Robotics and Computer-Integrated Manufacturing 34 (2015) 151–163

Contents lists available at ScienceDirect

Robotics and Computer-Integrated Manufacturing

journal homepage: www.elsevier.com/locate/rcim

Performance of state space and ARIMA models for consumer retail sales forecasting

Patrícia Ramos [a,b,*], Nicolau Santos [b], Rui Rebelo [b]

[a] School of Accounting and Administration of Porto, Polytechnic Institute of Porto, 4465-004 S. Mamede de Infesta, Portugal
[b] INESC Technology and Science, Manufacturing Systems Engineering Unit, 4200-465 Porto, Portugal

ARTICLE INFO

Article history:
Received 26 March 2014
Received in revised form
10 December 2014
Accepted 29 December 2014
Available online 22 January 2015

Keywords:
Aggregate retail sales
Forecast accuracy
State space models
ARIMA models

ABSTRACT

Forecasting future sales is one of the most important issues that is beyond all strategic and planning decisions in effective operations of retail businesses. For profitable retail businesses, accurate demand forecasting is crucial in organizing and planning production, purchasing, transportation and labor force. Retail sales series belong to a special type of time series that typically contain trend and seasonal patterns, presenting challenges in developing effective forecasting models. This work compares the forecasting performance of state space models and ARIMA models. The forecasting performance is demonstrated through a case study of retail sales of five different categories of women footwear: Boots, Booties, Flats, Sandals and Shoes. On both methodologies the model with the minimum value of Akaike's Information Criteria for the in-sample period was selected from all admissible models for further evaluation in the out-of-sample. Both one-step and multiple-step forecasts were produced. The results show that when an automatic algorithm the overall out-of-sample forecasting performance of state space and ARIMA models evaluated via RMSE, MAE and MAPE is quite similar on both one-step and multi-step forecasts. We also conclude that state space and ARIMA produce coverage probabilities that are close to the nominal rates for both one-step and multi-step forecasts.

1. Introduction

Sales forecasting is one of the most important issues that is beyond all strategic and planning decisions in any retail business. The importance of accurate sales forecasts to efficient inventory management at both disaggregate and aggregate levels has long been recognized [1]. Poor forecasts usually lead to either too much or too little inventory directly affecting the profitability and the competitive position of the company. At the organizational level, sales forecasting is very important to any retail business as its outcome is used by many functions in the organization: finance and accounting departments are able to project costs, profit levels and capital needs; sales department is able to get a good knowledge of the sales volume of each product; purchasing department is able to plan short- and long-term purchases; marketing department is able to plan its actions and assess the impact of different marketing strategies on sales volume; and finally logistics department is able to define specific logistic needs [2]. Accurate forecasts of sales have the potential to increase the profitability of retailers by improving the chain operations efficiency and minimizing wastes. Moreover, accurate forecasts of retail sales may improve portfolio investors' ability to predict movements in the stock prices of retailing chains [3]. Aggregate retail sales time series are usually preferred because they contain both trend and seasonal patterns, providing a good testing ground for comparing forecasting methods, and because companies can benefit from more accurate forecasts.

Retail sales time series often exhibit strong trend and seasonal variations presenting challenges in developing effective forecasting models. How to effectively model retail sales series and how to improve the quality of forecasts are still outstanding questions. Exponential smoothing and Autoregressive Integrated Moving Average (ARIMA) models are the two most widely used approaches to time series forecasting, and provide complementary approaches to the problem. While exponential smoothing methods are based on a description of trend and seasonality in the data, ARIMA models aim to describe the autocorrelations in the data. The ARIMA framework to forecasting originally developed by Box et al. [4] involves an iterative three-stage process of model selection, parameter estimation and model checking. A statistical framework to exponential smoothing methods was recently developed based on state space models called ETS models [5].

Despite the investigator's efforts, the several existing studies

* Corresponding author at: INESC Technology and Science, Manufacturing Systems Engineering Unit, 4200-465 Porto, Portugal.
E-mail addresses: patricia.ramos@inescporto.pt (P. Ramos), nicolau.f.santos@inescporto.pt (N. Santos), rrebelo@inescporto.pt (R. Rebelo).

http://dx.doi.org/10.1016/j.rcim.2014.12.015
0736-5845/

have not led to a consensus about the relative forecasting performances of these two modeling frameworks when they are applied to retail sales data. Alon [6] concluded that the Winters exponential smoothing method' forecasts of aggregate retail sales were more accurate than the simple and Holt exponential smoothing methods' forecasts. Additionally, Alon et al. [3] compared out-of-sample forecasts of aggregated retail sales made using artificial neural networks (ANNs), Winters exponential smoothing, ARIMA and multiple regression via MAPE (mean absolute percentage error). Their results indicate that Winters exponential smoothing and ARIMA perform well when macroeconomic conditions are relatively stable. When economic conditions are volatile (supply push inflation, recessions, high interest rates and high unemployment) ANNs outperform the linear methods and multi-step forecasts may be preferred. Chu and Zhang [7] also conducted a comparative study of linear and nonlinear models for aggregate retail sales forecasting. The linear models studied were the ARIMA model, regression with dummy variables and regression with trigonometric variables. The nonlinear models studied were the ANNs for which the effect of seasonal adjustment and use of dummy or trigonometric variables was investigated. Using multiple cross-validation samples evaluated via the RMSE (root mean squared error), the MAE (mean absolute error) and MAPE, the authors concluded that no single forecasting model is the best for all situations under all circumstances. Their empirical results show that (1) prior seasonal adjustment of the data can significantly improve forecasting performance of the neural network model; (2) seasonal dummy variables can be useful in developing effective regression models (linear and nonlinear) but the performance of these dummy regression models may not always be robust; (3) trigonometric models are not useful in aggregate retail sales forecasting. Another interesting example is by Frank et al. [8] forecast women's apparel sales using single seasonal exponential smoothing (SSES), the Winters' three parameter model and ANNs. The performance of the models was tested by comparing the goodness-of-fit statistics R^2 and by comparing actual sales with the forecasted sales of different types of garments. Their results indicated that the three parameter Winters' model outperformed SSES in terms of R^2 and forecasting sales. ANN model performed best in terms of R^2 (among three models) but correlations between actual and forecasted sales were not satisfactory. Zhang and Qi [9] and Kuvulmaz et al. [10] further investigated the use of ANNs in forecasting time series with strong trend and seasonality and conclude that the overall out-of-sample forecasting performance of ANNs, evaluated via RMSE, MAE and MAPE, is not better than ARIMA models in predicting retail sales without appropriate data preprocessing namely detrending and deseasonalization. Motivated by the particular advantages of ARIMA models and ANNs, Aburto and Weber [11] developed a hybrid intelligent system combining ARIMA type approaches and MLP-type neural networks for demand forecasting that showed improvements in forecasting accuracy. Encouraged by their results they proposed a replenishment system for a Chilean supermarket which led simultaneously to fewer sales failures and lower inventory levels. Motivated by the recent success of evolutionary computation Au et al. [12] studied the use of evolutionary neural networks (ENNs) for sales forecasting in fashion retailing. Their experiments show that when guided with the BIC (Bayesian Information Criterion) and the pre-search approach, the ENN can converge much faster and be more accurate in forecasting than the fully connected neural network. The authors also conclude that the performance of these algorithms is better than the performance of the ARIMA model only for products with features of low demand uncertainty and weak seasonal trends. Further, it is emphasized that the ENN approach for forecasting is a highly automatic one while the ARIMA modeling involves more human knowledge. Wong and Guo [13] propose a hybrid intelligent model using extreme learning machine (ELM) and a harmony search algorithm to forecast medium-term sales in fashion retail supply chains. The authors show that the proposed model exhibits superior out-of-sample forecasting performance over the ARIMA, ENN and ELM models when evaluated via RMSE, MAPE and MASE (mean absolute scaled error). However, they also observe that the performance of the proposed model deteriorated when the time series was irregular and random pointing that it may not work well with high irregularity and nonlinearity. Finally, Pan et al. [14] investigate the feasibility and potential of applying empirical mode decomposition (EMD) in forecasting aggregate retail sales. The hybrid forecasting method of integrating EMD and neural network models (EMD-NN) was compared with the direct NN model and the ARIMA model for aggregate retail sales forecasting. Data from two sampling periods with different macroeconomic conditions were studied. The out-of-sample forecasting results indicate that the performance of the hybrid NN model is more stable compared to direct NN model and ARIMA during volatile economy. However, during relatively stable economic activity, ARIMA performs consistently well. In summary, over the last few decades several methods such as Winters exponential smoothing, ARIMA model, multiple regression and ANNs have been proposed and widely used because of their ability to model trend and seasonal fluctuations present in aggregate retail sales. However, all these methods have shown difficulties and limitations being necessary to investigate further on how to improve the quality of forecasts. The purpose of this work is to compare the forecasting performance of state space models and ARIMA models when applied to a case study of retail sales of five different categories of women footwear from the Portuguese retailer Foreva. As far as we know it is the first time ETS models are tested for retail sales forecasting.

The remainder of the paper is organized as follows. The next section describes the datasets used in the study. Section 3 discusses the methodology used in the time series modeling and forecasting. The empirical results obtained in the research study are presented in Section 4. The last section offers the concluding remarks.

2. Data

The brand Foreva was born in September 1984. Since the beginning is characterized by offering a wide range of footwear for all seasons, the geographical coverage of Foreva shops in Portugal is presently vast; it has around 70 stores opened to the public most of them in Shopping Centers. In this study we analyze the monthly sales of the five categories of women footwear of the brand Foreva, Boots, Booties, Flats, Sandals and Shoes, from January 2007 to April 2012 (64 observations). These time series are plotted in Fig. 1. The Boots and Booties categories are sold primarily during the winter season while the Flats and Sandals categories are sold primarily during the summer season; the Shoes category is sold throughout the year. The winter season starts on September 30 one year and ends on February 27 next year. The summer season starts on February 28 and ends on September 29 of each year. With the exception to Flats series all the other series present a strong seasonal pattern and are obviously non-stationary. The Boots series remains almost constant in the first two seasons, decreases slightly in 2009–2010 recovering in 2010–2011 and then decreases again in 2011–2012. The Booties series also remains fairly constant in the first two seasons and then maintains an upward trend movement in the next three seasons. The Flats series seems more volatile than the other series and the seasonal fluctuations are not so visible. In 2007 the sales are clearly higher than the rest of the years. An exceptional increase of sales is observed in March and