Time-Series Forecasting Model for Automobile Sales in Thailand

Taweesin Apiwattanachai¹ and Juta Pichitlamken²

Abstract

Inventory management at car dealers is generally not efficient because dealers place orders based on their prior sales experiences to ensure that cars will be readily available to customers. If stocks are held at dealers for a long time, especially until the end of model life, it will be difficult to clear out those stocks. Customers are thus offered financial incentives (such as free insurance) which are then subsidized by car manufacturers. We propose to set up Information Database Center that collects sales data in real time. A modified Holt-Winter’s forecasting model is built to estimate customer demand, instead of using historical sales data. We evaluate our forecasting model by comparing forecasts (from the Holt-Winter’s and our modified model) with actual data.

Keywords: inventory management, time-series forecasting model, Holt-Winter’s method

¹ International Graduate Program in Industrial Engineering, Faculty of Engineering, Kasetsart University, Bangkok, Thailand
E-mail: taweesin_api@hotmail.com

² Industrial Engineering, Faculty of Engineering, Kasetsart University, Bangkok, Thailand; E-mail: juta.p@ku.ac.th
1. Introduction

In Thailand, automobile industry is considered one of the major industries which significantly contribute to the nation’s economy. As of 2007, automobiles and parts rank second in export values (9.5% of total) and accounts for 14% of the Gross Domestic Products (GDP) [1]. Figure 1 shows that the auto industry employs 10% of the Thai workforce. As a result, the Royal Thai Government has been promoting Thailand as the Detroit of Asia by providing funds and tax incentives, such as reducing import duty for vehicles under Free Trade Area (FTA) Agreements.

Because of many supports from the government to promote automobile industries in Thailand, competition among car manufactures is fierce. They realize that if they are unable to know customer’s buying behaviors well, their product would not be strong enough to survive in a highly competitive market. According to J.D. Power and Associates Reports on the top 10 reasons for avoiding a vehicle [2], top 3 main factors affecting decision-making of customers when buying a car are generally based on the following items:

1. Styling
2. Reliability
3. Price

Considering each factor, styling is personal preference. Nowadays, car reliabilities are not significantly different among brands due to compliance to international standards such as Economic Commission for Europe (ECE). Moreover, vehicles are periodically benchmarked to assess their strengths and weaknesses in comparison with competitors. As a result, automobile manufacturers aim to compete on prices. Manufacturers focus on reducing production cost to achieve competitive prices as well as to gain more profit.

Widely used cost reduction schemes are waste reduction, value engineering, and inventory reduction. Waste reduction considers both materials and production time. Value engineering approaches aim to measure value of a product in terms of quality, performance, and reliability (at an acceptable price) and to remove non value-added aspects where value is defined as worth/cost [3, 4]. In general, manufacturers mostly concentrate on reducing production inventory. However, we think that inventory of finished vehicles should be considered as well.

Dealers place orders based on their prior sales data. They often overestimate the order quantities to ensure that cars will be readily available for customers. Dealers are responsible for holding those stocks. However, if stocks are held at dealers for a long time, especially towards the end of model life, it would be difficult to clear out those stocks. Thus, customers are offered financial incentives (such as free auto insurance) which are then subsidized by car manufacturers. Some companies spend up to ten million Baht per year to clear out this dead stock.

If we have a better forecasting model for demands, there should be fewer dead stocks. We propose an Information Database Center that collects sales data in real time. From this data, a forecasting model is generated to estimate customers demand. Our model takes into account special characteristics of the Thai automobile market; for example, external factors such as petrol price, interest rate for loan, and average household income, together with seasonal characteristics due to annual sale events (see Figure 1 for example).

There are many forecasting techniques, such as, curve fitting (or regression) methods, smoothing methods, and seasonal smoothing methods. We choose the “Triple Exponential Smoothing” known as the Holt-Winter’s (HW) method because it is often applied to time series that exhibit trend and seasonality such as our case.

HW is essentially a quantitative method that uses mathematical recursive functions to predict trend and seasonality behaviors. It assumes that the future will follow the same pattern as the past. In our case, we have seasonal patterns corresponding to weekly, quarterly or annual periodicity; therefore, we should include factors in our model that utilize the historical information [5]. After the HW model is formulated, we evaluate it by comparing the forecasts with the actual data to see if it could help dealers place orders more effectively, e.g., to
minimize numbers of cars in their stock yards.

This paper is organized as follows: We present background on time-series forecasting models that are related to our work in Section 2. We outline our modeling procedure in Section 3. We present our results in Section 4. We conclude in Section 5.

2. Time-Series Forecasting Model

We discuss exponential smoothing models in Section 3.1 and the special case of an exponential smoothing model: Holt-Winter in Section 3.2.

2.1 Exponential smoothing

One of the forecasting techniques that can address a fairly predictable environment is time series [6-8]. Time series models include regression, decomposition and various adaptive methods. With such techniques, one essentially seeks to identify patterns in the data over time and moves to project the established patterns into the future. However, we have to keep in mind that these models assume that "what has happened in the past will continue to happen in the future," but, by definition, the future is unpredictable. If this basic assumption is violated, whether as a result of external or internal changes (e.g., the firm intends to launch a massive advertising campaign), the accuracy of the forecasts becomes questionable [9].

Time series analysis is extensively utilized in many areas, such as economic forecasting, budgetary analysis, and inventory studies. Users select a model-fitting method based on an application on hand as well as preference. These methods include Moving Average, Box-Jenkins, Autoregressive Integrated Moving Average (ARIMA), Box-Jenkins Multivariate model, and exponential smoothing. In this paper, we adopt the exponential smoothing because some studies, such as [10-11], show that exponential smoothing outperforms the more sophisticated Box-Jenkins models.

Exponential methods can be classified into three types: single exponential smoothing, double exponential smoothing and triple exponential smoothing or Holt-Winter’s (HW) method. In this paper, we concentrate on HW since automobile demand exhibit both trend and seasonality behaviors.

2.2 Holt-Winter’s method

Depending on the type of seasonality, HW models can be either a multiplicative seasonal model (Section 2.2.1) or an additive seasonal model (Section 2.2.2).

2.2.1 Multiplicative Seasonal Model

The multiplicative seasonal model is appropriate for time series in which amplitude of the seasonal pattern is proportional to the average level of the data [12]. It assumes that the time series is represented by Equation (1):

\[ F_t = (b_1 + b_2 t) S_t + ε_t \]  

where \( F_t \) is the forecast at time \( t \), \( b_1 \) is the base signal or the permanent component, \( b_2 \) is a linear trend component, \( S_t \) is a multiplicative seasonal factor, and \( ε_t \) is the random error component which is a standard normal random variable \( N(0,1) \).

From Model (1), we obtain the following recursive formula:

\[ F_t = (R_{t-1} + G_{t-1}) S_{t-L} \]

where \( R_t \) stands for the estimator of the permanent component \( b_1 \) at time \( t \), \( G_t \) stands for the estimator of the trend component \( b_2 \), \( L \) stands for the number of periods of historical data that is used to obtain forecasts.

Let \( y_t \) stand for the actual data at time \( t \). Parameters \( R_t \), \( G_t \), and \( S_t \) are updated as follows:

\[ R_t = \alpha \frac{y_t}{S_{t-L}} + (1 - \alpha)(R_{t-1} + G_{t-1}) \quad 0 \leq \alpha \leq 1 \]  
\[ G_t = \beta (S_t - S_{t-1}) + (1 - \beta)G_{t-1} \quad 0 \leq \beta \leq 1 \]  
\[ S_t = \gamma \left( \frac{y_t}{R_t} \right) + (1 - \gamma)S_{t-L} \quad 0 \leq \gamma \leq 1 \]

where \( \alpha \), \( \beta \) and \( \gamma \) are smoothing constants.

The value of forecast \( T \) periods after time \( t \) is given by:

\[ F_{t+T} = (R_{t-1} + TG_{t-1}) S_{t+T-L} \]  

To initialize seasonal factors, minimum historical data of one full seasons (or \( m \) periods) is needed:

\[ G_0 = \frac{y_m - y_1}{m-1} \]  
\[ R_0 = \frac{1}{n} \sum_{i=1}^{m} y_i \]  
\[ S_0 = y_1 - R_0 \]

2.2.2 Additive Seasonal Model

The additive seasonal model is suitable for data whose amplitude of seasonality is independent of the average level of the series. The additive seasonal model has the following form:

\[ F_t = b_1 + b_2 t + S_t + ε_t \]
where all variables have previously been defined in Equation (1).

Parameters $R_t$, $G_t$, and $S_t$ are updated as follows:

$$R_t = \alpha (y_t - S_{t-L}) + (1 - \alpha) (R_{t-1} + G_{t-1}) \quad 0 \leq \alpha \leq 1 \quad (10)$$

$$G_t = \beta (S_t - S_{t-1}) + (1 - \beta) G_{t-1} \quad 0 \leq \beta \leq 1 \quad (11)$$

$$S_t = \gamma (y_t - R_t) + (1 - \gamma) S_{t-L} \quad 0 \leq \gamma \leq 1 \quad (12)$$

The forecast for the next period is given by:

$$F_{t+1} = R_{t+1} + G_{t+1} + S_{t+1}. \quad (13)$$

3. Model Development

We consider one particular car model which has a lot of sales data, where we collect and categorize it. For this study, we divide the data into 5 groups; top grade automatic transmission (Top A/T), medium grade automatic transmission (Med A/T) and manual transmission (Med M/T), low grade automatic transmission (Low A/T) and manual transmission (Low M/T).

We choose either a multiplicative seasonal model (Equation (1)) or additive seasonal model HW model (Equation (9)) by considering Mean Absolute Percentage Error (MAPE) as shown in Equation (14). For this data set, time $t$ is in month and $M$ is 12 because automobiles are classified by year. The data used in this study are from January 2002 to April 2007.

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \frac{|y_t - F_t|}{y_t} \quad (14)$$

where $n$ stand for total number of data.

Both the multiplicative and the additive model require all 3 smoothing constants, $\alpha$, $\beta$ and $\gamma$, since automobile sale has both trend and seasonality behaviors. As a rule of thumb, values between 0.01 and 0.3 are used when the forecasts depend on a large number of past values, while larger values of smoothing constants are used when forecasts depend more heavily on a few recent values [13]. We use Excel’s solver to determine smoothing constants (Table 1).

Table 1: Smoothing constants of each configuration.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Multi</td>
<td>Add</td>
<td>Multi</td>
</tr>
<tr>
<td>Top A/T</td>
<td>0.30</td>
<td>0.59</td>
<td>0.30</td>
</tr>
<tr>
<td>Med A/T</td>
<td>0.30</td>
<td>0.33</td>
<td>0.30</td>
</tr>
<tr>
<td>Med M/T</td>
<td>0.60</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>Low A/T</td>
<td>0.81</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>Low M/T</td>
<td>0.66</td>
<td>0.54</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 2: MAPE of the multiplicative model and additive model.

<table>
<thead>
<tr>
<th></th>
<th>Multiplicative</th>
<th>Additive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top A/T</td>
<td>1.14</td>
<td>0.33</td>
</tr>
<tr>
<td>Med A/T</td>
<td>0.64</td>
<td>0.25</td>
</tr>
<tr>
<td>Med M/T</td>
<td>0.55</td>
<td>0.32</td>
</tr>
<tr>
<td>Low A/T</td>
<td>N/A</td>
<td>0.23</td>
</tr>
<tr>
<td>Low M/T</td>
<td>0.45</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 2 shows the resulting MAPE. Because the additive model has a lower MAPE, we only consider this model from now on. The forecasts can be graphically compared with the actual data as shown in Figure 2. Since the overall appearance of this automobile model was completely overhauled in December 2002, it might cause a drastic increase in customer demand. To better respond to this kind of fluctuation, we need to modify how our seasonal component is updated. In addition, we consider “smoothing” our data before we fit a forecasting model.

We smooth our data by using moving averages. For the data during the first season (when forecasting cannot be done), i.e., $t \leq m$, if $|y_t - y_{t-L}| / y_t \geq c$ for some threshold $c$, we replace $y_t$ with $(y_{t-L} + y_{t-L+1}) / 2$. After the first season (when forecasts exist), i.e., $t > m$, if $|y_t - F_t| / y_t \geq c$ for some threshold $c$, we replace $y_t$ with $(y_{t-L} + y_{t-L+1}) / 2$.

We consider the threshold $c$ of 10%, 20%, ..., 100% and see which value gives the lowest MAPE. Our results show that setting $c$ to 10% is best. Even after smoothing the data, our forecasting model still cannot capture the fluctuation pattern in
the data. Thus, we modify the seasonal component \( S_j \). The idea is that if \( S_j \) is a “real” seasonal component, its value should remain relatively constant from one season to the next. Thus, if error, defined as \( \left| \frac{y_i - F_j}{y_i} \right| \geq 10\% \), and

\[
\left| (y_i - F_j) - I_i \right| < \left| (y_i - F_j) - S_j \right|
\]

(15)

where \( I_i \) is the interpolated value of the seasonal component of the most recent past data and the future data which have error between forecasted value and actual value lower than 10%. Figure 3 shows the comparison between the original HW and our modified HW with the actual data. The original HW model produces highly inaccurate forecasts for December sales of every year due to the unusual observation in the first season. Our modified HW model can reduce the effect of this outlier and produce forecasts that better track the general trend in the data.

4. Model Evaluation

We evaluate our models with the data that are not used for model fitting: May 2007 to March 2008. Table 3 shows the MAPE of the forecasts that this automobile company came up with, the forecasts from the original HW and the forecasts from our modified model. We see that both HW models, altogether, perform better than the company’s forecasts for 4 out of 5 data sets. The modified HW, by itself, wins in 3 out of five cases, but the original HW wins for the Low M/T data set, and the company’s forecasts fares the best for the Low A/T data set.

![Figure 3: Comparison of two types of forecasts from the additive model with actual data for Top A/T configuration.](image)

Table 3: MAPE of the company’s forecasts, the original HW forecasts and the modified HW forecasts.

<table>
<thead>
<tr>
<th></th>
<th>Company’s forecasts</th>
<th>Forecasting by original HW</th>
<th>Forecasting by modified HW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top A/T</td>
<td>0.574</td>
<td>0.445</td>
<td>0.317</td>
</tr>
<tr>
<td>Med A/T</td>
<td>0.394</td>
<td>0.594</td>
<td>0.173</td>
</tr>
<tr>
<td>Med M/T</td>
<td>0.543</td>
<td>0.220</td>
<td>0.179</td>
</tr>
<tr>
<td>Low A/T</td>
<td>0.167</td>
<td>0.236</td>
<td>0.259</td>
</tr>
<tr>
<td>Low M/T</td>
<td>0.313</td>
<td>0.196</td>
<td>0.224</td>
</tr>
<tr>
<td>Average</td>
<td>0.398</td>
<td>0.338</td>
<td>0.230</td>
</tr>
</tbody>
</table>

To apply the modified HW into the inventory management for car dealers, historical sale record should be periodically uploaded onto Information Database Center. This data is then fed to some computer program to calculate sale forecasts which can be used for production planning.

5. Conclusion

Despite a better performance than the conventional forecasting method that the company currently uses, our proposed models can be improved further if we do a better job of determining smoothing constants (\( \alpha \), \( \beta \) and \( \gamma \) in Equations (2)—(4) and (10)—(12)). Although interpolated values can improve seasonal components, sometimes it might cause worse forecasts if interpolated value cannot be representative of seasonal component suitably due to a coincidence and high fluctuation of actual data during value interpolation, as we can see in Low A/T and Low M/T configurations in Table 3.

6. Acknowledgement

The first author would like to thank the International Graduate Program in Industrial Engineering for financial supports on tuition fee during the first year of study.

References

12 Time Series Forecasting using Holt-Winters Exponential Smoothing, Available at http://www-it.iitb.ac.in/~praj/acads/seminar/04329008_Exponential Smoothing.pdf
13 Exponential smoothing, Available at http://www.ist.massey.ac.nz/clai/161342/LectureNotes-07/Lectures%20(Ch%208).doc