ARIMA: The Models of Box and Jenkins

ERIC STELLWAGEN AND LEN TASHMAN

PREVIEW Foresight *tutorials are designed to be non-technical overviews of important methodologies, enabling business forecasters to make more informed use of their forecasting software. Our Fall 2012 issue contained Eric Stellwagen's tutorial "Exponential Smoothing: The Workhorse of Business Forecasting."*

Eric and Len now team up to discuss ARIMA, the models popularized by Box and Jenkins. They examine the pros and cons of ARIMA modeling, provide a conceptual overview of how the technique works, and discuss how best to apply it to business data.

A BIT OF HISTORY

In 1970, George Box and Gwilym Jenkins popularized ARIMA (Autoregressive Integrated Moving Average) models in their seminal textbook, *Time Series Analysis: Forecasting and Control* (Box and Jenkins, 1970). While the forecasting technique they describe is commonly known as an ARIMA model, many forecasters use the phrase "Box-Jenkins model" interchangeably, as we will here.

ARIMA models initially generated a lot of excitement in the academic community, due mostly to their theoretical underpinnings. If certain assumptions are met, ARIMA models yield *optimal forecasts*, a term essentially meaning that the errors from the model contain no information that could improve the forecasts. Methodologists call such errors *white noise*. This does not imply, however, that ARIMA models are necessarily superior to alternatives, especially if the data do not conform to the necessary assumptions – and business data often do not.

ARIMA models did not initially enjoy widespread use in the business community. Mostly this was due to the difficult, time-consuming, and highly subjective procedure described by Box and Jenkins to identify the proper form of the model for a given data set. To make matters worse, empirical studies showed that despite the ARIMA model's theoretical superiority over other forecasting methods, in practice the models did not routinely outperform other time-series methods. So here is a notable example of the contrast between the theoretical property of "optimality" and actual forecasting accuracy assessed on real data.

One particularly important empirical study (Makridakis and colleagues, 1984) found that Exponential-Smoothing models – the subject of *Foresight*'s previous tutorial (Stellwagen, 2012) – outperformed Box-Jenkins 55% of the time on a sample of 1,001 data sets. This is still a good showing for Box-Jenkins (after all, it outperformed Exponential Smoothing 45% of the time), so the lesson here is this: *forecasters should switch between different methods as appropriate, rather than taking a one-size-fits-all approach*.

The challenge for a corporate forecaster is to determine which data sets are best suited to Box-Jenkins, and then to identify the proper form of the models. Most of today's software packages use algorithms to automatically select the proper form of an ARIMA model. Such automatic approaches have been shown to outperform the manual identification procedures proposed by Box and Jenkins (Ord and Lowe, 1996), and deserve much of the credit for making Box-Jenkins models accessible and useful to the business forecasting community.

CONCEPTUAL OVERVIEW

Although *multivariate* forms of ARIMA models exist, in business this method is mostly used as a *time-series* forecasting technique, which bases the forecast solely on the history of the item being forecast.

ARIMA models, like other time series models, are appropriate when you can assume a reasonable amount of continuity between the past and the future, i.e., that future patterns and trends will resemble current patterns and trends. This is a reasonable assumption in the short term, but becomes more tenuous the further out you forecast. Hence these models are best suited to shorter-term forecasting – say, 12 months or less.

Box-Jenkins models are similar to Exponential Smoothing models in that they (1) can account for trend and seasonal patterns, (2) can be automated, and (3) are *adaptive* and adjust to new information in a structured way. ARIMA models differ from Exponential Smoothing models in that they incorporate an additional type of information about the data. Called *autocorrelations*, this aspect of the data is not a trend or seasonal pattern, but a *carryover* pattern from one time period to subsequent periods. Manual (and some automatic) identification of ARIMA models is based on autocorrelation calculations.

Due to their use of the autocorrelation information, Box-Jenkins models are more mathematically complex than are Exponential Smoothing models and therefore harder to understand conceptually. In this article, we will provide an overview of how an ARIMA model works and show how its forecasts can be explained. If you're interested in learning more about Box-Jenkins models, they are covered in virtually every academic textbook on time series forecasting.

Forecasters should switch between different methods as appropriate, rather than taking a one-size-fits-all approach.

COMPONENTS OF AN ARIMA MODEL

ARIMA models attempt to identify patterns in the historical data. Their goal is to identify the *process* that is generating and influencing the historical pattern, which is called the *data generating process*.

An ARIMA model has three components, each of which helps to model a certain type of pattern. The "AR" or *autoregressive* component attempts to account for the patterns between any one time period and previous periods. The "MA" or *moving average* component (which is better understood as an error feedback term) measures the adaptation of new forecasts to prior forecast errors. The "I" or *integrated* component connotes a trend or other "integrative" process in the data. In the face of trends, the *differences* from one month to the next must be modeled rather than the monthly data themselves. Once these forecasts are made, the differences must be integrated back into monthly levels. Sometimes a second order of differencing is needed to accommodate very strong trends.

The AR and MA components have an associated *model order* indicating the duration or persistence of a pattern – how the current value of the data is affected by previous values (lags) of itself. So an autoregressive term of order 1 – an AR1 – suggests a carryover pattern from one time period only to the next, while an AR2 for monthly sales data indicates that sales in any one month is

Figure 1. A Nonseasonal ARIMA Model

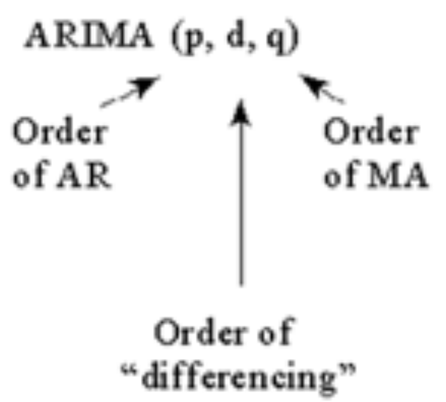


Figure 2. Illustrative Time Series of Sales and Lagged Sales

Month	Sales Y(t)	Lagged Once Y(t -1)	Lagged Twice Y(t -2)	Lagged 3 times Y(t -3)	Lagged 4 times Y(t -4)
Jan	22				
Feb	50	22			
Mar	34	50	22		
Apr	58	34	50	22	
May	59	58	34	50	22
June	53	59	58	34	50
July	90	53	59	58	34
Aug	85	90	53	59	58
Sept	73	85	90	53	59
Oct	52	73	85	90	53
Nov	114	52	73	85	90
Dec	131	114	52	73	85
Jan2	91	131	114	52	73
Feb2	127	91	131	114	52
Mar2	122	127	91	131	114
Apr2	158	122	127	91	131
May	134	158	122	127	91
June2	237	134	158	122	127
July2	270	237	134	158	122
Aug2	261	270	237	134	158

affected by sales of the previous two months. Similarly, an MA1 would relate current sales to the error in forecasting sales last month.

A *nonseasonal* Box-Jenkins model is often symbolized as ARIMA(p,d,q) where "p" indicates the order of the AR term, "q" the order of the MA term, and "d" the number of times the data must be *differenced* to de-trend or otherwise facilitate ARIMA modeling.

Seasonal Box-Jenkins models are symbolized as ARIMA(p,d,q)*(P,D,Q), where p,d,q indicates the model orders for the short-term components of the model, and P,D,Q indicates the model orders for the seasonal components of the model.

With Exponential Smoothing, the forecaster (or statistical algorithm) must determine if there are trends and seasonal patterns in the historical data and then choose appropriate forms of these attributes. With ARIMA, the analogous tasks are to find the appropriate orders of the AR, MA, and differencing components. Theoretically, the model orders could take on any integer values; in practice they are usually 0, 1, 2, or 3. This still yields hundreds of different models to consider – one of the reasons why manual identification is so involved. Manual identification of the model orders requires knowledge and experience akin to radiological examination of patient X-rays. But most ARIMA software solutions offer automatic model selection, as if the computer were reading the radiographs.

Figure 3a. Autocorrelation at Lag 1

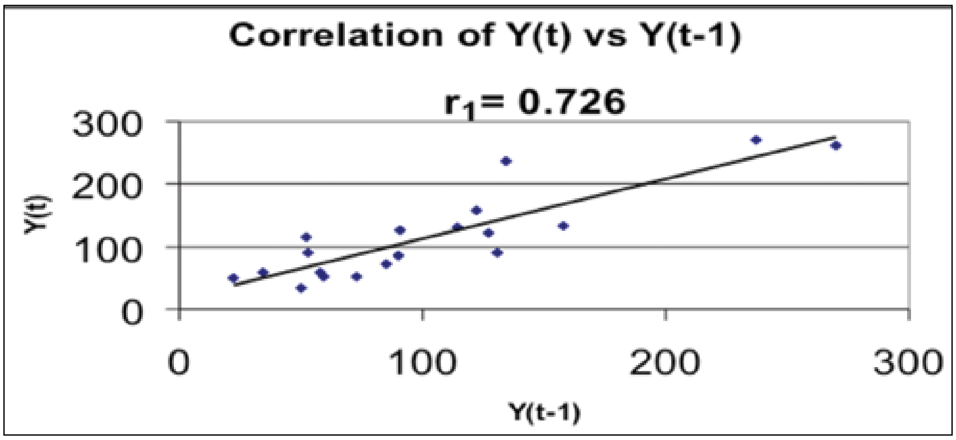
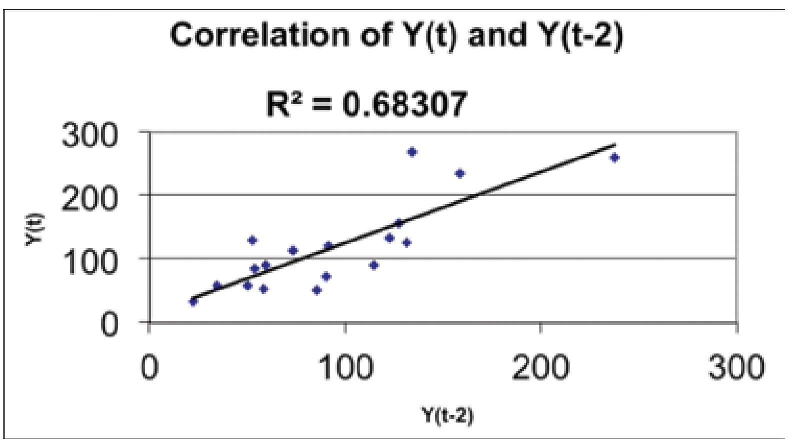


Figure 3b. Autocorrelation at Lag 2



UNDERSTANDING ARIMA MODELS: LAGGED VARIABLES AND AUTOCORRELATIONS

Figure 2 illustrates lagged variables for a hypothetical time series of 20 months of sales. The sales history is shown in the second column, while lagged sales series fill the remaining columns. For example, in the column for the series Lagged Once, each figure represents sales of the prior month. Similarly, in the far right column, each figure represents sales occurring 4 months ago.

Autocorrelation Coefficients

A correlation coefficient between the sales series and any lagged sales series is called an *autocorrelation coefficient* (we are correlating sales with a lagged value of itself). The precise formula is found in any standard text chapter on ARIMA, but it is very similar to that of a standard correlation coefficient.

Figures 3a and 3b illustrate a pair of autocorrelation coefficients for the sales data in Figure 2. Figure 3a shows that the autocorrelation coefficient for lag 1 is about 0.7, suggesting a moderate positive correlation; that is, a tendency for a sales change in any one month to be followed by a change in the same direction the following month. This "momentum" in the sales history is a key component of ARIMA modeling.

Figure 3b shows a 0.5 autocorrelation between sales two months apart, a somewhat weaker "link" than existed between sales of successive months. The pattern displayed by autocorrelation coefficients at successive lags is important for ARIMA modeling.

Autocorrelation Plots

Figure 4 displays an *autocorrelation function* (ACF) or *correlogram*, which is available in all commercial software for ARIMA that supports manual identification of the model orders. The graph displays the autocorrelation coefficients at different lags. You can recognize from Figure 3 the values of 0.7 and 0.5 for the autocorrelation coefficients at lags 1 and 2.

What do we learn from this graph? The pattern of autocorrelation coefficients is used by software programs as one of several "clues" toward selection of an appropriate ARIMA model. The business user therefore need not learn the details of reading such patterns, but should recognize simply that ARIMA models are based in part on relationships between a times series and lagged values of itself.

ARIMA MODEL BUILDING

An ARIMA Model for Nonseasonal Data

The series history is quite long (100 months) and lacks any apparent trend and seasonality; however, there is a pattern to the autocorrelation plot, shown in Figure 6, and this pattern suggests that ARIMA modeling is promising. An automatic ARIMA algorithm selected an ARIMA(2,0,0) model for this data set (an AR2).

Figure 4. An Autocorrelation Plot

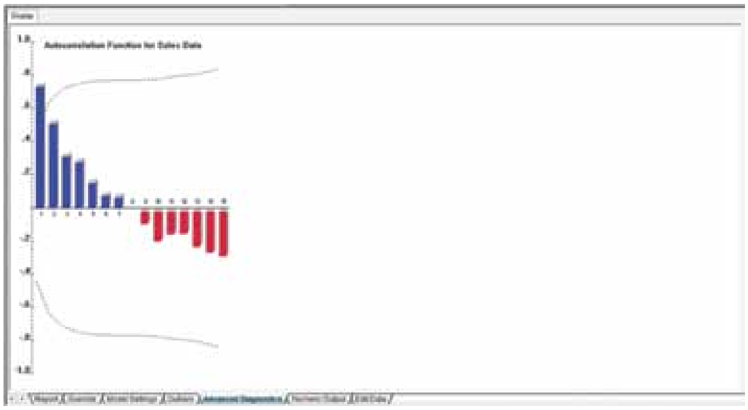
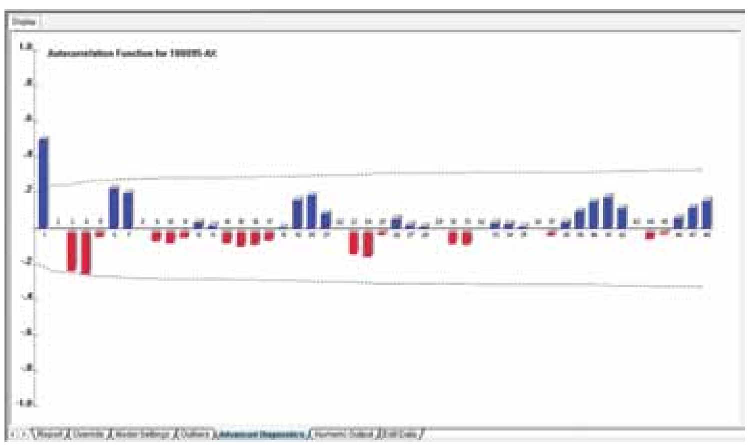


Figure 5: Plot of a Historical Time Series of Monthly Sales of an Aftermarket Part

Figure 6. The Autocorrelation Function for the Aftermarket Product Sales Data



The ARIMA(2,0,0) model form: $Y_t = \text{constant} + a_1 * Y_{t-1} + a_2 Y_{t-2}$

Using the Model to Forecast

Once the software algorithm fits the model to the historical data, the model coefficients are estimated. The resulting forecasting equation was:

$$Y_t = 20.43 + 0.72^* Y_{t-1} - 0.43^* Y_{t-2}$$

To forecast one month ahead, plug in the prior two months' values:

$$\begin{aligned} Y_{t+1} &= 20.43 + 0.72^* Y_t - 0.43^* Y_{t-1} \\ &= 20.43 + 0.72^* 32.4 - 0.43^* 26.7 \\ &= 32.1 \end{aligned}$$

To forecast two months ahead, we "bootstrap" on the one-month-ahead forecast:

$$\begin{aligned} Y_{t+2} &= 20.43 + 0.72^* Y_{t+1} - 0.43^* Y_t \\ &= 20.43 + 0.72^* 32.1 - 0.43^* 32.4 \\ &= 29.5 \end{aligned}$$

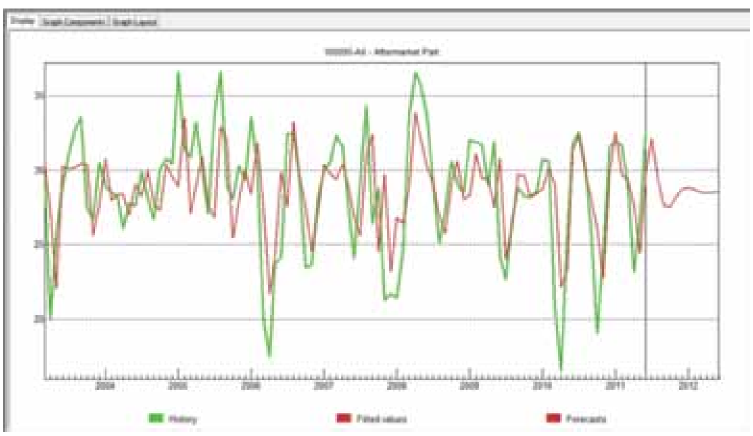
Model Validation: Diagnostic Tests

An important step in the ARIMA modeling process is *model validation*, which consists of scrutinizing a variety of diagnostics to determine whether the model selected is "healthy" – as if it were a medical patient – and hence ready to forecast, or rather that the presence of some "ailment" is detected. In the latter case, the model must be "cured" before it is used to forecast.

Most of the diagnostic tests are performed on the model errors – the difference between actual sales and the model's estimates across history. The key tests involve an *Error Autocorrelation Function* and can include a *Ljung-Box* test and a *Durbin Watson statistic*. These tests seek to determine if the errors repeat themselves in an identifiable pattern. Errors that show no discernible repeating pattern are called *white noise*, which is taken to mean that there is no evidence of model ailments. A pattern of error repetition suggests that the model selected did not fully account for the autocorrelations in the historical data and hence has the potential to be improved.

Figure 7a. Exponential Smoothing Model Fit and Forecasts

Figure 7b. ARIMA (2,0,0) Fit and Forecasts



ARIMA VS. EXPONENTIAL SMOOTHING

Our example sales data provide an illustration of ARIMA at its best, especially in relation to exponential smoothing.

The sales data do not appear to contain trend or seasonality, so the appropriate smoothing model – Simple Exponential Smoothing – yields the straight-line forecasts shown in Figure 7a. The model fit to history was poor, and a diagnostic test performed on the errors of this model showed a pattern of error repetition.

The deficiency of the smoothing model lies in its inability to make use of the autocorrelations in the data. The ARIMA model, on the other hand, captured these autocorrelations, resulting in an improved fit to the data and white-noise errors. Observe in Figure 7b that the forecasts are not simply a horizontal line but reveal a pattern similar to the autocorrelations apparent in Figure 6.

ARIMA models have the potential to outperform Exponential Smoothing models (as shown in the Aftermarket Part example), but have more stringent data requirements.

The principal advantage for ARIMA lies in its incorporation of information on autocorrelations in the historical data, information beyond the scope of most smoothing procedures. The sales data used in our example are (1) lengthy, (2) lack strong trends and seasonal variation, and (3) exhibit autocorrelation patterns. For series with these characteristics, we have illustrated how ARIMA can outperform ES.

The ARIMA models we've discussed here are only a portion of the available family of ARIMA procedures. While Exponential Smoothing models are typically limited to procedures that extrapolate the historical data, ARIMA can be extended to incorporate causal variables, somewhat similar to the manner in which regression models are built. A good example is in the Vosen and Schmidt article on page 38 of this issue, where an ARIMA model is augmented by the inclusion of several causal variables.

Multivariate ARIMA, sometimes abbreviated as ARIMAX, requires specialized software and is of more benefit to users with advanced statistical and modeling knowledge.

Both ARIMAX and Exponential Smoothing contain procedures that can model *disruptions* to the historical data – outliers, special events, structural shifts – but the ARIMAX procedures are richer and more nuanced.

Many businesses have new or semi-new products that have short histories (less than 24 months, for example). For short series, Exponential Smoothing is often (but not always) still feasible for identifying possible trend and seasonal behavior, while ARIMA is handicapped in identifying the autocorrelation patterns. The result is that when applied to short series, ARIMA often defaults to a simple extrapolative procedure such as the *naïve forecast* of no change from the most recent time period.

Until software mastered the ability to automatically find appropriate ARIMA models, ARIMA was a tough sell in the business world. And it remains a far greater challenge to explain an ARIMA forecast to a nontechnical audience than it is to explain an Exponential Smoothing forecast. Exponential Smoothing is based on tangible components of the data – level, trend, seasonality – each of which lends itself to a straightforward interpretation. ARIMA's autocorrelation-based procedures are anything but straightforward to decipher. Nonetheless, because of their strengths with selected data sets, ARIMA models are an important resource that should be in the toolbox of most business forecasters.

REFERENCES

Box, G.E.P. & Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*, Revised Edition, San Francisco: Holden Day.

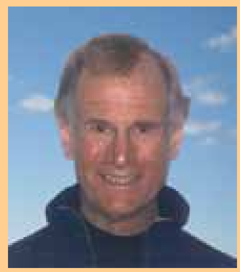
Makridakis, S. et al. (1984). *The Forecasting Accuracy of Major Time Series Methods*, Chichester: Wiley.

Ord, K. & Lowe, S. (1996). Automatic Forecasting, *The American Statistician*, Volume 50, Number 1, 88-94.

Stellwagen, E. (2012). Exponential Smoothing: The Workhorse of Business Forecasting, *Foresight*, Issue 27 (Fall 2012), 23-28.

Eric Stellwagen
is cofounder of Business Forecast Systems, maker of Forecast Pro software solutions.

estellwagen@forecastpro.com



Len Tashman
is *Foresight*'s Editor.
lentashman@forecasters.org