeler only considers ARIMA models.

Expert Modeler considers seasonal models. This option is only enabled if a periodicity has been defined for the active dataset. When this option is selected, the Expert Modeler considers both seasonal and nonseasonal models. If this option is not selected, the Expert Modeler only considers nonseasonal models.

Events and Interventions. Enables you to designate certain input fields as event or intervention fields. Doing so identifies a field as containing time series data affected by events (predictable recurring situations, for example, sales promotions) or interventions (one-time incidents, for example, power outage or employee strike). The Expert Modeler does not consider arbitrary transfer functions for inputs identified as event or intervention fields.

Input fields must have a measurement level of Flag, Nominal,
or Ordinal and must be numeric (for example, 1/0, not True/False, for a flag field), before they will be included in this list.

Outliers

Detect outliers automatically. By default, automatic detection of outliers is not performed. Select this option to perform automatic detection of outliers, then select the desired outlier types. See the topic Handling Outliers for more information.

Streaming TS Model Options
Handling Outliers

Univariate Series (TSMODEL algorithms)

Users can let the Expert Modeler select a model for them from:

- All models (default).
- Exponential smoothing models only.
- ARIMA models only.

2. gross domestic product (GDP)
Gross Domestic Product (GDP) is a measure of the total value of all goods and services produced within a country in a given period of time. It is used to measure the size and health of an economy and is considered one of the most important indicators of economic performance. GDP is calculated by adding up the total value of all goods and services produced in a country, including consumer spending, government spending, investments, and exports minus imports.

2.1 The standard measurement of the value added produced via the production of goods and services in a nation over a specific time period is the gross domestic product (GDP). Consequently, it also accounts for the revenue generated by that manufacturing, or the total amount spent on finished goods and services (less imports). Although GDP is the most significant indicator for measuring economic activity, it is insufficient to evaluate people's material well-being, for which other metrics may be more adequate. This indicator is based on nominal GDP, which can be measured in US dollars and US dollars per person. Nominal GDP is also known as GDP at current prices or GDP in value (current PPPs). Every OECD nation compiles theirdata from the System of National Accounts for 2008 (SNA). This indicator is less useful for comparing trends over time because developments are influenced by changes in prices and PPPs as well as real growth.

2.2 GDP per quarter

The standard measurement of the value added produced via the production of goods and services in a nation over a specific time period is the gross domestic product (GDP). Consequently, it also accounts for the revenue generated by that manufacturing, or the total amount spent on finished goods and services (less imports). Although GDP is the most significant indicator for measuring economic activity, it is insufficient to evaluate people's material well-being, for which other metrics may be more adequate. This indicator's foundationon real GDP, which is defined as GDP that is adjusted for price fluctuations over time (also known as GDP at constant prices or GDP in volume). Also, the data is seasonally adjusted. The indicator can be measured in a variety of ways, including volume index (2015=100), percentage change from the previous quarter, and percentage change from the same quarter last year. Every OECD nation gathers its data in accordance with the 2008 System of National Accounts (SNA).

2.3 Forecast for Real GDP

Real gross domestic product (GDP), which refers to the volume level of GDP, is GDP expressed at constant prices. By quantifying the values of all the goods and services produced in a given year in terms of a base period, one can derive constant price estimates of GDP. Prediction isbased on an evaluation of the global and national economies made using a combination of model-based analysis and professional judgment. Growth rates in relation to the prior year are used to measure this indicator.

2.4 Real GDP long-term forecast The trend in real GDP is shown here, along with long-term baseline predictions (up to 2060). The forecast is based on an evaluation of the global and national economies made using a combination of model-based analysis and expert opinion. This statistic is measured in USD using purchasing power parities and constant prices (PPPs) .

3.Data Analysis

3.1 Time Series Modeler

Model Description

			Model Type
Model ID	الكهرباء، الغاز والماء	Model_1	ARIMA(4,0,0)

Each estimated model has an item in the table of model descriptions, along with the type and model identifier. The name (or label) of the associated dependent variable plus a name given by the system make up the model identifier. Sales of men's clothing are the dependent variable in the current case, and Model 1 is the label given to it by the computer system.
Both ARIMA models and exponential smoothing are supported by the Time Series Modeler. The different varieties of exponential smoothing models are listed along with their common names, like Holt and Winters' Additive. The typical nomenclature for ARIMA model types is ARIMA(p,d,q)(P,D,Q), where p denotes the order of autoregression, d the order of differencing (or integration), q the order of moving average, and (P,D,Q) denotes the corresponding seasonal orders for each.
The Expert Modeler has discovered that a seasonal ARIMA model with one order of differencing best captures the sales of men's clothes. The single order of differencing reflects the increasing trend that was visible in the data, and the seasonal character of the model accounted for the seasonal peaks that we noticed in the series plot.

Model Summary

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile						
					5	10	25	50	75	90	95
Stationary R-squared	.907	.	.907	.907	.907	.907	.907	.907	.907	.907	.907
R-squared	.935	.	.935	.935	.935	.935	.935	.935	.935	.935	.935
RMSE	1181.036	.	1181.036	1181.036	1181.036	1181.036	1181.036	1181.036	1181.036	1181.036	1181.036

MAPE	9.297	.	9.297	9.297	9.297	9.297	9.297	9.297	9.297	9.297	9.297
MaxAPE	202.145	.	202.145	202.145	202.145	202.145	202.145	202.145	202.145	202.145	202.145
MAE	592.442	.	592.442	592.442	592.442	592.442	592.442	592.442	592.442	592.442	592.442
MaxAE	5951.998	.	5951.998	5951.998	5951.998	5951.998	5951.998	5951.998	5951.998	5951.998	5951.998
Normalized BIC	14.534	.	14.534	14.534	14.534	14.534	14.534	14.534	14.534	14.534	14.534

R-Squared

Goodness-of-Fit, often known as the coefficient of determination, is a measurement of a linear model's accuracy. It is the percentage of variation in the dependent variable that the regression model is able to account for. It has a value between 0 and 1. Little numbers show that the model does not adequately fit the data **(R-Squared=0.93)**.

Root Mean Square Error (RMSE)

the mean square error's square root. a measurement that uses the same units as the dependent series to express how much a dependent series deviates from the level anticipated by the model.

MAPE

Mistake in Mean Absolute Percentage. a way to gauge how much a dependent series deviates from the level predicted by the model. It can be used to compare series using various units because it is independent of the units being used.

Mean Absolute Error (MAE)

determines how far the series deviates from the level predicted by the model. The original series units have MAE information.

Maximum Absolute Percentage Error, or MaxAPE. The biggest predicted inaccuracy, in percentage form. When making forecasts, it's helpful to consider the worst-case situation.

MaxAE

Absolute Maximum Error. The biggest predicted error, expressed in the dependent series' units. It is helpful for imagining the worst-case scenario for your forecasts, much as MaxAPE. When the absolute error for a large series value is marginally larger than the absolute error for a small series value, for instance, the maximum absolute error and maximum absolute percentage error may occur at different series points. In that instance, the larger series will experience the highest absolute error.

The smaller series value is where the smaller absolute % error will occur.

BIC normalized

Bayesian information criterion normalized. a broad evaluation of a model's overall fit that strives to take model complexity into consideration. The amount of parameters in the model and the duration of the series are penalized in this score, which is based on the mean square error. The statistic is simple to compare across many models for the same series since the penalty eliminates the benefit of models with more parameters.
The Ljung-Box Q metric
An assessment of the residual mistakes in this model's randomness.
df. Variations in freedom
how many model parameters are open to change while estimating a certain target.
Value of the Ljung-Box statistic's significance
If the significance level is lower than 0.05, the residual errors are not random.
Statistical summaries. This section includes various summary statistics, such as mean, minimum, maximum, and percentile values, for the various columns.
All of the models' fit statistics are calculated and presented in the Model Fit table. It gives a succinct overview of how well the reestimated parameterized models fit the data. The table lists the mean, standard error (SE), lowest and maximum values for each statistic across all models. Additionally, it includes percentile values that show how the statistic is distributed throughout models. A percentage of models fall below the stated value for the fit statistic for each percentile. As an example, MaxAPE (maximum absolute percentage error) values for 95% of the models are less than 3.676.
We will concentrate on two statistics, MAPE (mean absolute percentage error) and MaxAPE, out of the many that are presented (maximum absolute percentage error). Absolute percentage error measures how much a dependent series deviates from the level anticipated by the model and gives you a sense of how certain your forecasts are. Among all models, the mean absolute percentage error ranges from a minimum of 0.669% to a maximum of 1.026%. All models have a maximum absolute percentage error that ranges from 1.742% to 4.373%. Thus, the greatest uncertainty (the mean value of MaxAPE) is approximately 2.5%, with a mean uncertainty in each model's predictions of about 1%.
with an absolute worst case of 4%. Your level of risk tolerance will determine if these values indicate an acceptable level of uncertainty.
R-squared, which is used to express the descriptive statistics of a model, indicates the coefficient of good fit. If the number is larger than 0.87 and greater than 0.05, a good model is one that accurately represents the data.

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	DF	Sig.	
-والماء الغاز الكهرباء، Model_1	0	.907	6.997	14	.935	0

Ljung-Box Q statistics. An assessment of the residual mistakes in this model's randomness.
df. levels of independence. how many model parameters are open to change while estimating a certain target.
Ljung-Box statistic significance value, abbreviated as Sig. If the significance level is lower than 0.05, the residual errors are not random.
The Ljung-Box test is an example of:
H0: The distribution of the data is random (i.e. the correlations in the population from which the sample is taken are 0, so that any observed correlations in the data result from randomness of the sampling process).

Ha: The data show serial correlation; they are not dispersed randomly.
Accept Null Hypothesis, Data is Random, sig greater 0.05.
Results
The modified Box-Pierce statistic, sometimes referred to as the Ljung-Box statistic, offers a clue as to whether the model is properly stated. A significance level of 0.05 or below suggests that the model cannot fully explain for the observed series' structure.
Nine points that were deemed to be outliers were found by the expert modeler. There is no need for you to remove any of these statements from the series because they have all been adequately represented.
Note: When the Sig value is more than 0.05 and the residual error test is used, the results indicate that the data are **random** and appropriate for prediction (**value of Sig. = 0.9350)**.
(Sig more 0.05 accept Null Hypothesis , Data is Random)

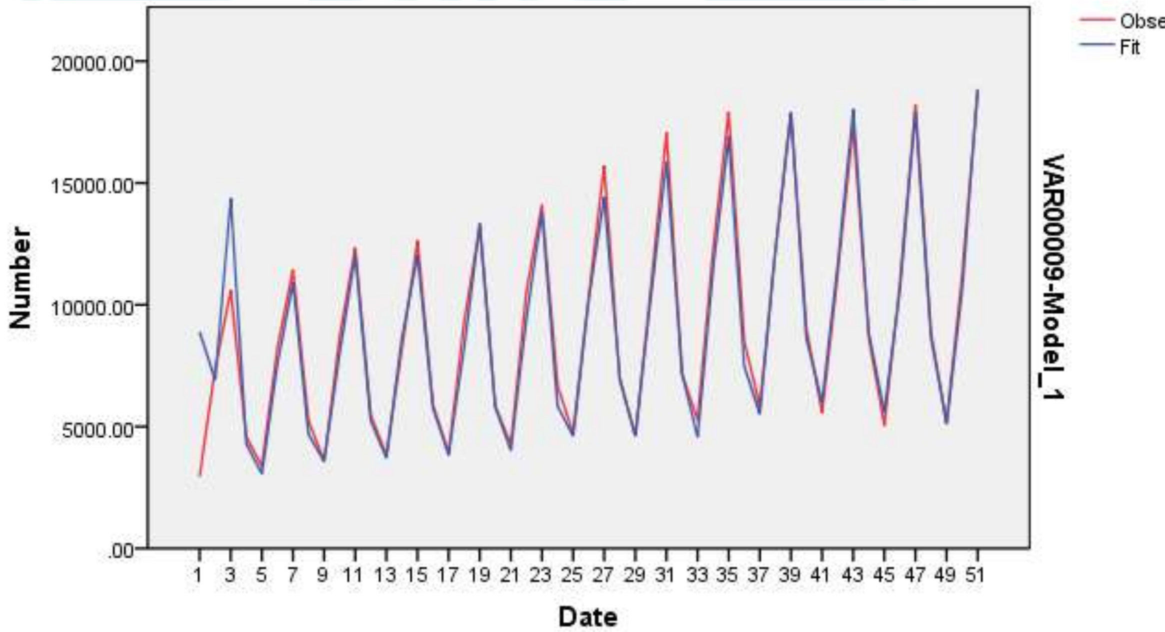
ARIMA Model Parameters

					Estimate	SE	t	Sig.
-والماء الغاز الكهرباء، Model_1	الغاز الكهرباء، والماء	Natural Logarithm	Constant		8.923	.702	12.706	.000
			AR	Lag 1	.480	.149	3.213	.002
				Lag 2	-.477-	.152	- 3.130-	.003
				Lag 3	.468	.151	3.106	.003

	Lag 4	.521	.152	3.432	.001

Each estimated model's entry is identified by its model identifier in the table of ARIMA model parameters, which displays values for all of the parameters in the model. The dependent variable and any independent variables that the Expert Modeler found to be important will be listed along with the rest of the model's variables for our purposes. We already know that there are two significant predictors from the model statistics table. The number of catalogs sent and the number of phone lines available for ordering are listed in the model parameters table.
This table gives an estimate of the model's coefficients; we note that the model's level of significance is Sig=0.00. 0.05 or less .

The ARIMA model parameters table shows values for all of the parameters, indicating that the model parameter is significant and forecasting and that the coefficients are statistically significant.
Note :
we note that the level of significance Sig= 0.00. Less than 0.05, which indicates that the coefficients are statistically significant, also effective and predictable , the model parameter is significant , able to forecasting



The predicted values show good agreement with the observed values, indicating that the model has satisfactory predictive ability. Notice how well the model predicts the seasonal peaks. And it does a good job of capturing the upward trend of the data.

Note : As in the diagram we observe the compatibility between the observed and real values.

Thus we have predicted a model that represents the data well by using all statistically significant measures , We observed contingency between observed value (red) and real value (blue) .

4. Conclusion

For the descriptive statistics of the model, R-squared represents the coefficient of good fit if the value is greater = 0.93 more than 0.05 this mean the model represent data exactly (good model) .

This table provides an estimate of the coefficients of the model, from the model we note that the level of significance Sig= 0.00. Less than 0.05, which indicates that the coefficients are statistically significant, also effective and predictable .

References

- Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 1994. Time series analysis: Forecasting and control, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.
- Brockwell, P. J., and R. A. Davis. 1991. Time Series: Theory and Methods, 2 ed. : Springer-Verlag.
- Gardner, E. S. 1985. Exponential smoothing: The state of the art. Journal of Forecasting, 4, 1-28.
- Harvey, A. C. 1989. Forecasting, structural time series models and the Kalman filter. Cambridge: Cambridge University Press.
- Makridakis, S. G., S. C. Wheelwright, and R. J. Hyndman. 1997. Forecasting: Methods and applications, 3rd ed. ed. New York: John Wiley and Sons.
- Melard, G. 1984. A fast algorithm for the exact likelihood of autoregressive-moving average models. Applied Statistics, 33:1, 104-119.
- Pena, D., G. C. Tiao, and R. S. Tsay, eds. 2001. A course in time series analysis. New York: John Wiley and Sons.
- Forecasting: Principles and Practice (2nd ed) Rob J Hyndman and George Athanasopoulos Monash University, Australia
- https://otexts.com/fpp2/what-can-be-forecast.html
- Hyndman, R.J., & Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on This online version of the book was last updated on 21

November 2022.
- The print version of the book (available from Amazon and Google) was last updated on 8 May 2018.

- Wickham, H. (2016). ggplot2: Elegant graphics for data analysis (2nd ed). Springer.
  - Introductory Time Series with R (Use R!) 2009th Edition, Kindle Edition

https://www.amazon.com/Introductory-Time-R-Use-ebook/dp/B00HWUXJTK/

- https://www.tableau.com/learn/articles/time-series-analysis-books
- The Analysis of Time-series: An Introduction with R by Chris Chatfield and Haipeng Xing Resources in R for Time Series Analysis
- On-line and off-line resources Several Springer Texts in Statistics cover Time Series Analysis using R, including:
- Time Series Analysis & Its Applications: With R Examples (3rd ed)– Robert Shumway and David Stoffer, Springer (2011)
- In addition Stoffer's own web-site includes a useful R Time Series Tutorial at http://www.stat.pitt.edu/stoffer/tsa2/R_time_series_quick_fix.htm
- Introductory Time Series with R Examples – Paul Cowpertwait and Andrew Metcalfe, Springer (2009)
- Time Series Analysis in R Examples (2nd ed) – Jonathan Cryer and Kung-Sik Chan, Springer (2009).