Appendix I

Sales Analysis with Application to Sparkling Beverage Products Sales in Southern Thailand

Sales Analysis with Application to Sparkling Beverage Products Sales in Southern Thailand

Wassana Suwanvijit (Corresponding author)
Faculty of Economics and Business Administration, Thaksin University
140 Karnchanavanich Road, Muang, Songkla, 90000, Thailand
E-mail: swassana@gmail.com.

Chamnein Choonpradub
Faculty of Science and Technology, Prince of Songkla University
181 Charernpradit Road, Rusamilar, Pattani, 94000, Thailand
E-mail: cchamnein@bunga.pn.psu.ac.th

Nittaya McNeil
Faculty of Science and Technology, Prince of Songkla University
181 Charernpradit Road, Rusamilar, Pattani, 94000, Thailand
E-mail: nittaya@bunga.pn.psu.ac.th

Abstract

Managers need to know about their business performance and how to gain rapid insight into fast-growth products and profit opportunities. This case study used monthly sales revenue from sparkling beverages of data collected routinely from 14 provinces of Southern Thailand during years 2000-2006. Linear regression models of log-transformed sales revenue per 1000 population were used to analyse per capita revenue and study product preference with respect to flavour and branch location. Population data was obtained from the 2000 Population and Housing Census of Thailand. The model contains quarterly effects and interactions associated with flavour-branch and branch-year. The average annual per capita consumption value rate was 297 baht for the population of 8.087 million residents in Southern Thailand. Branches in Samui and Phuket had higher rates than other branches. In addition, areas with different proportions of Muslims in their population had different beverage preferences.

Keywords: Sales revenue, Linear regression model, Sparkling beverages, Consumption rate, Product preference, Southern Thailand

1. Introduction

Generally, business owners and managers need to know about their sales trends and profitability, not only overall but by departments or products within each department. They also have to consider the potential of their business improvement from different areas, by gathering, classifying, comparing, and studying company sales data. Most managers don't have the time or expertise to analyse their sales data themselves. Nevertheless, there are many studies about sales analysis. In the case study of carbonated soft drink consumption and bone mineral density in adolescence by McGartland et al. (2003), adjusted regression modeling was used to investigate the influence of carbonated soft drinks on bone mineral density. Residual demand analysis (Higgins et al. 2005) could be applied to test whether carbonated soft drinks were a relevant product market under the Merger Guidelines. Goktolga et al. (2006) analyzed socio-demographic factors affecting a decrease in consumption of chicken meat because of Avian Influenza (Bird Flu) in Turkey using a multiple bounded probit model. Descriptive and multiple regression analyses (Probart et al. 2006) had some good features to identify the factors associated with the offering and sale of competitive foods and school lunch participation. Helasoja et al. (2007) used logistic regression to fit socio-demographic patterns of drinking and binge drinking in Estonia, Latvia, Lithuania and Finland, 1994-2002. Ratio analysis, historical trend analysis and linear regression analysis used by Bureau (2007) were found to be useful for analyzing the factors and examining the opportunities critical to the success of the food and beverage industry in India.

This study illustrates methods for graphing and statistical modeling available to businesses and demonstrates how each can be implemented using freely available software. Linear regression models are used to see how well these methods can be applied to a case study in southern Thailand. From an academic perspective the study demonstrates how easily business analysis can be taught to both students and managers. From a managerial perspective, the study demonstrates that developing accurate analysis capability need not be expensive or overly time-consuming. The purpose of this study is to understand the main factors that influence sparkling beverage sales in Southern Thailand and to find the annual per capita consumption value rate. The study investigates the associations between product preference with respect to flavour, branch location and socio-demographic factors that are useful for planning and decision making. Our study aims are to find a suitable statistical model to describe the variation in quarterly sales data of a major sparkling beverage company in southern Thailand reported from 2000 to 2006.

Thailand is divided into four geographical regions. The southern region occupies about 14% of the total land area. There are 14 provinces in southern Thailand with a total area 71,798 square kilometers. Sparkling beverages are traditionally popular products in the south, although some consumers prefer more healthy beverages. Sparkling beverages are also primarily used as mixers for consumption with alcoholic drinks.

2. Methods

The outcome in this study is sales revenue. Determinants are flavour, branch location and quarterly period. Flavours were identified for five types of products, namely “Cola flavour”, “Orange flavour”, “Red flavour”, “Green flavour” and “Lime flavour”. Two types of returnable packages (10 ounce and 1 liter) and three types of non-returnable packages (1.25 liter, 2 liter and 325 ml) were identified. In our study area 20 branches were sampled from 14 provinces. Provinces with more than one branch were Surat Thani, Nakorn Sri Thammarat, Phang-nga, Chumphon and Narathiwat. Branch locations were grouped into 3 areas (Muslim, tourist and normal areas) based on product preferences and consumer behavior in each branch location.

Data for each month were available in computer files with records for sales revenue separated by flavour and branch location. Records from years 2000 to 2006 were stored in a MySQL database. SQL was used to create sales revenue data in baht by month, flavour and branch location. All graphical and statistical analyses were performed using R software (R Foundation for Statistical Computing, Vienna, Austria).

This quantitative research focused on using statistical graphics and statistical models, including linear regression, for analyzing sales revenue and its dependence on flavour, branch location and quarter. Sales revenue data generally have positively skewed distributions so it is conventional to transform them by taking logarithms. Cells with zero counts were adjusted to avoid the problems of taking the logarithm of zero.

Linear regression models of log-transformed sales revenue per 1000 population were used to analyze the sparkling beverage sales per capita revenue and to study the main factors including product preference with respect to flavour and packaging.

The annual per capita consumption rate was computed using sales revenue divided by the number of years and population obtained from the 2000 Population and Housing Census of Thailand.

The Southern Thailand market can be grouped by branches location and consumer preferences into tourist area, Muslim area and normal areas as shown in Figure 3.

3. Statistical model

The simplest linear model takes the additive form

$$Y_{ijtk} = m + b_i + f_j + a_t + q_k, \tag{1}$$

where Y_{ijtk} is the natural logarithm of the quarterly revenue in 1000s of baht per 1000 population for branch i (b_i), flavour j (f_j), and quarter k (q_k) of year t (a_t). In this formula m is the overall mean of the log-transformed quarterly revenue.

Since this additive model does not allow for different flavour preferences in different regions, we also consider a more general model of the form

$$Y_{ijtk} = m + c_{ij} + a_t + q_k \qquad (2)$$

In this model c_{ij} is an interaction between branch and flavour (c_{ij}). Generalizing further, we also consider the model

$$Y_{ijtk} = m + c_{ij} + d_{it} + q_k \qquad (3)$$

Model (3) thus allows for interactions between branch-flavour (c_{ij}) and branch-year (d_{it}).

These statistical methods are described in further detail by Venables and Ripley (2002, chapter 6). After fitting the models, we plotted confidence intervals for parameters after back-transforming so that the parameter estimates were expressed in terms of the original data, that is, in 1000 baht per 1000 population. To do this, it was necessary to incorporate an additional scale parameter for each factor to ensure that the mean revenue associated with each factor based on the fitted model matched the overall observed mean revenue.

4. Results

Figure 4 shows the overall distribution before and after transforming the data by taking natural logarithms of sales revenue per 1000 population. The sales revenue data were symmetry with homogeneously of variance. This log-normal distribution can be used to provide an estimate of consumption rate in each quarter.

From Table 1, the results indicate that Model (3) is the best regression model to use in this study because it has the highest r-squared value, which was very close to 1.

From Figure 5, it is clear that the residual plot of Interaction Model (3) was the most linear.

Figure 6 shows plots of sales revenue in baht (left panel) and per 1000 population (right panel) versus fitted value. The model predicts the proportions in the 2800 cells very well.

Table 2 shows an annual per capita consumption rate in each branch location compared with population in the corresponding catchment area. The average annual per capita consumption rate was 297 baht for the total population of 8.087 million residents.

Figure 7 shows sales revenue per 1000 population by branch location and flavour. Samui (Sm) and Phuket (Pk) branches had higher rates than the other branches, while Nakhon Sri Thammarat (Nk) branch had the lowest consumption rate.

Figure 8 shows sales revenue of each flavour grouped by branch location and consumer preferences for Muslim, tourist, and normal areas. Areas with different proportions of their population being Muslim, had different beverage preferences due to consumers preferring colour and lime products.

5. Discussion and conclusion

The log-transformed quarterly sales revenue trends can be modeled using linear regression. From this model, we found that quarterly effects and interactions associated with flavour-branch and branch-year are main factors that influence the sparkling beverage sales in southern Thailand.

We also found that the average annual per capita consumption value rate was 297 baht for the population of 8.087 million residents in southern Thailand. The annual per capita consumption rate can be used to find market opportunities which lead to greater share and growth in each area. Samui and Phuket branches, both tourist areas, had higher consumption rates than other branches. This may be because most tourists prefer these drinks over other types, are familiar with sparkling beverages. Nakhon Sri Thammarat (Nk) branch had the lowest consumption rate, thus this branch has an opportunity to increase their sales growth. This province also has the highest population in southern Thailand. A company would need to do the market development including more activities or promotions to drive volume and sales growth rate.

We studied the associations between product preference with respect to flavour, branch location and socio-demographic factors by grouping branch location into 3 types, namely tourist, Muslim and normal areas. Areas with Muslim population had different preferences compared with other areas because the Muslims prefer cola products and colour flavours (orange, red, green) of sparkling beverage as well. This is very useful fact findings for managers to know where is the right place to push each product in and to understand how to create more sales into each area.

Other studies have also used linear regression analysis (McGartland et al, 2003; Probart et al, 2006., and Bureau, 2007) but their studies did not take the location into account and did not include any analysis on per capita consumption rate. In our study, linear regression models were applied to both sales revenue and consumption rate analysis with application to a case study of sparkling soft drink products in southern Thailand. The statistical model used in this study was very suitable to answer our research questions and the results are very useful for managers to understand their performance, consumers needs in each branch location and opportunities to boot up the sales growth in each areas. Having the best model, managers can provide a useful basis for sales analysis, incorporating the results into their company plans and strategies. However, further study in this data will be expanded to more complex business forecasting.

Acknowledgements

We are grateful for Prof. Don McNeil and Greig Rundle for their helpful advice and suggestions. We also would like to thank Khun Dumrongrugs Apibalsawasdi for his helpful guidance.

References

Bureau. (2007). Indian Food and Beverages Forecast (2007-2011). [Online] Available: http://www.bharatbook.com/detail.asp?id=50059 (January 12, 2008)

Cox, D.R. & Wermuth, N. (1996). Multivariate Dependencies-Models, Analysis and Interpretation. London: Chapman & Hall.

Goktolga, Ziya Gokalp., Gunduz, Orhan. (2006). The Analysis of Socio-Demographic Factors Affecting Decrease in Consumption of Chicken Meat Because of Avian Influenza (Bird Flu) in Turkey: Case of Multiple Bounded Probit Model. *Journal of Applied Sciences Research*, 2(12), 1350-1354.

Helasoja, Ville et al. (2007). The sociodemographic patterning of drinking and binge drinking in Estonia, Latvia, Lithuania and Finland, 1994-2002. BMC Public Health, 7, 241.

Higgins., Richard, S., Kaplan., David, R.and McDonald., Michel, J. (2005). Residual Demand Analysis of the Carbonated Soft Drink Industry. [Online] Available: http://www.springerlink.com/content/h7n6w33821up1w24 (July 13, 2007)

Kleinbaum, D.G., Kupper, L.L., Muller, K.E., & Nizam, A. (1998). Applied regression analysis and multivariable methods, 3rd ed, Duxbury Press, Pacific Grive, C.A.

McGartland, C., Robson, P.J., Murray, L., Cran, G., Savage, M.J., Watkins, D., Rooney, M., & Roreham, C. (2003). Carbonated Soft Drink Consumption and Bone Mineral Density in Adolescence: The Northern Ireland Young Hearts Project.. *Journal of bone and mineral research*, 18, 1563-1569.

National Statistical Office. (2002). The 2000 population and housing census, southern region. Bangkok: Statistical Data Bank and Information Dissemination Division, National Statistical Office.

Nichols, Donald. (2005). The Simple Regression Model. [Online] Available: http://dnichols.wustl.edu/413sp2005/simple%20regression%20sp2005.pdf (June 5, 2007)

Probart C., McDonnell E., Hartman T., Weirich J., & Bailey-Davis L. (2006). Factors associated with the Offering and sale of competitive foods and school lunch participation. *Journal of the American Dietetic Association*, 106, 242-247.

R Development Core Team. (2007). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statisical Computing, Vienna.

Syspro. (2007). Sales Analysis. [Online] Available: http://www.syspro.com/PDF/sales_analysis.pdf (March 24, 2008)

Venables, W.N., & Ripley, B.D. (2002). Modern Applied Statistics with S. Springer. Queensland.

Venables, W.N., & Smith, D.M. (2004). An Introduction to R. [Online] Available: http://cran.r-project.org/doc/manuals/R-intro.pdf (March 30, 2007)

Table 1. Result of fitting linear regression model

Model	R^2	df	RSS
1	0.913	2767	211.9
2	0.959	2691	100.7
3	0.973	2577	65.1

Table 2. Population (2000 census) of sparkling beverage branch locations and annual consumption in Southern Thailand

Branch	Symbol	Population (x 1000)	Annual consumption (Baht/population)
Nakhon Sri Thammarat	Nk	1,375	92
(Exclude: Thung Song district)			
Hat Yai	Hy	1,258	267
Surat Thani	Sr	746	183
(Exclude: Phunphin and Samui district)			
Narathiwat	Nv	600	69
(Exclude: Kolok district)			
Pattani	Pn	599	92
Trang	Tr	597	142
Phatthalung	Pl	501	84
Yala	Yl	418	172
Chumphorn	Cp	380	76
(Exclude: Lang Saun district)			
Krabi	Kb	338	169
Phuket	Pk	250	697
Satun	St	219	144
Phang-nga	Pg	192	105
(Exclude: Takuapa district)			
Ranong	Rn	162	208
Thung Song	Ts	148	325
Phunphin	Pp	91	428
Lang Saun	Ls	69	291
Kolok	Kl	65	386
Takuapa	Tk	44	583
Samui	Sm	35	1,424
Total		8,087	
Average annual consumption (Baht per population)			297

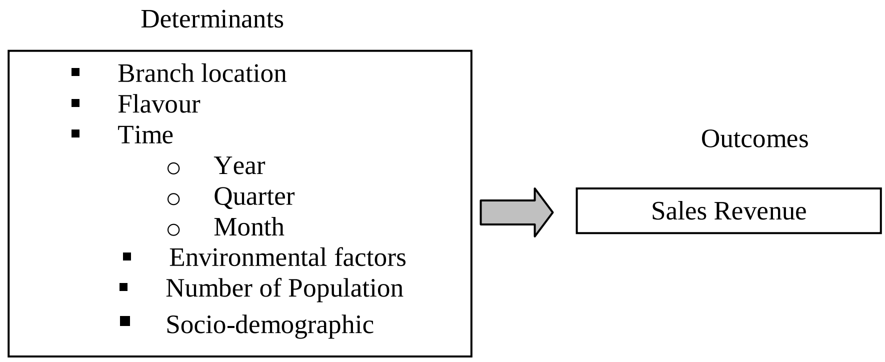

Figure 1. Data path diagram

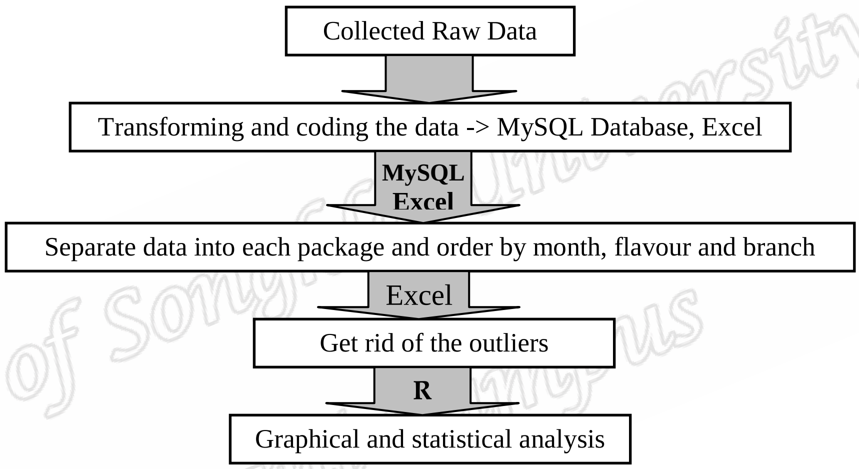

Figure 2. Data collection and management

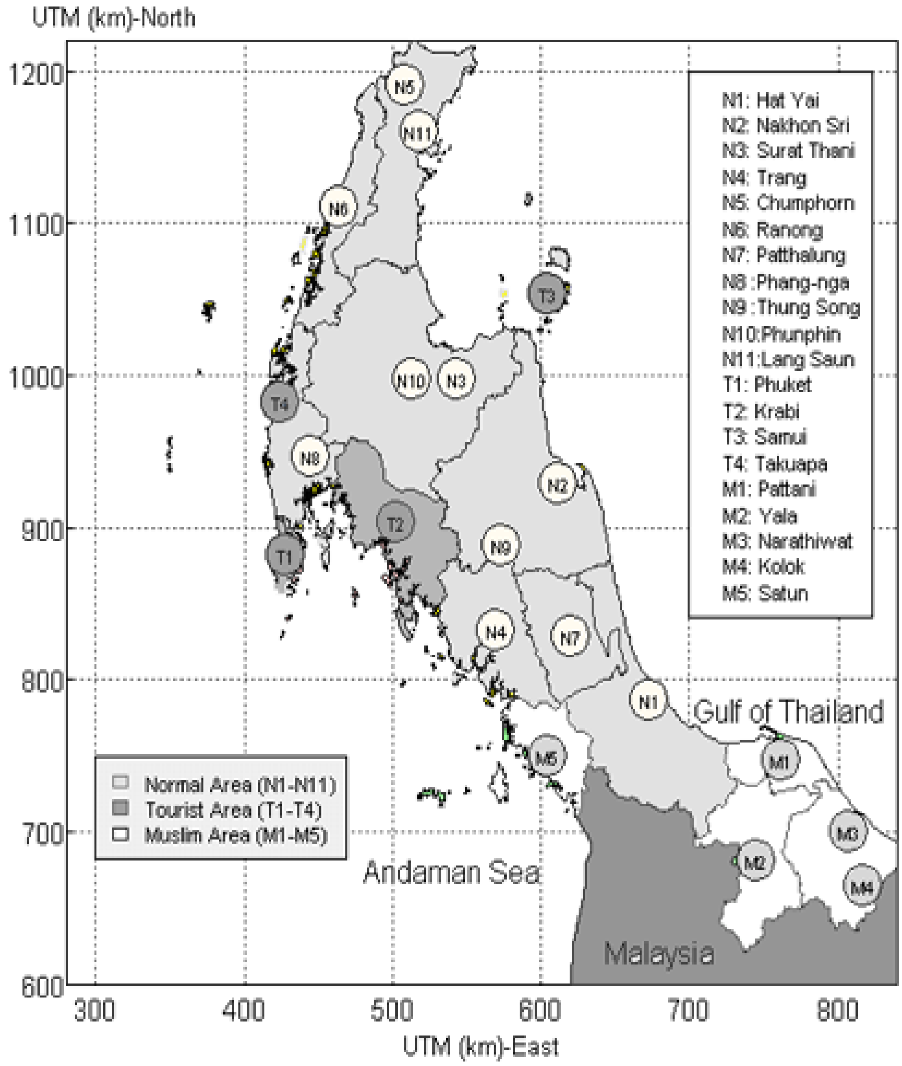

Figure 3. The locations of the branches and their definitions as tourist, Muslim or normal areas

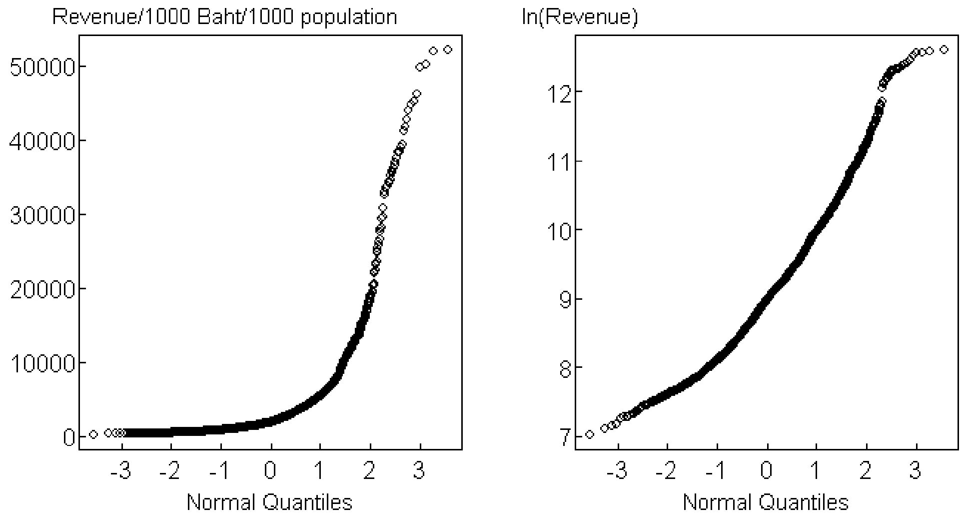

Figure 4. Sales revenue distribution before and after transforming to
ln (Baht/1000 population)

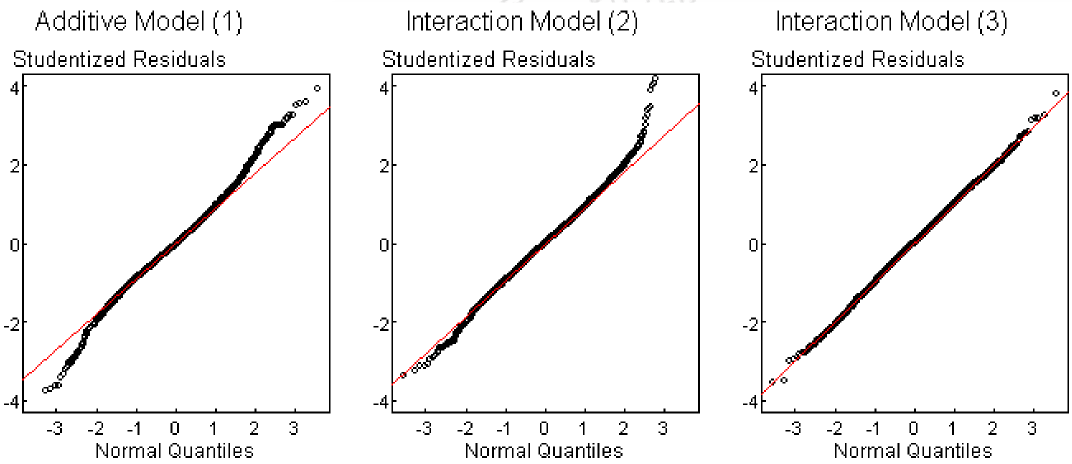

Figure 5. Comparison between residual plots for three models

Figure 6. Plots of sales revenue and fitted values for Model (3)

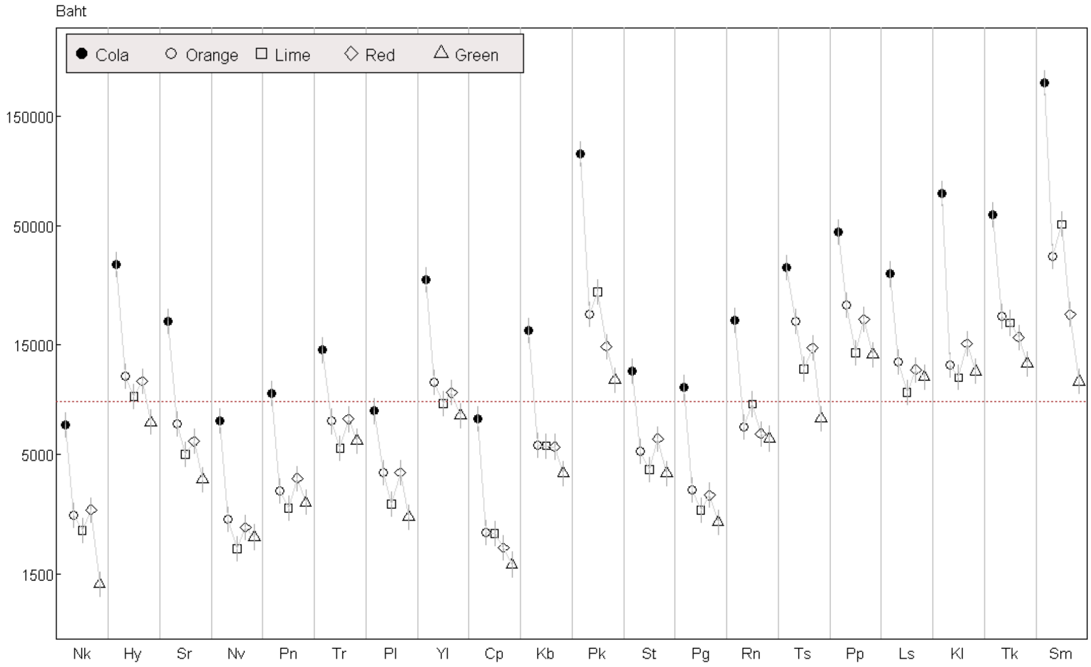

Figure 7. Sales revenue per 1000 population by branch location and flavour

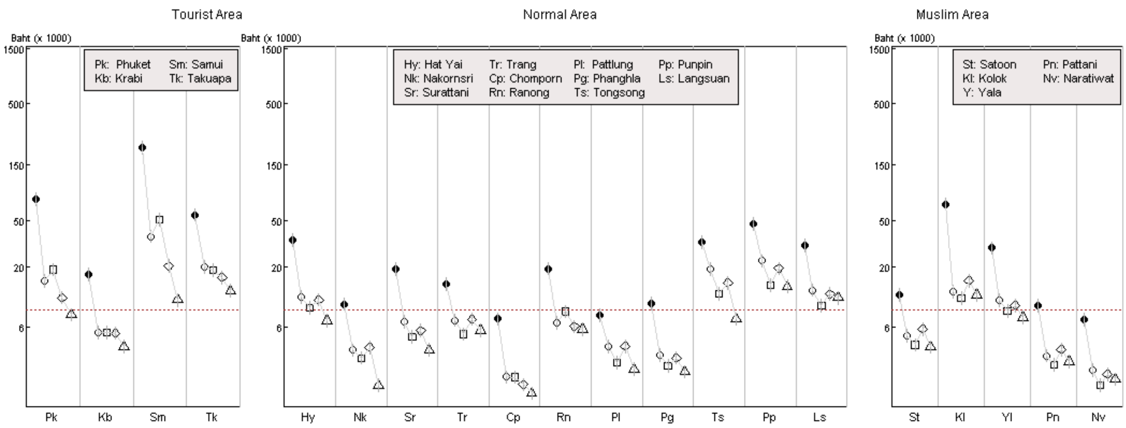

Figure 8. Sales revenue grouped by area per 1000 population

Appendix II

Statistical Model for Short-Term Forecasting Sparkling Beverage Sales in Southern Thailand

Statistical Model for Short-Term Forecasting Sparkling Beverage Sales in Southern Thailand

Wassana Suwanvijit, (Email: swassana@gmail.com), Thaksin University, Thailand
Chamnein Choonpradub, (Email: cchamnein@bunga.pn.psu.ac.th), Prince of Songkla University,Thailand
Nittaya McNeil, (Email: nittaya@bunga.pn.psu.ac.th), Prince of Songkla University, Thailand

ABSTRACT

This study developed a simple statistical model for forecasting a company's sparkling beverage sales in the 14 provinces of Southern Thailand. Data comprised sales revenue from January 2000 to December 2006 obtained from the company. We fitted an observation-driven multiple regression model to log-transformed monthly revenue containing season of year (month), location and beverage flavour as factors, as well as lagged observations for the preceding four months. The model gave a r-squared of 0.95 and was effective for forecasting revenues for up to 12 future months. Using such models for forecasting sales revenue can assist company managers with planning more effectively.

Keywords: sales revenue, multiple regression model, Sales forecasting, Southern Thailand, Sparkling beverages.

INTRODUCTION

Forecasts are basic inputs for many kinds of decisions in business organizations. Forecasts help managers by reducing some of the uncertainty, thereby enabling them to develop more useful plans. Business forecasting involves more than predicting demand. Forecasts are also used to predict profits, revenues, costs, productivity changes, prices and law materials (Kran 2008). Sales forecasting helps to set sales targets and to plan production, marketing, distribution, etc. (Lingham 2004). Accurate sales forecasts facilitate effective and efficient allocation of scarce resources. Over-estimates of demand lead to several problems. First, excess inventory uses up valuable shelf space and leads to obsolescence. Next, scarce working capital blocked up in inventory carrying charges. Third, storage charges are incurred to store excess inventory in public or private warehouses. Finally, margins are reduced when excess inventory is removed through end-of-year clearance sales. Under-estimates of demand lead to a different set of problems. First, stock-outs lead to wasted shelf space. Next, insufficient inventory leads to lost sales and consequent lost margins. Third, failure to keep up with customer demand may necessitate the use of limited and expensive overtime production leading to lower profitability. Finally, and most importantly, the firm may lose customers when prospects facing an empty store shelf try an alternative brand or go to an alternative store and are satisfied by the competitive offering. Given the detrimental impact of inaccurate forecasts, marketers use a variety of sales forecasting techniques in order to forecast sales accurately. Lin and Hsu (2002) applied the Grey model to forecast sales of eight subcategory non-alcoholic beverages in Taiwan between 2001 and 2003. The advantages of the Grey model are that it can use only a few data to estimate an unknown system and it can use a first-order differential equation to characterize the unknown system's behavior. TSO (2002) described how to use the translation model (a mathematical model in which the process of human-language translation) for forecasting long-term electricity demand for the Republic of Ireland. Celia (2004) used a fuzzy forecasting model for women's casual sales. RNCOS (2007) studied the South Korean food, beverages and tobacco market forecast until 2011 by using an in-depth study and evaluation of the past, current, and future market trends in the food, beverage and tobacco industry of South Korea. Newberne (2007) demonstrated the use of the Holt-Winters model (a model used when data exhibits both trend and seasonality) on common healthcare data series. Dodds (2008) used the Bass model (applied from Riccati equation with constant coefficients) in long-term new product forecasting.

The purposes of this paper are to apply multiple linear regression analysis to short-term forecasting sparkling beverages' sales revenue, to illustrate methods for graphing and statistical modeling available to businesses, and to demonstrate how each method can be implemented using freely available software.

Methods

Steps to forecast are: 1) decide what to forecast, 2) evaluate & analyze appropriate data, 3) select & test the forecasting model, and 4) generate the forecast and monitor forecast accuracy over time.

In this study, the outcome was sales revenue. Determinants were flavour, branch location and seasonal period. Flavours were identified for five types of products, namely "cola product" (cola flavour), "colour products" (orange, red and green flavour) and "lime product" (lime flavour). There were two types of returnable packages (10 oz and 1 liter) and three types of non-returnable packages (1.25 liter, 2 liter and 325 ml). There are 20 branches in 14 provinces.

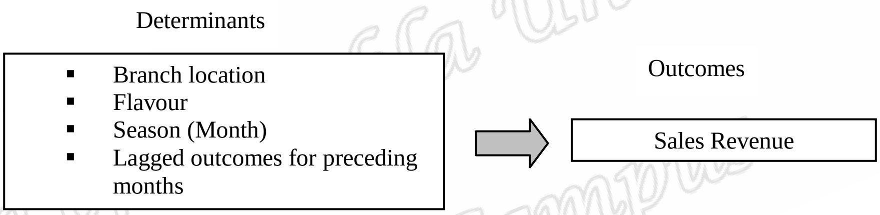

Figure 1: Data Path Diagram

Data for each month were available in computer files with records for sales revenue separated by flavour and branch location. After correcting or imputing data entry errors, records from years 2000 to 2006 were stored in an MySQL database. MySQL and Excel programs were used to create sales revenue in baht by month, flavour and branch location. All graphical and statistical analyses were performed using R (Venables and Smith 2004).

This quantitative research focused on using statistical graphics and statistical models, including multiple linear regression for forecasting sales revenue.

Sales revenues generally have crazy skewed distributions, so it is essential to transform them by taking logarithms. Sales revenue by month and branch location are sometimes nearly zero for small branches and some adjustment is needed to avoid taking logarithms of 0.

Log-transform to ensure that statistical assumptions of symmetry and variance homogeneity errors were satisfied.

The time series of log-transformed monthly sales revenue was plotted to ensure that the sales contains trend and seasonal effects before choosing the best method and model for forecasting.

We fitted a multiple linear regression model to the data and compared results using R software (Venables and Smith 2004). Then, the model of log-transformed sales revenue, which contains seasonal effects and time-lagged terms, was applied for 12-month forecasting.

Statistical Model

We fitted multiple linear regression models to the data, after transforming, using natural logarithms to ensure that statistical assumptions of normality and constant variance were satisfied.

The predictor variables compressed (a) the interactions between branch and flavour, (b) month of the year, and (c) the (log-transformed) sales revenues in the preceding four months. If Y_t is the sales revenue in branch *i*, flavour *j*, of year *y*, month *t*, *s* is the "season-month" (January, February,…) and ε_t is a series of independent normally distributed errors with mean 0, we write

$$\ln \mathrm{Y}_t = \alpha + \mathrm{A}_{ij} + \gamma_s + \delta_1 \ln \mathrm{Y}_{t-1} + \delta_2 \ln \mathrm{Y}_{t-2} + \delta_{11} \ln \mathrm{Y}_{t-11} + \delta_{12} \ln \mathrm{Y}_{t-12} + \varepsilon_t$$

(1)

where α , A and ($\delta_1, \delta_2, \delta_{11}, \delta_{12}$) are parameters in the model denoting an initial value, a trend, and two further coefficients denoting the influence of the sales in the previous four months, respectively, and $\gamma_1 = 0, \gamma_2, \gamma_{11}, \gamma_{12}$ is a set of seasonal effects indicating how the sales revenue varied with month of the year. Forecasts for ln Y_{t+k} (*k* months in the future) are obtained by substituting the estimated values for the coefficients into the right-hand side of (1), using the forecast values themselves for values of *k* > 1. However, to obtain forecasts for $\mathrm{Y}_{t+k'}$ (1) must be transformed back by exponentiation and the forecast is then the mean of $\mathrm{Y}_{1'}$ which has a log-normal distribution with expected value

$$\mathrm{E}[\mathrm{Y}_t] = \exp(\mu)$$

(2)

where μ is the mean, the forecast of Y_{t+k} is

$$\mathrm{E}(\mathrm{Y}_{t+k}) = \exp(\alpha + \mathrm{A}_{ij} + \gamma_s + \delta_1 \ln \mathrm{Y}_{t+k-1} + \delta_2 \ln \mathrm{Y}_{t+k-2} + \delta_{11} \ln \mathrm{Y}_{t+k-11} + \delta_{12} \ln \mathrm{Y}_{t+k-12})$$

(3)

We used associative models that used explanatory variables to predict future sales revenue. The model is a multiple regression model since more than one predictor variable is used to predict sales. The goodness-of-fit of the sales forecasting model is checked with such statistics as r-squared and the standard error of regression relative to the mean and standard deviation of the response variable sales. Later, the partial explanatory power of each predictor variable is checked for expected sign and significance. The error terms are scanned for potential heteroskedasticity (serial auto correlation of the error term) in order to satisfy the assumptions underlying the use of multiple regression analysis.

Prelim Analysis

Figure 2 shows the overall distribution before and after transforming the data by taking natural logarithms of sales revenue. The sales revenue data were symmetry and variance homogeneously for residually. This log-normal distribution can be used to provide an estimate of consumption rate in each month.

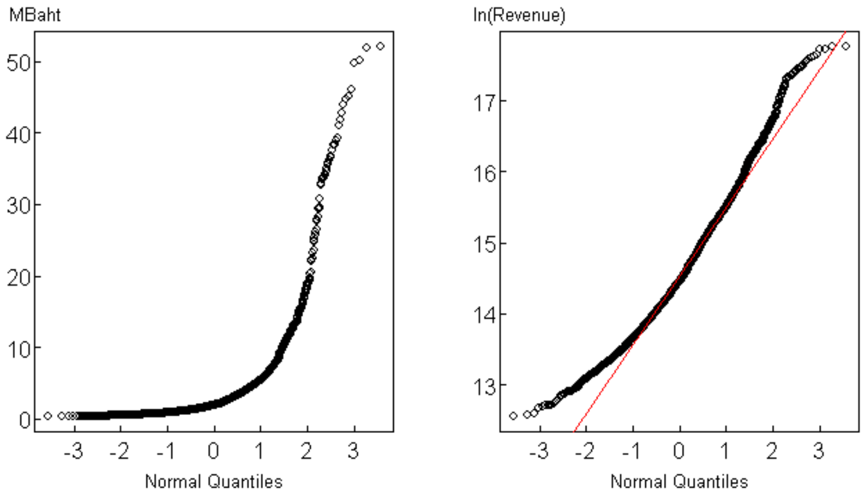

Figure 2: Sales revenue distribution before and after transforming to ln (Baht)

Figure 3 shows a time series plot of the log-transforming actual monthly sales revenue from the years 2000 to 2006. The sales varied between 79.9 and 201.6 MBaht (average value for 2000–2006 is 120.1 MBaht). The sales continuously increased to peak at 201.6 MBaht in 2006. The data contains seasonal effects and there was an upward trend in the last few years.

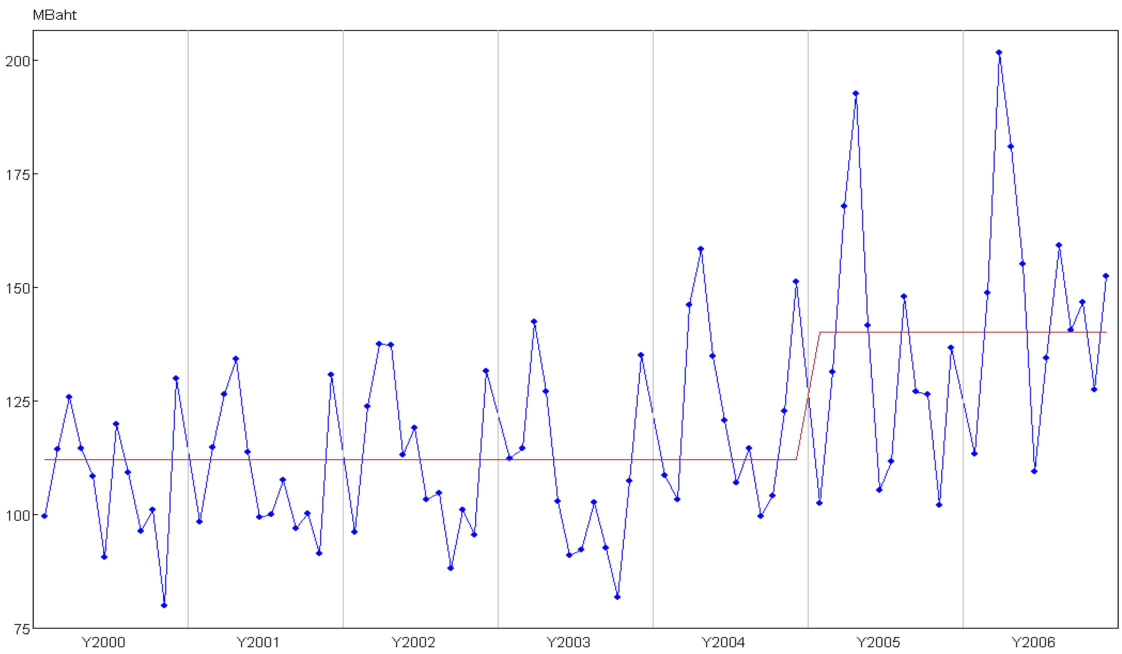

Figure 3: Time series of (log-transforming) actual sales revenue

Results

Figure 4 shows plots of the observed total sales revenue with the natural log-transformed (left panel) and sales revenue (right panel) versus fitted value after fitting the model given by (3).

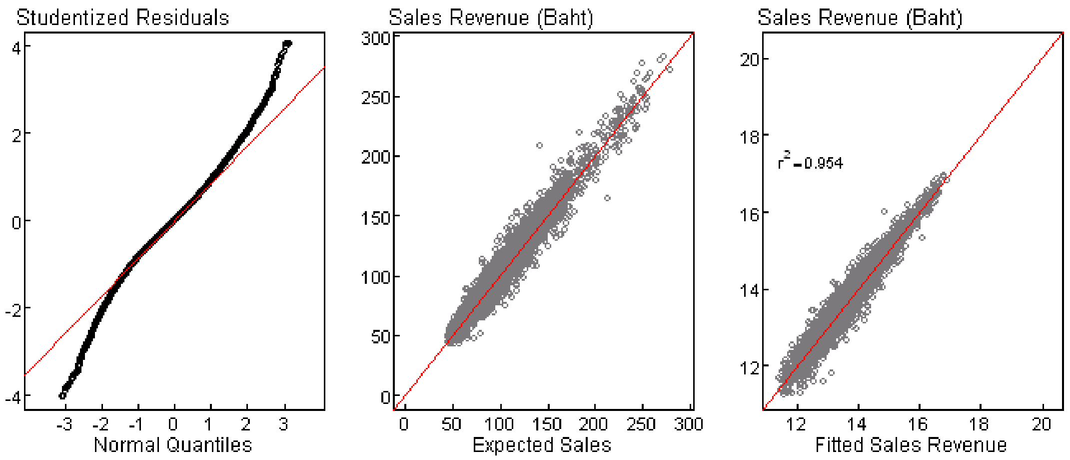

Figure 4: Plots of sales revenue and fitted values from fitting multiple linear regression model to logarithms of monthly sales revenue

Figure 5 shows a plot of the time series of data based on the model given by (3). The monthly forecasts of sales revenue in the company during the twelve months of 2007 were 120.7, 142.7, 172.9, 179.8, 139.8, 121.4, 119.8, 136.2, 128.4, 128.1, 126.0 and 146.5 MBaht.

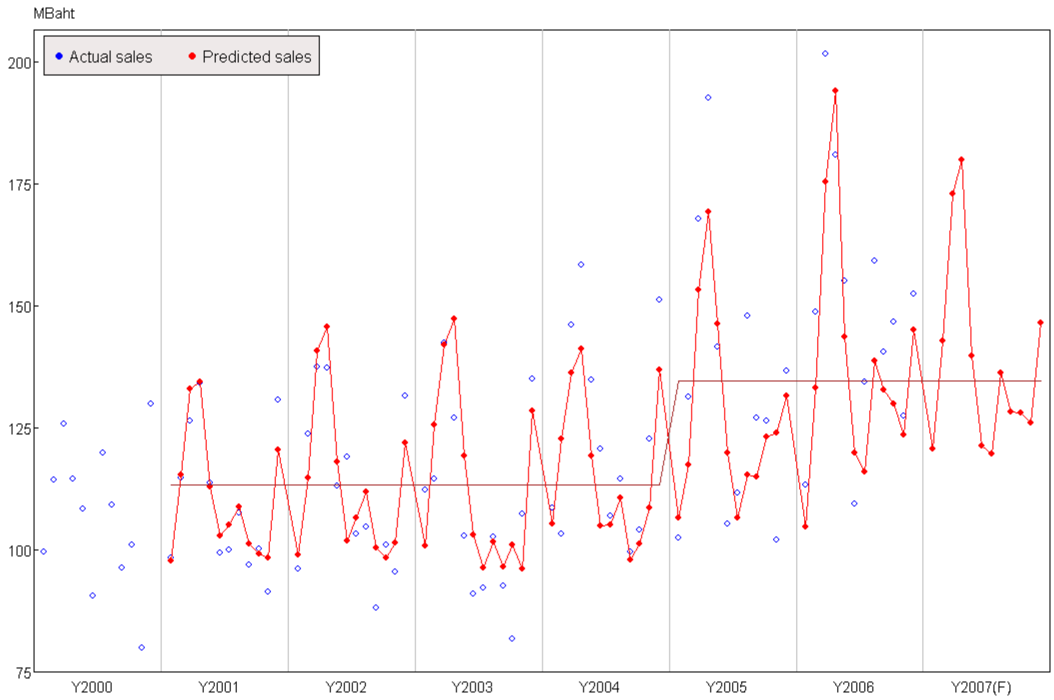

Figure 5: Comparison between actual and forecasting sales revenue

Table 1 shows the forecast and actual values of sales revenue (MBaht) and percentage errors in 2006. The average percentage error of predicted sales revenue was very slight.

Month	Forecast	Actual	% error
January	104.8	113.4	7.6%
February	133.2	148.7	10.4%
March	175.5	201.6	12.9%
April	193.9	181.0	7.1%
May	143.6	155.1	7.4%
June	120.0	109.4	9.7%
July	116.1	134.5	13.7%
August	138.8	159.1	12.8%
September	132.8	140.6	5.5%
October	129.9	146.7	11.5%
November	123.6	127.4	3.0%
December	145.1	152.6	4.9%
Average	138.1	147.5	6.4%

Table 1: Forecast and actual sales revenue of sales revenue (MBaht) in year 2006

Figure 6 shows a plot of the time series of data (all flavours) with the forecasts based on the model given by (3) comparison by package types (returnable and non-returnable packages).

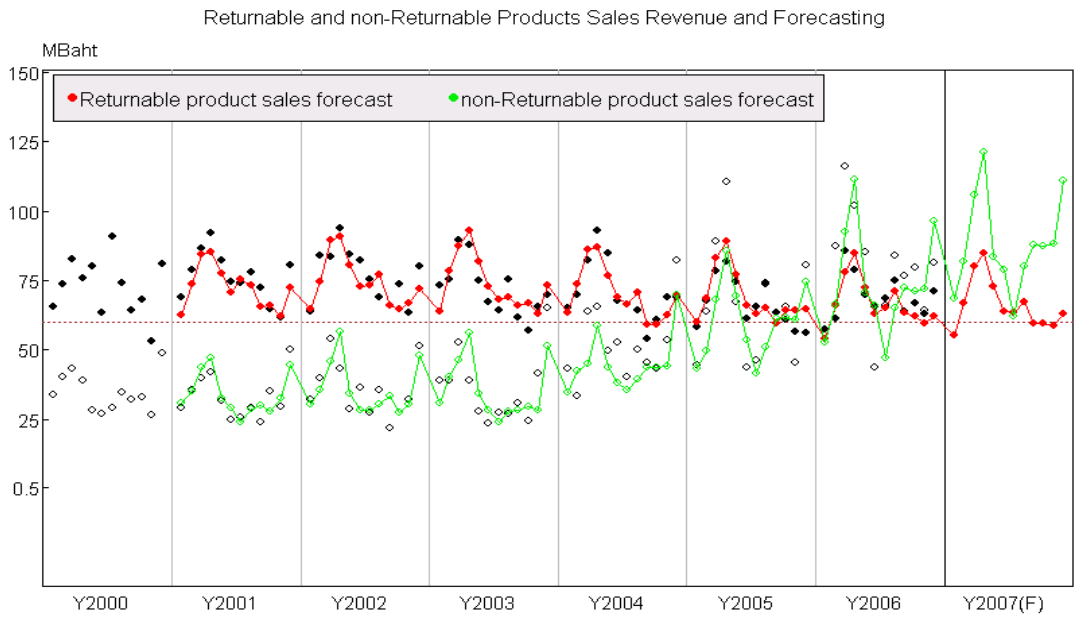

Figure 6: Comparison between actual and forecasting sales revenue grouped by package types

Figure 7 shows a plot of the time series of returnable products sales revenue with the forecasts based on the model given by (3) grouped by flavour (cola, orange, red, green and lime). There was a slightly downward trend in sales revenues.

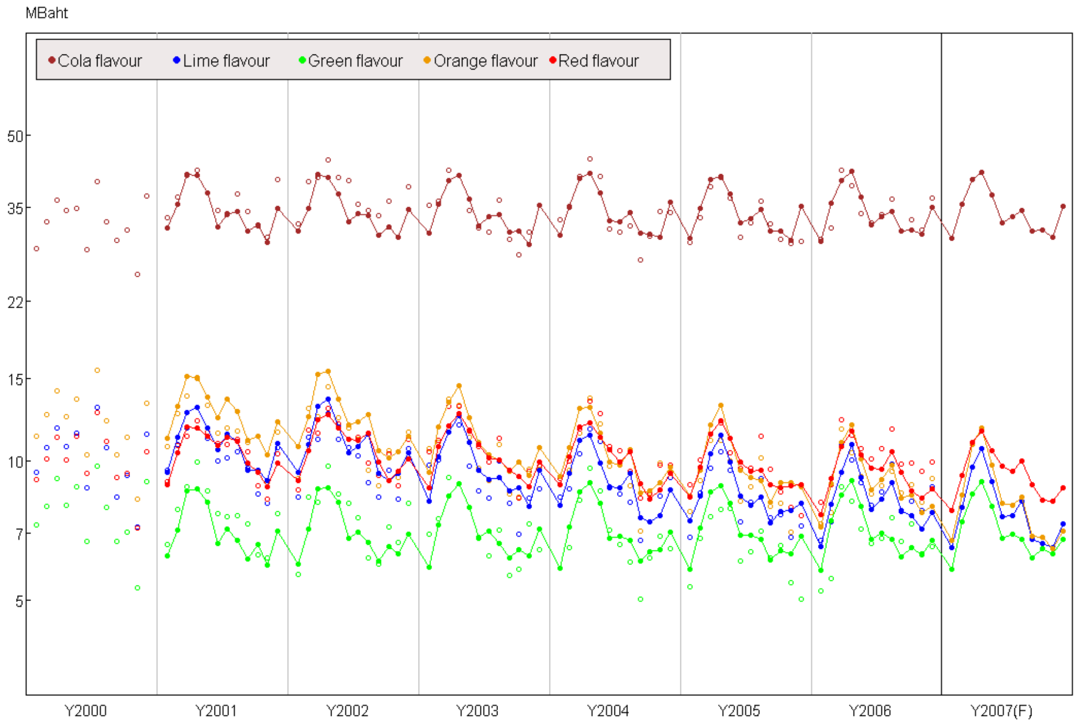

Figure 7: Comparison between actual and forecasting sales revenue of returnable products grouped by Flavour

Figure 8 shows a plot of the time series of non-returnable products sales revenue with the forecasts based on the model given by (3) grouped by flavour (cola, orange, red, green and lime). There was a moderate upward trend in sales revenue of each flavour.

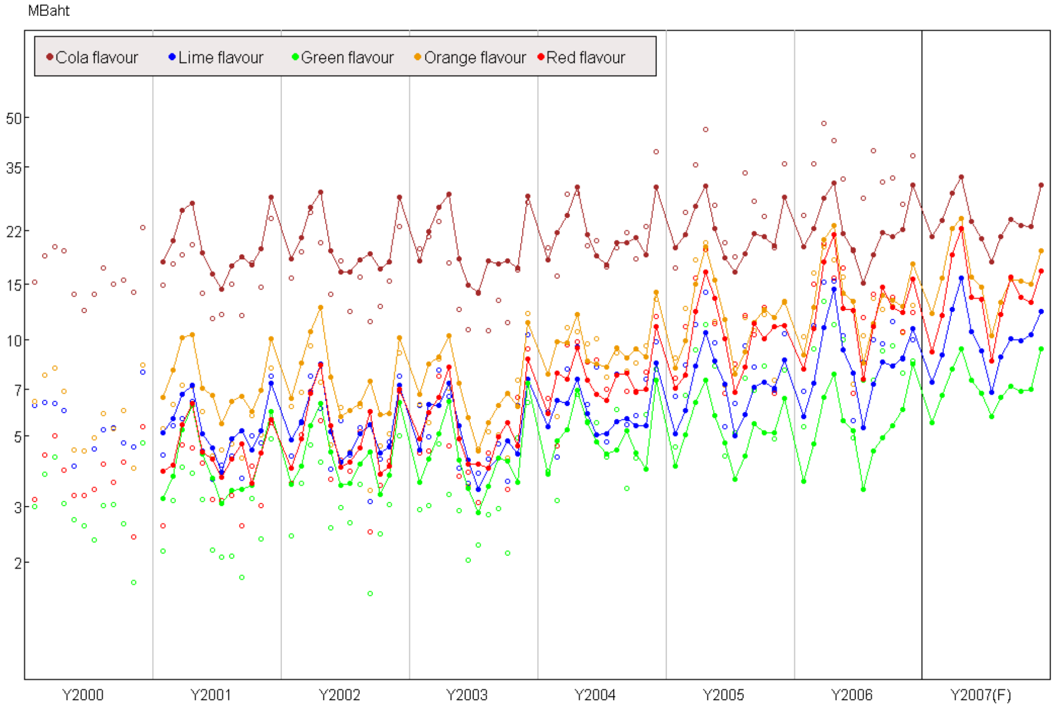

Figure 8: Comparison between actual and forecasting sales revenue of non-returnable products grouped by Flavour

Discussion

Figure 3 shows the time series of monthly sales revenue from January 2000 to December 2006 and reveals that the sales have increased substantially in the last few years, it could be due to an expanding in modern trade channel effects consumer behavior and life style of Southerner. The seasonal effects found in our study could be related both to regional climatic changes and human activities due to there are long week-end and holidays, including to consumers often consume sparkling beverage in the dry (hot) season (extends from February to April).

Figure 4 shows plots of sales after fitting an observation-driven multiple regression model to log-transformed monthly revenue containing season of year (month), location and beverage flavour as factors, as well as lagged observations for the preceding four months. The model predicts the proportions in the 8,400 cells very well since it gave a r-squared of 0.95**.**

Figure 5 shows a plot of the time series of data with the forecasts based on the model given by (3). The fitted value predicts the sales revenue very well and the model was effective for forecasting revenues for up to 12 future months. The results from Table 1 indicate that the percent error of sales forecasting in year 2006, compared with actual sales, was approximated 6.4%, which was very slight. In addition, the model could give the 12-months of sales revenue in year 2007 forecasting that contains both the seasonal and time-lagged term.

Figure 6 shows a plot of the time series of data with the forecasts based on the model given by (3), grouped by package types (returnable and non-returnable packages). As we can seen, the returnable products is in a negative situation trend, but even so the non-returnable products sales have a moderate growth of sales revenue since the consumers behavior change, most consumers prefer convenience packages such as PET bottles and can.

Figure 7 shows a plot of the time series of returnable products sales with the forecasts based on the model given by (3), grouped by flavour (cola, orange, red, green and lime). There was a slightly downward trend in sales revenues of all flavours, except Cola product. Colour products has a fluctuation due to in the early years orange flavour has the most value in colour product. However, red flavour becomes has more value than orange in a last few years. It is possible that now there are a lot of orange juice products in the market, so consumers have more opportunities to buy this kind of product. Even if the sales revenue from sparkling beverage company in the Southern Thailand accounts for a large part of market share in the south, there are some direct and indirect competitors.

Figure 8 shows a plot of the time series of non-returnable product sales with forecasts based on the model given by (3), grouped by flavour (cola, orange, red, green and lime). There is a slight growth trend with the Cola flavour, while there is moderate growth with lime and colour flavours.

AUTHOR BIOGRAPHIES

Wassana Suwanvijit is a lecturer at faculty of Economics and Business Administration, Thaksin University, Thailand, and a marketing advisor at Sparkling Beverage Company. She also is currently a PhD candidate in Research Methodology.

Chamnein Choonpradub (PhD, Assistant Professor) is a lecturer for the Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University (Pattani campus), Thailand.

Nittaya McNeil (PhD) is a lecturer for the Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University (Pattani campus), Thailand.

CONCLUSIONS

The observation-driven multiple regression model to log-transformed monthly revenue containing season of year (month), location and beverage flavour as factors, as well as lagged observations for the preceding four months to log-transformed monthly revenue model, was effective for forecasting total sales revenues, including sales revenue grouped by flavours and package types for up to 12 future months. The model can be applied for forecasting in other business data. Using such models for forecasting sales revenue can assist company managers in planning more effectively.

ACKNOWLEDGEMENTS

We are grateful for Prof. Don McNeil for his helpful advice and suggestions. We also would like to thank Khun Dumrongrugs Apibalsawasdi for his guidance.

REFERENCES

1. Bureau., Indian Food and Beverages Forecast (2007-2011), June, 2007.
2. Caruana, Albert., Step in forecasting with seasonal regression: A case study from the carbonated soft drink market, Vol. 10, No. 2, pp.94-102, 2001.
3. Celia, Frank., Les, Sztandera., Balaji, Vemulapali., Asish, Garg., Amar, Raheja., Fuzzy Forecasting Model for Women's Casual Sales. National Textile Center Annual Report, November, 2004.
4. Cox, D.R. and Wermuth, N., Multivariate Dependencies-Models, Analysis and Interpretation, London: Chapman & Hall, 1996.
5. Dodds, Wellesley., An Application of the Bass Model in Long-Term New Product Forecasting. Journal of Marketing Research (JMR), 2008.
6. Hanke, John E., Wichern, Dean W., Business Forecasting., Pearson Education, New Jersy, 2005.
7. Higgins, Richard S., Kaplan, David R. and McDonald, Michel J., Residual Demand Analysis of the Carbonated Soft Drink Industry. Empirica, Vol.22, No.2, pp.115-126, 2005.
8. Kran, T.A., Equipment Productivity, LMDA, 2008.
9. Lin, Chin-Tsai., Hsu, Pi-Fang., Forecast of non-alcoholic beverage sales in Taiwan using the Grey theory. Asia Pacific Journal of Marketing and Logistics, Vol.14, No.4, pp.3-12, 2002.
10. Lingham, Leo., Marketing-Acquisition & Analysis of information, 2004. http://en.allexperts.com/q/Marketing-1090/Marketing-Acquisition-Analysis-information.htm. [March 20, 2009)
11. Newbeme, Joan H., Captain, USAF., MSCHolt-Winters Forecasting in Healthcare, Mike O'Callaghan Federal Hospital, 2007.
12. R Development Core Team., R:A language and environment for statistical computing, Vienna, 2007.
13. RNCOS., South Korean Food, Beverages and Tobacco Market Forecast till 2011, December, 2007.
14. Tinney, Mary-Colleen., Retail Sales Analysis: Champagne and Sparkling Wine Sales Rising, Wine Communications Group, 2007.
15. TSO., A New Methodology for forecasting long term electricity demand for the Republic of Ireland. Electricity Demand Model Review, Generation Capacity Planning, September, 2002.
16. Venables, W.N., Ripley B.D., Modern Applied Statistics with S. Springer, Queensland, 2002.
17. Venables, W. N., Smith, D.M., An Introduction to R, 2004. http://cran.r-project.org/doc/manuals/R-intro.pdf [March 30, 2007]
18. Williams., Developing Sales Forecasts, 2007. http://web.csustan.edu/market/williams/Chapter%2010.htm [March 14, 2008]

Appendix III

Long-Term Sales Forecasting using Lee-Carter and Holt-Winters Methods

Long-Term Sales Forecasting using Lee-Carter and Holt-Winters Methods

Wassana Suwanvijit, (Email: wassana@tsu.ac.th), Thaksin University, Thailand
Prof. Thomas Lumley, PhD (Email: tlumley@u.washington.edu), University of Washington, USA
Assist. Prof. Chamnein Choonpradub, PhD (Email: cchamnein@bunga.pn.psu.ac.th), Prince of Songkla University (Pattani), Thailand
Nittaya McNeil, PhD (Email: nittaya@bunga.pn.psu.ac.th), Prince of Songkla University (Pattani), Thailand

ABSTRACT

This study developed a statistical model for long-term forecasting sparkling beverage sales in the 14 provinces of Southern Thailand. Data comprised the series of monthly sales from January 2000 to December 2004 obtained from the company. We applied a classical Lee-Carter mortality forecasting approach as well as exponential smoothing Holt-Winters with additive seasonality method to log-transformed monthly sales containing season of month and branch location as factors. The model produced excellent estimates in sales predicting for up to 24 future months of 20 branches compared with actual data in each branch during the years 2005-2006. The model also gave more accurate results than using separate forecasting method whereas it was parsimonious in the number of parameters used.

Keywords: Long-term, Sales forecasting, Lee-Carter approach, Holt-Winter method, Sparkling beverage

INTRODUCTION

Research aim

The research aims to investigate suitable statistical methods for long-term forecasting of sparkling beverage sales revenue data collected routinely in 14 provinces of Southern Thailand during years 2000-2004, including to demonstrate how each method can be implemented using free available software.

Initial assumption of the paper

Forecasting is about predicting the future as accurately as possible. Sales forecast is the amount of a product that the company actually expects to sell during a specific period. A simplistic but useful approach is to begin with a sales forecast where an initial assumption is made that the future sales trend will follow to historical sales or have a similar pattern. Based on our initial assumption, we can identify the general level of sales. We can also determine whether there is a pattern or trend, such as an increase or decrease in sales revenue over time.

Reasoning for the focus of the paper

Long-term sales forecasting is a difficult area of management. However, it is a subject of great interest as a result of the persistent tendency of company performance and it is indispensable for business planning and strategy. Forecasts help managers by reducing some of the uncertainty, thereby enabling them to develop more useful plans. It also provides useful information for intelligent business decisions making. Modern organizations require sales forecasts and long-term sales forecasts are used in strategic planning. For these *reasons*, the long-term sales forecast is the main *focus* of this *paper*.

Previous researches

Several statistical models have been used for business data forecasting in the previous researches. Software World (2009) developed computer model for forecasting beer consumption in UK by applying Mathematics techniques of correlation analysis, regression analysis. The model would have been accurate within 10% on 90% of occasions for forecasting monthly consumption. RNCOS (2008) forecast Philippines, beverages and tobacco market till 2011 using ratio analysis, historical trend analysis and linear regression analysis. Information has been sourced from books, newspapers, trade journals, and white papers, industry portals, government agencies, trade associations, monitoring industry news and developments, and through access to more than 3000 paid databases. The report provides detailed overview of the consumption patterns of the Philippines in various food segments. The beverage segment talks about the type of beverages, their sales and consumption patterns among Philippines while the tobacco segment provides a brief description of the tobacco industry in the country. Boonruangthaworn (2007) studied about forecasting techniques for soft drink industry in order to predict the demand of a new product so that the production planning and inventory control are more efficient. They collect the historical demand data from 2004 to 2006. Then, implement both qualitative and quantitative forecasting models and compare their forecasting error such as MSE, MAD and MAPE. Then, the best model is selected as an input of the production planning. Finally, they compare total cost by using the result from both techniques. The results show that Least Square is the best forecasting method with the lowest forecasting error. In addition, by using this technique, the related cost can be reduced by 90% compared to that of the current technique. The result show that this quantitative forecasting technique is more appropriate than the current qualitative forecasting technique by considering cost and service level criteria. Higgins et al. (2005) analysis the residual demand to test whether carbonated soft drinks is a relevant product market using weekly A.C. Nielsen Scanner price and quantity data for carbonated soft drink products purchased in supermarkets in the United States. The results suggest that a market for carbonated soft drinks is too narrow for purposes of merger analysis according to the Merger Guidelines established by the United States Department of Justice and the Federal Trade Commission. In the case study of carbonated soft drink consumption and bone mineral density in adolescence by McGartland et al (2003), adjusted regression modeling was used to investigate the influence of carbonated soft drinks on bone mineral density. Lin et al (2002) forecasted non-alcoholic beverage sales in Taiwan. The study applied the Grey dynamic model to forecast sales of eight sub-category non-alcoholic beverages in Taiwan between 2001 and 2003. The accuracy of the new forecasting model exceeds 95 percent. The model estimates that the total beverages market will grow, but growth rates will vary for individual sub-categories. In relation to current growth, from 2001 to 2003, tea drinks, carbonated drinks, functional drinks and sports drinks will experience decreased market growth, while bottled water and fruit and vegetable juices will be a high growth market and coffee drinks and other drinks will enjoy improved sales. These results provide a valuable reference for the Taiwanese beverage industry developing marketing plans.

Research and epistemological approaches

The research and epistemological approaches that have been adopted to conduct this research are Lee-Carter model and Holt-Winters exponential smoothing method. These are different methods compared with the methods used for sales forecasting in the previous researches. However, these approaches basically involve the theoretical framework that has to be adopted in this study.

Originality of the paper and contribution to knowledge

This study attempts to extend the knowledge about long-term forecasting to produce useful finding for both company and further researches. The originality of the paper lies in the methodologies we apply for long-term forecasting of sparkling beverage sales revenue and the accurate of the sales forecast compared with actual value and the results from using the separate forecast method. A new forecasting model is presented and the model can be extended into other specific business data analysis and forecasting using the same methods.

Area of study

Thailand is divided into four geographical regions. The southern region occupies about 14% of the total land area. There are 14 provinces in southern Thailand with a total area 71,798 square kilometers. Sparkling beverages are traditionally popular products in the south, although some consumers prefer more healthy beverages. Sparkling beverages are also primarily used as mixers for consumption with alcoholic drinks. There are 20 branches in Southern Thailand. The branch locations are shown in Figure 1.

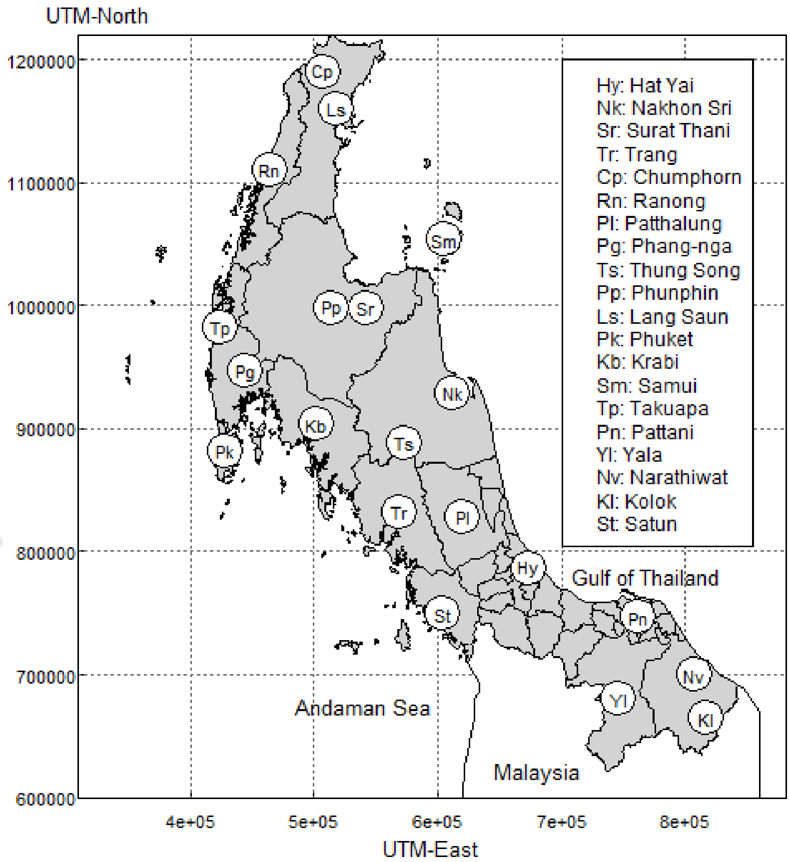

Figure 1. Branch Locations in Southern Thailand

Preliminary analysis and results

The monthly sales revenue was plotted to assess trend and branch effects before choosing the best model for forecasting. Let Y_x be the logarithm of the sales for branch x (b_x) (where x = 1, .., 20) and ε_x is a vector of error terms, the simple linear regression model can be written as

$$\log \mathrm{Y}_x = b_x + \varepsilon_x \qquad (1)$$

Figure 2 shows the result from the linear regression fitted given by (1) in the left panel and the time series plot of some branches in the right panel. The preliminary analysis using historical data indicates that the monthly sales in each branch location have increased substantially. However, each branch has different patterns in the growth rate compared with the others. The branch location is a factor of interest in the models since it affects the sales trend. The time series plot also shows seasonal effect on the sales in each branch.

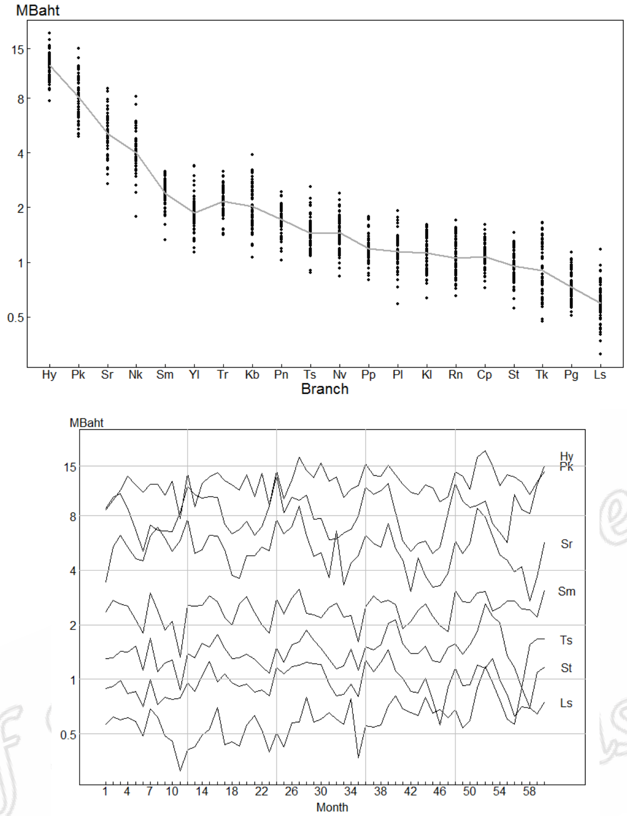

Figure 2: Monthly sales by branch (left panel) and time series plot by branch (right panel)

THEORETICAL FRAMEWORK

In this study, the statistical method for long-term sales forecasting was selected based on empirically tested theories (Baker 1999). Figure 3 shows theoretical basis of the study. The theoretical perspective in this study is based on statistical forecasts using extrapolation models and quantitative forecasting method. Quantitative forecasting methods are used when historical data on variables of interest are available. These methods are based on an analysis of historical data concerning the time series of the specific variable of interest and possibly other related time series. It is preferable to use this method rather than judgmental extrapolations because extrapolation methods become more useful and less expensive as one can work directly with time-series data on sales. Extrapolation methods use only historical data on the series of interest and quantitative extrapolation methods make no use of managements' knowledge of the series. The theory assumes that the causal forces that have affected a historical series will continue over the forecast horizon. Extrapolation of sales can be used to predict for the situation that is of interest that often adequate for the decisions that need to be made.

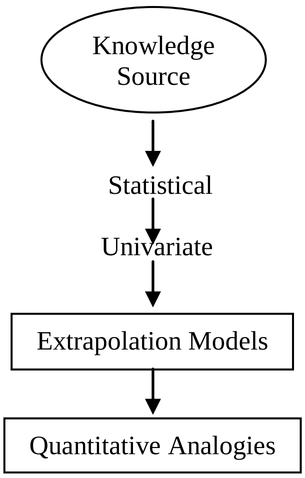

Figure 3: Theoretical basis of study

RESEARCH MODEL

Steps to forecast are: 1) decide what to forecast, 2) evaluate & analyze appropriate data, 3) select & test the forecasting model, and 4) generate the forecast and monitor forecast accuracy over time. Figure 4 shows the research model of this study. We indentify what to forecast and research aims in the first step. Then, the in depth literature review and literature review are explored. Next, we select the right sales forecast methodology based on empirically tested theories. After that, we implement the forecasting methodology including to evaluate and analyze the forecasting results and the errors in order to find the suitable model. Finally, we do the discussion, summary and recommendation for further studies.

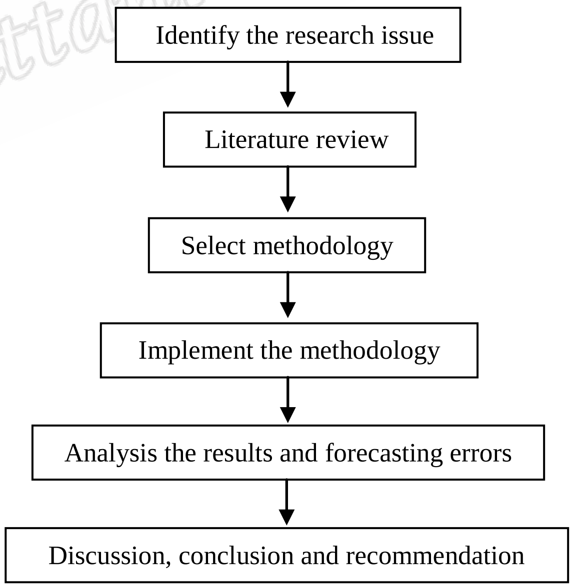

Figure 4: Research model of study

LITERATURE REVIEWS

From the theoretical framework as shown in Figure 3, the theoretical perspective in this study is based on statistical forecasts using extrapolation models and quantitative forecasting method. The non-linear Lee-Carter approach was designed for long-term forecasting based on a lengthy time series of historic data. The most popular and cost effective of extrapolation models and quantitative forecasting method are based on exponential smoothing, which implements the useful principle that the more recent data are weighted more heavily. The Holt-Winters exponential weight method is also popular for long-term forecasting of business data. Therefore, a classical Lee-Carter mortality forecasting approach and exponential smoothing Holt-Winters method were applied for long-term sales forecast in this study**.**

The non-linear Lee-Carter approach (Lee and Carter 1992), is widely used in both the academic literature and practical applications and it has become the "leading statistical model of mortality forecasting in the demographic literature" (Deaton and Paxson 2004). The method was designed for long-term forecasting based on a lengthy time series of historic data. In the typical application, this model is fitted to past data to obtain parameter estimates, then it produces an excellent fit to mortality trends by linearized trends and thereby adds confidence to extrapolations. Since the Lee-Carter model is computationally simple to apply and it has given successful results, it was popular used for long-term forecasts of age specific mortality rates from various countries and time periods, such as USA (Lee and Carter 1992), Canada (Lee and Nault 1993), Chile (Lee and Rofman 1994), Japan (Wilmoth 1996), the seven most economically developed nations (G7) (Tuljapurkar et al 2000), Belgium (Brouhns et al 2002) and Sweden (Wang 2007). The model can also be applied for seasonality and non-linearity data such as using to price a risky coupon survivor bond (Denuit et al 2007) including to describe seasonal variation and non-linear for quarterly industrial production (Franses et al 2005). The original principal component of the Lee -Carter model is

$$\log(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \qquad (2)$$

with mortality $m_{x,t}$ at age x and time t, fixed age effect a_x equal to the average observed log death rate and an age-specific impact b_x of a time-specific general mortality index $k_t \varepsilon_{x,t}$, here is a set of random disturbances. This single parameter k_t maps as the average age pattern of mortality deviation from a_x to the actual pattern. b_x is the first principle component which is estimated by singular value decomposition method. Constraints imposed to obtain a unique solution are the b_x sum to unity and the k_t sum to zero. The subsequent estimation of the mortality index k_t as a time series linear forecasting model: ARIMA (autoregressive integrated moving average) process results in a simple random walk with drift. The method adjusts k_t by refitting to total observed deaths.

The underlying principle of the Lee-Carter method is the extrapolation of past trends and makes no effort to incorporate knowledge about medical, behavioral, or social influences on mortality change. This method allows age-specific death rates to decline exponentially without limit and assumes that the model errors have the same variance over all ages. Several advantages of the model are claimed such as a parsimonious demographic model is combined with statistical time-series methods, the method involves no subjective judgments, forecasting is based on persistent long-term trends and probabilistic confidence intervals are provided for the forecasts (Lee and Carter 1992). The model is also computationally simple to apply with robustness in the context of linear trends in age-specific death rates. Differ criteria have been proposed for the Lee-Carter method. Wilmoth (1993) developed two alternative one-stage estimation strategies which are a weighted least square (WLS) and a maximum likelihood (MLE) technique. Bell (1997) suggested using a multivariate time series model for all coefficients and do not exploit the orthogonality of the coefficient series. There have been several extensions of the Lee-Carter method such as non-parametric smoothing, Kalman filtering, and multiple principle components.

Exponential smoothing is a procedure for continually revising a forecast in the light of more recent experiences. Exponential smoothing assigns exponentially decreasing weights as the observation get older. In other words, recent observations are given relatively more weights in forecasting than the older observations (Kalekar 2004). Exponential smoothing methods are among the most widely used forecasting techniques in industry and business, in particular the well-known Holt-Winters methods

(Holt 1957 and Winters 1960) that allow us to deal with univariate time series with contain both trend and seasonally factors. Their popularity is due to their simple model formulation and good forecasting results (Gardner 1985). Holt-Winters is popular for mass produced forecasts, for example in production planning, because of its simplicity (ONS 2008). Kotsialos et al (2005) used of a damped-trend Holt-Winters method and feedforward multilayer neural networks to forecast sales data from two German companies up to 52 periods ahead. Bermudez et al (2005) applied additive Holt-Winters forecasting procedure to the series of monthly total UK air passengers from the year 1949 to 2005. Newberne (2007) demonstrated the use of the Holt-Winters model on common healthcare data series.

METHODOLOGIES

Research method

This quantitative research focused on the classical Lee-Carter model and the well-known Holt-Winters with additive seasonality method to log-transformed monthly sales containing season of month and branch location as factors. We fitted the models to the data. Then, the models of sales revenue, which contains trend and seasonal effects were applied to estimate in sales predicting for up to 24 future months of 20 branches in Southern Thailand. All graphical and statistical analyses were performed using R (R development core team 2008).

Procedure and data collection

Figure 5 show procedures in this study. Monthly data was available in computer files with records for sales revenue separated by branch location. After correcting or imputing data entry errors, records from years 2000 to 2004 were stored in a MySQL database. MySQL and Excel programs were used to create sales revenue in Baht by month and branch location. Sales revenues generally have skewed distributions, so it is essential to transform them by taking logarithms. Log-transformations can also ensure that statistical assumptions of symmetry and variance homogeneity of errors are satisfied. Figure 6 shows the overall distribution before and after transforming the data by taking logarithms of the sales. It shows that the distribution is more symmetric after transforming the data.

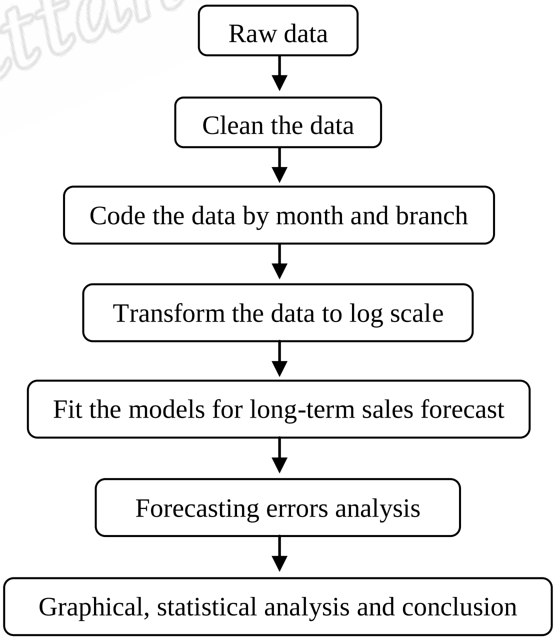

Figure 5: Procedures of study

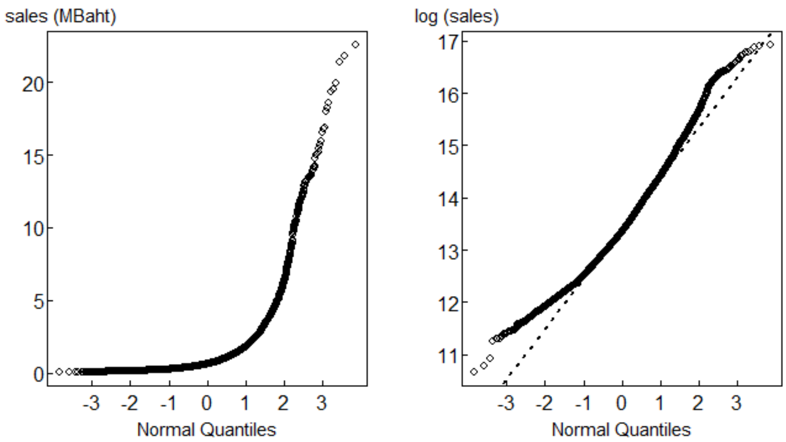

Figure 6: Sales distribution before and after transforming to log (Baht)

Lee-Carter model

Let Y_{xt} be the logarithm of the sales for branch x (where x = 1, .., 20) at month t (where t =1,…, 60), the Lee-Carter model with principal component is

$$\log Y_{xt} = a_x + b_x k_t + \varepsilon_{xt} \tag{3}$$

where a_x is the average sales by branch which is constant over time
b_x is the changes in the sales at branch x in response to changes in k_t over time
k_t is the temporal trend of sales changes over time
ε_{xt} is a vector of error terms

with the constraints: $\sum b_x = 1,\ \sum k_t = 0$ (4)

Lee-Carter model with 2 components extension can be written as

$$\log Y_{xt} = a_x + b_{x1}k_{t1} + b_{x2}k_{t2} + \varepsilon_{xt} \tag{5}$$

Lee-Carter model with 3 components extension can be written as

$$\log Y_{xt} = a_x + b_{x1}k_{t1} + b_{x2}k_{t2} + b_{x3}k_{t3} + \varepsilon_{xt} \tag{6}$$

We estimated the average sales (a_x) by

$$a_x = \log \prod_{t=t_1}^{t_n} Y_{xt}^{\frac{1}{n}} \tag{7}$$

For the least squares estimation, the Singular Value Decomposition (given by the R function “svd”) was applied to the average sales over time t for each branch x.

$$\log Y_{xt} - a_x = UDV', \tag{8}$$

where D is a diagonal matrix containing singular values and both U and V are orthogonal matrices. The parameters b_{x1}, b_{x2}, b_{x3} are set equal to the first, second and third column of U respectively, and the k_{t1}, k_{t2}, k_{t3} values are set equal to the product of the first, second and third column of V and the leading singular value d_1, d_2, d_3 respectively along with the normalizations given in (4). In order to make more accurate in the forecasting results, we adjust b_{x1}, b_{x2}, b_{x3} by comparison with an average of the last 12 months observation data.

In the next step, we fitted the Lee-Carter model given by (3), (5) and (6) by using estimated parameters including to the b_{x1}, b_{x2}, b_{x3} adjusted from the previous step. Then we compared the results from fitting the Lee-Carter model with the 60 months observation data. After that we adjust k_{t1}, k_{t2}, k_{t3} for the best fit of each model. The estimated parameters for the Lee-Carter model with 1-3 components are shown in Figure 7.

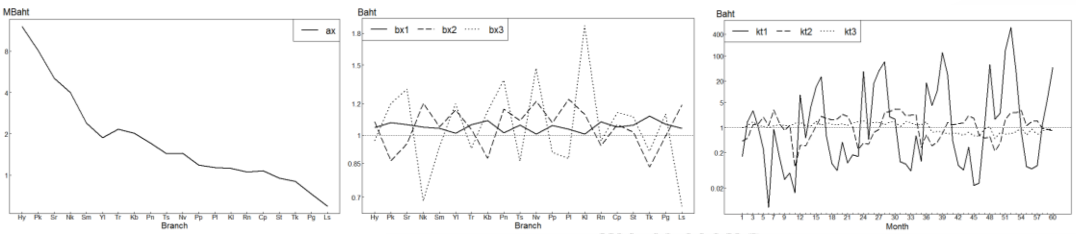

Figure 7: Parameters (a_x, b_{x1}, b_{x2}, b_{x3}, k_{t1}, k_{t2} and k_{t3}) for the Lee-Carter model plots

Holt-Winters method

Since the historical data series are seasonal with linear trend, we forecast k_{t1}, k_{t2}, k_{t3} adjusted values for up to 24 months ahead as well as their 95% robust prediction intervals by using Holt-Winters exponential smoothing with additive seasonality forecasting method. The Holt-Winters prediction function (for time series with period length p) is

$$\hat{K}_{t+h} = a_t + hb_t + s_{t+1+(h-1)\bmod p} \tag{9}$$

where a_t, b_t and s_t are given by $a_t = \alpha\,(K_t - s_{t-p}) + (1-\alpha)\,(a_{t-1} + b_{t-1})$

$$b_t = \beta\,(a_t - a_{t-1}) + (1-\beta)\,b_{t-1}$$

$$s_t = r(K_t - a_t) + (1-r)\,s_{t-p}$$

Once k_{t1}, k_{t2}, k_{t3} forecast are built, the monthly sales forecast in each branch can be predicted easily by just fitting the Lee-Carter model using k_{t1}, k_{t2}, k_{t3} forecasted as well as a_x given by (7) and the b_{x1}, b_{x2}, b_{x3} adjusted. So, let z be the 24 months in the future, the prediction of $\mathrm{Y}_{x,t+z}$ for Lee-Carter model with 3 components extension was thus obtained from

$$\log \hat{\mathrm{Y}}_{x,t+z} = a_x + b_{x1}k_{t1+z} + b_{x2}k_{t1+z} + b_{x3}k_{t1+z} + \varepsilon_{x,t+z} \tag{10}$$

Measures

As Baker (1992) pointed out that traditional error measures, such as mean square error, do not provide a reliable basis for comparison of methods. The Mean Absolute Percentage Error (MAPE) is more appropriate because it is invariant to scale and is not overly influenced by outliers. For comparisons using a small set of series, it is desirable, also, to control for degree of difficulty in forecasting. For these reasons, as a measure of fitting and forecast accuracy of each model and compared with separate forecasts by branch, we compute the mean absolute error (MAPE) from

$$MAPE = \frac{\sum ABS(\%error)}{n} \tag{11}$$

Analyses

In the analysis, we simply take the quantitative measures of the forecasting results. We seek to the suitable statistical model by comparing the forecasting results with the actual value during years 2005-2006 and the results from using separate forecast method, including to measure the forecasting errors of each model.

RESEARCH RESULTS

In order to find the best model for long-term sales forecasting, we fitted the Lee-Carter 1-3 components compare with separate forecast including to the actual data during years 2005 -2006 (24 months). The forecasting results of some branches can be compared in Figure 8. As can be seen from Figure 8 that the predicted lines of each model is quite closed to the actual values. However, the forecasting results using the Lee-Carter model with 3 components extension can give the best fit than other models and there are some forecasting errors that need to be considered in the next step.

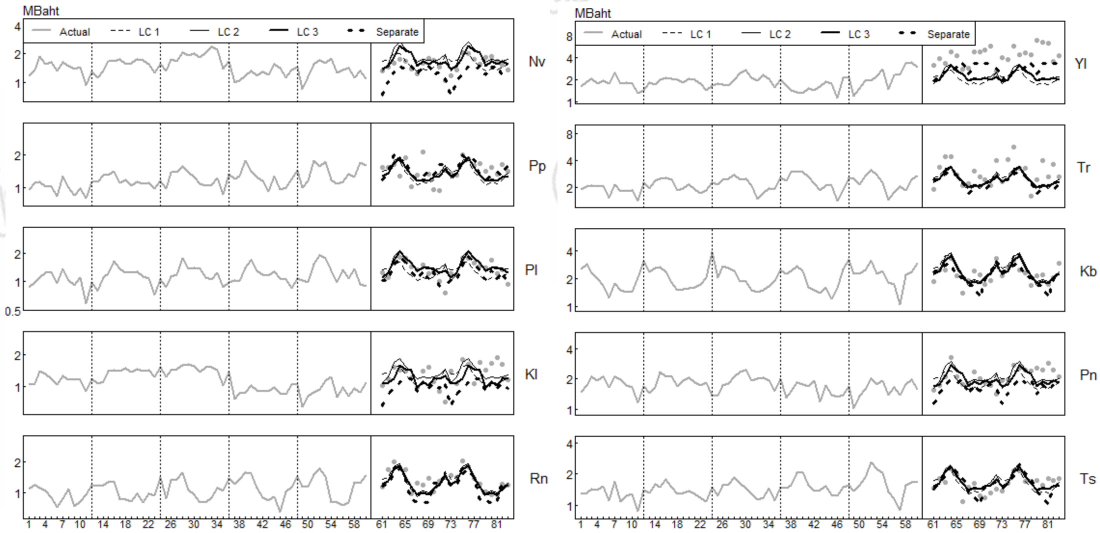

Figure 8: Forecasting results using Lee-Carter models 1-3 components compare with the results from using separate forecast method

Forecasting errors

As can be seen from Figure 8 that there is a rapid trend in Yala branch during the years 2005-2006 that may caused high forecasting errors. So, we evaluated the forecast accuracy by comparison both the total forecasting error and the errors excluding the Yala (Yl) branch from the formulas given by (11). The forecasting errors are shown in Table 1. The table shows that the mean absolute percent error from using the Lee-Carter model with 3 components extension is less than using the other model or the separate method. The errors from forecasting sales in 19 branches (exclude Yala branch) are less than from forecasting sales in all branch locations.

Table 1: Forecasting errors comparison between each model

Methods	Mean absolute percent error (MAPE)	
	Total	Excl. Yala branch
Lee-Carter 1 component	1.52%	1.35%
Lee-Carter 2 components	1.35%	1.20%
Lee-Carter 3 components	1.32%	1.16%
Separate forecast	1.47%	1.39%

Figure 9 shows the distribution of total forecasting errors in each model in the box plot compares with the distribution of forecasting errors excluding Yala branch. It is clear that the forecasting results using Lee-Carter model with 3 components extension is a suitable model since the median value of error (as can be noticed from line in the box) is very small and the model can give the less far out of errors than other models.

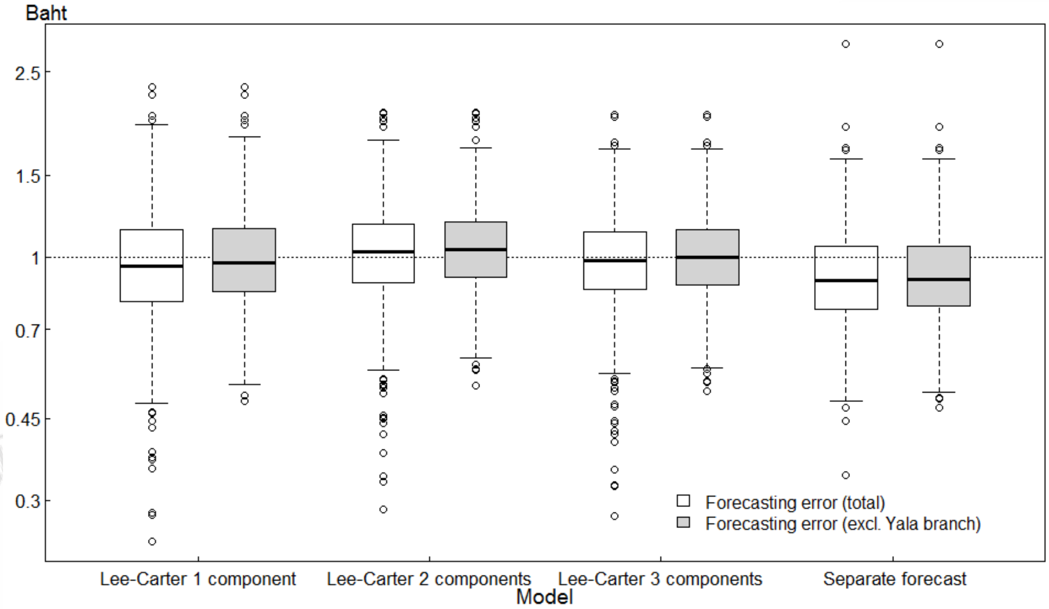

Figure 9: Forecasting errors of each model in total compare with the errors exclude Yala branch

From the forecasting errors analysis as shown in Table 1 and Figure 9, we found that the Lee-Carter model with 3 components extension was the best model for long-term forecasting of sparkling beverages sales in Southern Thailand.

Prediction Interval

It is useful to consider the 95% prediction intervals with the point forecast. Figure 10 shows some of the forecasting results with 95% prediction intervals comparison between Lee-Carter model with 3 components extension and separate forecasts. As can be seen from Figure 10 that although there are some of the observed data are outside of the predicted lines but they are well-covered by the prediction intervals. The plot displays that the forecasting results using the Lee-Carter model with 3 components extension are well-covered by the prediction intervals than the separate method.

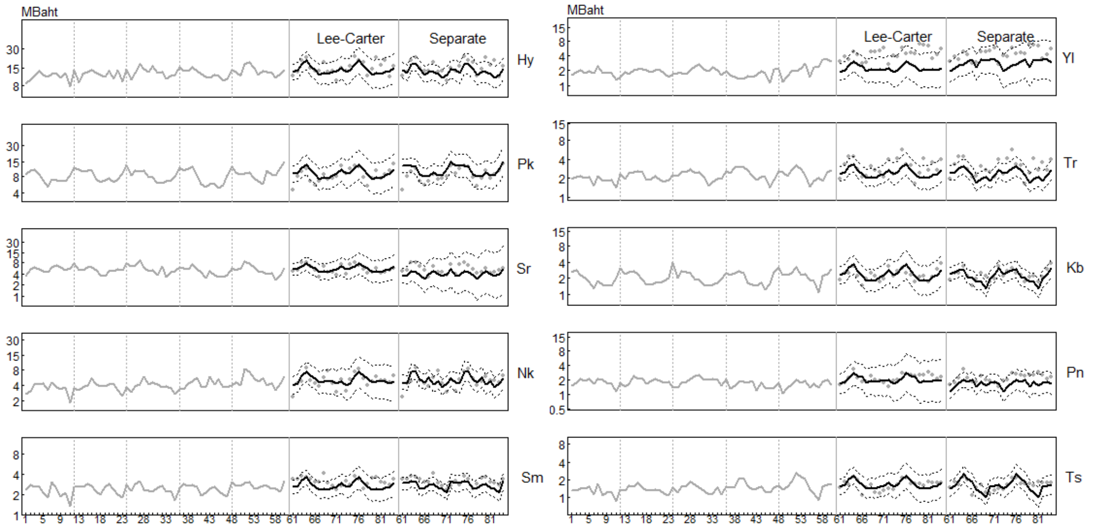

Figure 10: The forecast results using Lee-Carter 3 components with 95% prediction interval compare with the results from using separate forecast and actual data

Figure 11 shows the scatter plot comparison between actual sales in the forecasting periods (January 2005 to December 2006) and 24 months predicted sales in 20 branches. The plot indicates that the forecasting results are well fit to the actual value and the smoothing line.

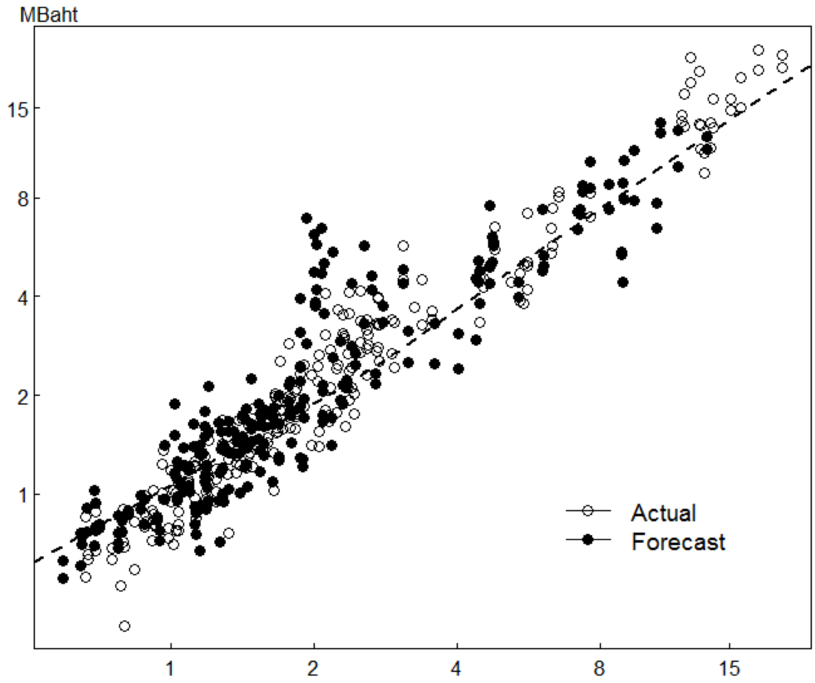

Figure 11: Actual versus forecasted sales during January 2005 - December 2006

Figure 12 shows time series of the actual sales during years 2000 - 2004 (60 months) and the sales forecast of some branches from fitting Lee-Carter model with 3 components extension during years 2005 – 2006 (24 months ahead). The plot reveals that the forecasting sales is appropriate fitted to the last period trends and can give the reasonable trend in the future months.

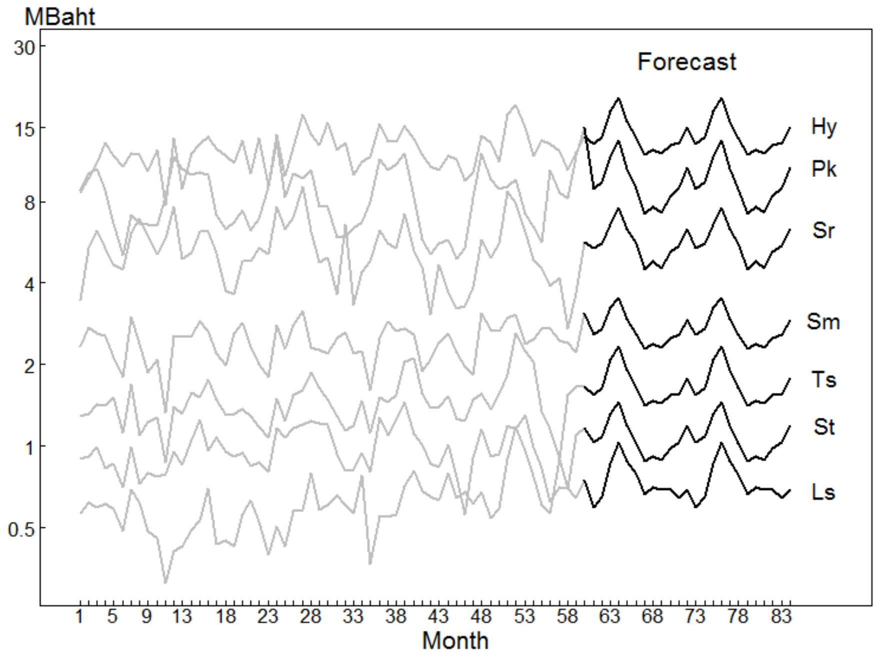

Figure 12: Monthly actual sales by branch with 24 future months sales forecasting using the Lee-Carter model with 3 components extension

DISCUSSION

The sales in each branch have increased substantially. This could be due to an expansion in modern trade outlets affecting purchasing behaviours and consumers life style in Southern Thailand. There is a significant sales growth in Yala branch during the forecasting periods due to the market execution and distribution penetration by the company. So, it is a good opportunity for the company to focus and do some more promotions or marketing strategies especially for this area.

The seasonal effects found in our study could be related both to regional climatic changes and to human activities. In the dry season (extending from February to April), hot weather and long holidays lead to greater consumption of sparkling beverages.

Degree of answering the research objectives and the support of the initial assumptions

Forecasting is an active research area for many decades. The accuracy of forecasting hence researchers for improving the effectiveness of forecasting models. In this paper, we applied a Lee-Carter model with 3 components extension and exponential smoothing Holt-Winters with additive seasonality method for sparkling beverages sales data with logarithmic transformation. This is a suitable case-study since the model produced excellent estimates in sales predicting for up to 24 future months of 20 branches compared with actual data in each branch during the years 2005 - 2006. As can be seen from Table 1 that the model gave more accurate results than using separate forecasting method. The Lee-Carter model takes advantage of the unique strength of non-linear modeling and it produces an excellent fit and sensible estimates in long-term sales predicting for many branches in the same time. It is also parsimonious in the number of parameters used and its concept is not much complicated for a person with minimal knowledge of statistics. The exponential smoothing Holt-Winters with additive seasonality method allows us to deal with univariate time series with contain both trend and seasonally factors. It contains simple model formulation and can give good forecasting results.

The model can support of the initial assumptions that the future sales trend will follow to historical sales or have a similar pattern. The model can also support our research aims in investigating suitable statistical methods for long-term sales forecasting and implementing the method using free available software (R).

Relating the findings to earlier work

In the previous researches or earlier works, the long-term sales forecasts usually work by using the separate forecasting methods. For example, in this case forecasters need to forecast the sales in each branch location separately or do the 20 times forecasts. This study attempted to extend the knowledge about long-term forecasting through the development of a statistical model. By using such our model, the forecasters can save the forecasting time and cost since it is needs only one time fitting the forecasting model given by (10), to achieve the sales forecast for up to 24 future months of 20 branches**.**

Theoretical implications

The models can be extended into other specific business data analysis and forecasting using the same methods and theoretical framework.

Practical implications

For practical use, the model is designed to deal with the following two points: accuracy and timing used. As a case study, we applied the model to sparkling beverages sold in Southern Thailand. The model enables us to obtain a practical sales forecast in 20 branches, and furthermore provides valuable information on reprint decision-making. Using such models for forecasting sales revenue can assist company managers in long-run planning more effectively.

Limitations

There is a limitation of the Lee-Carter model that the model is well-fitted in case of the sales trend in each branch has a similar pattern. As can be noticed from the Yala branch case, the forecasting results was not good enough due to there was a rapid sales growth during the predicting periods which was a different trend compared with other branches. Some substantial interactions between the branch-month factor and the other factors existed could not be fitted using the model. In addition, the forecasting results using Holt-Winters method always depend on the last period trend.

Recommendation for further researches

For further studies, sales data in recent year, new packages and new flavours data may be considered in further sales forecast. The models can be easily adapted for per capita consumption forecasting or extended to other business data forecast.

CONCLUSION

In this study, we applied a classical Lee-Carter mortality forecasting approach as well as exponential smoothing Holt-Winters with additive seasonality method to log-transformed monthly sales containing season of month and branch location as factors**.** By using a basic forecasting process, we found that the Lee-Carter model with Holt-Winters method can produces an excellent fit and it can give sensible estimates in long-term sales forecasting for 20 branches in the same time. The model also parsimonious in the number of parameters used and its concept is not much complicated for a person with minimal knowledge of statistics. The model also gives more accurate results than using separate forecasting method. The models can be extended into other specific business data analysis and forecasting using the same methods.

ACKNOWLEDGEMENTS

We would like to thank the Office of the Higher Education Commission, Thailand for supporting by grant fund under the program Strategic Scholarships Fellowships Frontier Research Networks
(Specific for Southern region) for the Join PhD Program Thai Doctoral degree for this research. We are grateful for Prof. Don McNeil for his helpful advice and suggestions. We also would like to thank Khun Dumrongrugs Apibalsawasdi for his guidance.

AUTHORS INFORMATION

Wassana Suwanvijit: Instructor for the Faculty of Economics and Business Administration, Thaksin University, Songkhla, Thailand and currently a PhD candidate in Research Methodology.

Thomas Lumley, **PhD**: Professor for the Department of Biostatistics, University of Washington, USA.

Chamnein Choonpradub, PhD: Assistant Professor for the Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University (Pattani campus), Thailand.

Nittaya McNeil, **PhD**: Instructor for the Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University (Pattani campus), Thailand.

REFERENCES

1. Baker, Michael J. The IEBM Encyclopedia of Marketing, International Thompson Business Press, pp. 278-290, 1999.
2. Bell, W.R. Comparing and assessing time series methods for forecasting age-specific fertility rates, *Journal of Official Statistic*, Vol. 13, pp. 279-303, 1997.
3. Bermudez, J.D., Segura, J.V. and Vercher, E. Holt-Winters forecasting: an alternative formulation applied to UK air passenger data, *Centro de Investigation Operativa (CIO)*, 2005.
4. Bohr, Neils. and Danish. Introduction to Forecasting, 2007. http://www.answers.com/forecasting?cat=biz-fin [March 12, 2009]
5. Boonruangthaworn, Duangrat. Forecasting Techniques for New Product: Case Study of Soft Drink Industry, 2007. http://aitm.agro.ku.ac.th/students/s01/abstracts/2007_duangrat_en.pdf [January 22, 2008]
6. Booth, H. Demographic Forecasting: 1980 to 2005 in review, *International Journal of Forecasting*, Vol. 22, pp. 547-581, 2006.
7. Booth, H., Hyndman R.J., Tickle, L. and DeJong, P. Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions, *Demographic Research*, Vol. 15, pp. 289-310, 2006.
8. Booth, H., Maindonald, J. and Smith, S. Applying Lee-Carter under conditions of variable mortality decline, *Population Studies*, Vol. 56, pp. 325-336, 2002.
9. Brouhns, N., Denuit, M. and Vermunt, J. A Poisson Log-bilinear Regression Approach to the Construction of Projected Lifetables, *Insurance: Mathematics and Economics*, Vol. 31, pp. 373-393, 2002.
10. Bureau. Indian Food and Beverages Forecast (2007-2011), June, 2007.
11. Caruana, Albert. Step in forecasting with seasonal regression: A case study from the carbonated soft drink market, Vol. 10, No. 2, pp. 94-102, 2001.
12. Chatfield, C., Koehler, A. B., Ord, J. K. and Snyder, R. D. A new look at models for exponential smoothing, *Journal of the Royal Statistical Society*, *Series D: The Statistician*, Vol. 50, No. 2, pp. 147-159, 2001.

13. Currie, Iain. The Extended Lee-Carter Family, Actuarial Teaching and Research Conference, Queen's University, 2009.
14. Currie, I.D., Durban, M. and Eilers, P.H.C. Smoothing and forecasting mortality rates. *Statistical Modelling*, Vol. 4, No. 4, pp. 279-298, 2004.
15. Dekimpe, M. G. and Hanssens, D. M. Time-series models in marketing: Past, present and future. *International Journal of Research in Marketing*, Vol. 17, pp. 183-193, 2000.
16. Denuit, Michel., Devolder, Pierre. and Goderniaux, Anne-C6cile. Securitization of Longevity Risk: Pricing Survivor with Wang Transform in the Lee-Carter Framework, *The Journal of Risk and Insurance*, Vol. 74, No. 1, pp. 87-113, 2007.
17. Franses, Philip Hans. and Dijk, Dick Van. The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production, *International Journal of Forecasting,* Vol. 21, pp. 87-102, 2005.
18. Gardner, E.S. Jr. Exponential smoothing: the state of the art, *Journal of Forecasting*, Vol. 4, pp. 1-28, 1985.
19. Gelper, S., Fried, R. and Croux, C. Robust forecasting with exponential and Holt–Winters smoothing, *Journal of Forecasting*. Vol. 29, No. 3, pp 285-300, 2010.
20. Girosi, Federico. and King, Gary. Understanding the Lee-Carter Mortality Forecasting Method, Working Paper, Harvard University, 2007.
21. Granger, Clive W.J., Jeon, Yongil. Long-term forecasting and evaluation, *International Journal of Forecasting,* Vol. 23, pp. 539–551, 2007.
22. Girosi, Federico. and King, Gary. Understanding the Lee-Carter Mortality Forecasting Method, Working Paper, Harvard University, 2007.
23. Hanke, John E. and Wichern, Dean W. Business Forecasting, Pearson Education, New Jersy, 2005.
24. Higgins, Richard S., Kaplan, David R. and McDonald, Michel J. Residual Demand Analysis of the Carbonated Soft Drink Industry, 2005. http://www.springerlink.com/content/h7n6w33821up1w24/ [July 13, 2009]
25. Holt, C.C. Forecasting seasonals and trends by exponentially weighted moving averages, *ONR Research Memorandum*, Carnigie Institute, Vol. 52, 1957.
26. Hyndman, Rob J. and Ullah Md. Shahid. Robust forecasting of mortality and fertility rates: A functional data approach, *Computational Statistics & Data Analysis*, Vol. 51, pp. 4942-4956, 2007.
27. Hyndman, Rob J. Business forecasting methods, 2009. http://www. robjhyndman.com/papers/businessforecasting.pdf [January 22, 2010]
28. Inoue, Atsushi. and Kilian, Lutz. On the selection of forecasting models, *Journal of Econometrics,* Vol. 130, pp. 273–306, 2006.
29. Jan G. De Gooijer., Rob J. Hyndman. 25 years of time series forecasting, *International Journal of Forecasting*, Vol. 22, pp. 443-473, 2006.
30. Kalekar, Prajakta S. Time series forecasting using Holt-Winter exponential smoothing, Kanwall Rekhi School of Information Technology, 2004.
31. Koehler, Anne B., Snyder, Ralph D. and Ord, J. Keith. Forecasting models and prediction intervals for the multiplicative Holt-Winters methods. *International Journal of Forecasting*, Vol. 17, pp. 269-286, 2001.
32. Koissi, M-C., Shapiro, A.F. and Hognas, G., Evaluating and extending the Lee-Carter models for mortality forecasting: Bootstrap confidence intervals, *Insurance Mathematics and Economics*, Vol. 38, pp. 1-20, 2006.
33. Kotsialos, Apostlos., Papageorgiou, Markos. and Poulimenos, Antonios. Long-Term Sales Forecasting Using Holt-Winters and Neural Network Methods, *Journal of Forecasting*, Vol. 24, pp. 353-368, 2005.
34. Lawton, Richard. How should additive Holt-Winters estimates be corrected?, *International Journal of Forecasting*, Vol. 14, pp. 393-403, 1998.
35. Lazar, Dorina. On forecasting mortality using Lee-Carter method, Babes Bolyai University, 2004. http://www.stat.ucl.ac.be/Samos2004/proceedings2004/Lazar2.pdf [April 14, 2010].
36. Lee, R.D. and Carter, L. R., Modeling and forecasting U.S. Mortality, *Journal of American Statistical Association*, Vol. 87, No. 41, pp. 659-671, 1992.

37. Lee, R.D. The Lee-Carter Method for Forecasting Mortality, With Various Extensions and Applications (with discussion), *North American Actuarial Journal*, Vol. 4, No. 1, pp. 80-93, 2000.
38. Lee, R.D. and Miller, T. Evaluating the performance of the Lee-Carter mortality forecasts, *Demography*. Vol. 38, pp. 537-549, 2001.
39. Lee, R. D., Nault, F. Modeling and forecasting provincial mortality in Canada, paper presented at the World Congress of the International Union for Scientific Study of Population, Montreal, 1993.
40. Lee, R. D., Rofman, R. Modeling and Projecting Mortality in Chile, *Notas Polacion*, Vol. 22, No. 59, pp. 183-213, 1994.
41. Lin, Chin-Tsai. and Hsu, Pi-Fang., Forecast of non-alcoholic beverage sales in Taiwan using the Grey theory. Asia Pacific Journal of Marketing and Logistics, Vol.14, No.4, pp.3-12, 2002.
42. McGartland, C., Robson, P.J., Murray, L., Cran, G., Savage, M.J., Watkins, D., Rooney, M., & Roreham, C. Carbonated Soft Drink Consumption and Bone Mineral Density in Adolescence: The Northern Ireland Young Hearts Project, *Journal of bone and mineral research*, Vol. 18, pp. 1563-1569, 2003.
43. Newbeme, Joan H. and Captain, USAF. MSC Holt-Winters Forecasting in Healthcare, Mike O' Callaghan Federal Hospital, 2007.
44. ONS. Documentation of UKCeMGA Methods used for National Accounts, 2008. http://www.statistics.gov.uk. [April 14, 2010].
45. Pan, R. Holt–Winters Exponential Smoothing, Wiley Encyclopedia of Operations Research and Management Science, 2010.
46. Racharit, N. and Suraseangsung, S. Forecasting models for Thai Mortality Rates Using the Lee-Carter Method, *Journal of Demography*, Vol. 22, No. 2, 2006.
47. Renshaw, A.E. and Haberman, S. Lee-Carter mortality forecasting: a parallel generalised linear modeling approach for England and Wales mortality projections, *Applied Statistics*, Vol. 52, pp. 119-137, 2003.
48. R Development Core Team. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2008.
49. RNCOS. South Korean Food, Beverages and Tobacco Market Forecast till 2011, 2007
50. Software World. Computer model for forecasting beer consumption in U.K. A.P. Publications Ltd. 2009. http://www.thefreelibrary.com [April 14, 2010].
51. Sorin, Sylvain. Exponential weights algolithm in continuous time, Spinger-Virlag, 2007.
52. Steffens, P., Kaya, M. and Albers, S. Long term sales forecasts of innovations - An empirical study of the consumer electronic market, proceedings AGSE Entrepreneurship Research Exchange, 2007.
53. Suwanvijit, W., Choonpradub, C. and McNeil, N. Sales analysis with application to sparkling beverage product sales in Southern Thailand, *International Journal of Business and Management*, Vol. 4, No. 7, pp. 43-51, 2009.
54. Suwanvijit, W., Choonpradub, C. and McNeil, N. Statistical model for short-term forecasting sparkling beverage sales in Southern Thailand, *International Business and Economics Research Journal*, Vol. 8, No. 9, pp. 73-81, 2009.
55. Tanaka, K. A sales forecasting model for new-released and nonlinear sales trend products, *Expert Systems with Applications*, Vol. 37, No. 11, pp. 7387-7393, 2010.
56. Tinney, Mary-Colleen., Retail Sales Analysis: Champagne and Sparkling Wine Sales Rising, Wine Communications Group, 2007.
57. Tuljapurkar, S., Li, N. and Boe, C. A Universal Pattern of Mortality Decline in the G7 Countries, *Nature*, Vol. 405, pp. 789-792, 2000.
58. Venables, W. N. and Ripley B.D. Modern Applied Statistics with S. Springer, Queensland, 2002.
59. Venables, W. N. and Smith, D.M. An Introduction to R, 2004. http://cran.r-project.org/doc/manuals/R-intro.pdf [March 30, 2009]
60. Wang, Jenny Zheng. Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model, *Mathematical Statistics*, Stockholm University, 2007. http://www.math.su.se/matstat [April 4, 2010].

61. Williams. Developing Sales Forecasts, 2007. http://web.csustan.edu/market/williams/Chapter%2010.htm [March 14, 2010].
62. Wilmoth, John R. Mortality Projections for Japan: A Comparison of Four Methods, *Health and Mortality Among Elderly Populations*, Oxford University Press, pp. 266-287, 1996.
63. Wilmoth, J.R. Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. *Technical Report. Dept. of Demography*, University of California, Berkeley,1993.
64. Winters, P. R. Forecasting sales by exponentially weighted moving averages, *Management Science*, Vol. 6, pp. 324–342, 1960.
65. Yelland, P., Kim, S. and Stratulate, R. A Bayesian model for sales forecasting at Sun Microsystems,
Interfaces, Vol. 40, No. 2, pp.118-129, 2010.