Supply chain management of consumer goods based on linear forecasting models

Patrícia Ramos (patricia.ramos@inescporto.pt)
INESC TEC, ISCAP, Instituto Politécnico do Porto
Rua Dr. Roberto Frias, 378 4200-465, Porto, Portugal

José Manuel Oliveira (jmo@inescporto.pt)
INESC TEC, Faculdade de Economia, Universidade do Porto
Rua Dr. Roberto Frias, 378 4200-465, Porto, Portugal

Abstract

In this work we apply linear forecasting models to a very broad collection of retail sales of consumer goods from a Portuguese retailer. This allows us to draw conclusions for guidelines within this field, and also to contribute to general observations relevant to the main field of forecasting. For each retail series the model with the minimum value of the AIC for the in-sample period is selected from all admissible models for further evaluation in the out-of-sample. Both one-step and multiple-step forecasts are produced. The results show that ARIMA models outperform state space models in out-of-sample forecasting judged by MAPE.

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

Introduction

Demand forecasting is one of the most important issues that is beyond all strategic and planning decisions in any business organization. The importance of accurate demand forecasts in successful supply chain operations and coordination has been recognized by many researchers (Wong and Guo, 2010; Arlot and Alain, 2010). A poor forecast would result in either too much or too little inventory, directly affecting the profitability of the supply chain and the competitive position of the organization. Forecasting future sales is crucial to the planning and operation of retail business at both high and low levels. At the organizational level, forecasts of sales are needed as the essential inputs to many decision activities in various functional areas such as marketing, sales, production/purchasing, as well as finance and accounting (Agrawal and Schorling, 1996; Chopra and Meindl, 2007).

Retail sales often exhibit strong trend and seasonal variations, presenting challenges in developing effective forecasting models. Historically, modeling and forecasting seasonal data is one of the major research efforts and many theoretical and heuristic methods have been developed in the last several decades (Alon et al., 2001; Chu and Zhang, 2003; Zhang and Qi, 2005; Kuvulmaz et al., 2005, Pan et al., 2013). Exponential smoothing and Autoregressive Integrated Moving Average (ARIMA) models are the 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. (1994) 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 innovations state space models called ETS models (Hyndman et al., 2008a).

Despite the investigator's efforts, the several existing studies have not led to a consensus about the relative forecasting performances of these two modeling frameworks when they are applied to retail sales data. The purpose of this work is to compare the forecasting performance of state space models and ARIMA models when applied to a very broad collection of retail sales of four different categories of consumer goods from the Portuguese retailer Jerónimo Martins. As far as we known it's 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.

Data

Jerónimo Martins is a Portugal-based international group operating in food distribution, food manufacturing and services sectors. Involving operations in retail and wholesale formats, the Jerónimo Martins Group is the leader in food distribution in Portugal, with the brands Pingo Doce (leader in supermarkets) and Recheio (leader in cash & carry), in food store chains in Poland (Biedronka) and in Colombia (Ara).

The work presented in this paper was developed using 67 time series of sales of consumer goods of a Pingo Doce supermarket of around 1500 m^2 between January 2007 and July 2012 (67 months). Figure 1 shows the time plot of the number of different products sold per day in this supermarket during that period. It can be seen that the number of different products sold per day is increasing and it has an annual seasonal behavior.

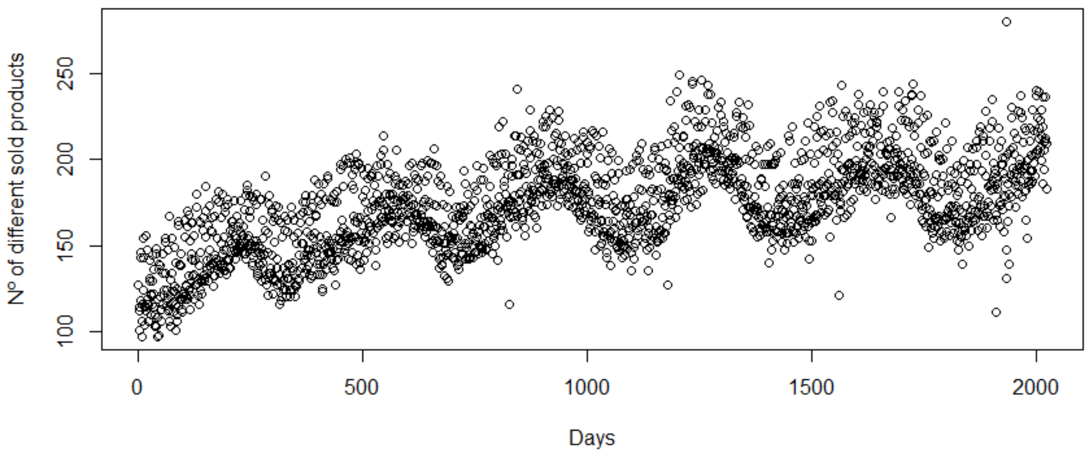

Figure 1 – Number of different products sold per day between January 2007 and July 2012.

To illustrate the broad collection of time series analyzed in this work, Figure 2 shows the time plot of monthly sales of six products sold between January 2007 and July 2012 (67 observations). All these series are obviously non-stationary exhibiting strong trend and/or seasonal patterns providing a good testing ground for comparing the two forecasting methods.

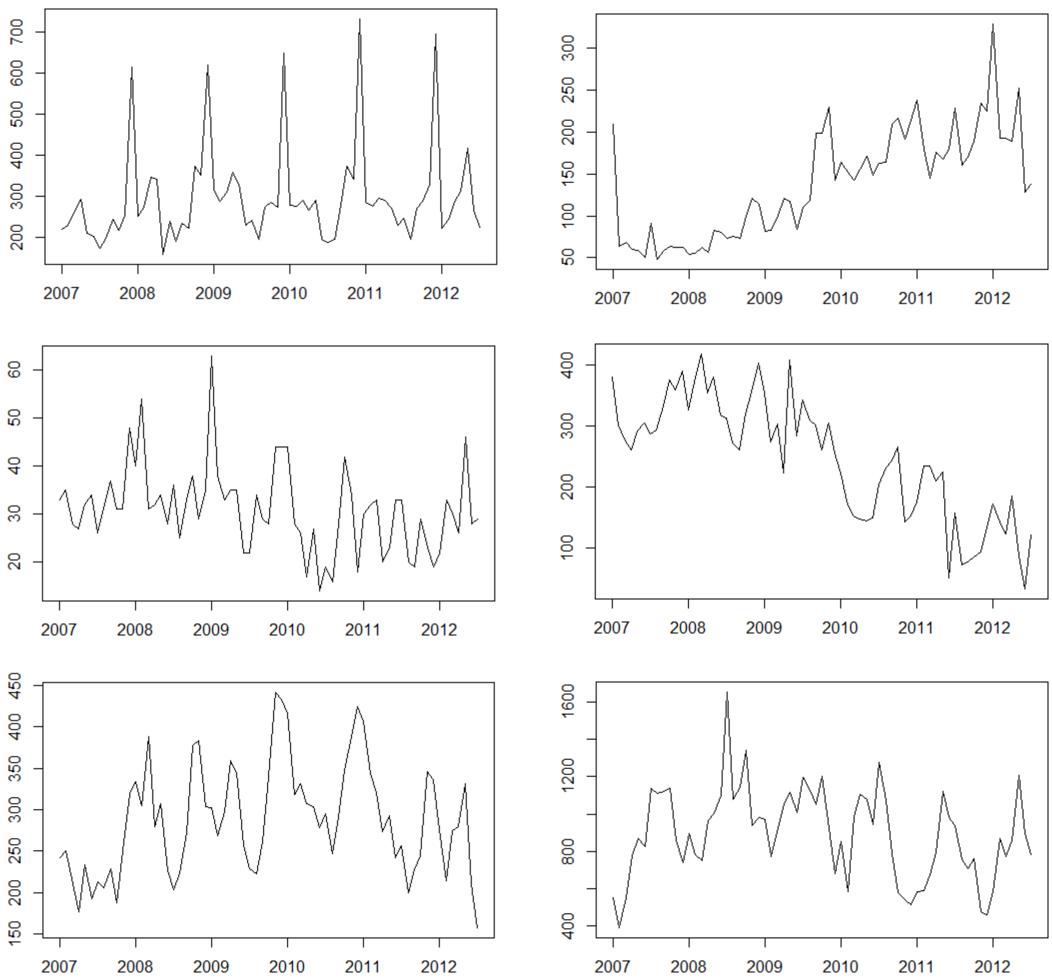

Figure 2 – Time series of consumer goods sold by Pingo Doce..

Forecasting models

ETS models

Exponential smoothing methods have been used with success to generate easily reliable forecasts for a wide range of time series since the 1950s (Gardner, 1985; Gardner, 2006). In these methods forecasts are calculated using weighted averages where the weights decrease exponentially as observations come from further in the past. The most common representation of these methods is the component form. Component form representations of exponential smoothing methods comprise a forecast equation and a smoothing equation for each of the components included in the method. The components that may be included are the level component, the trend component and the seasonal component. By considering all the combinations of the trend and seasonal components, fifteen exponential smoothing methods are possible. Each method is usually labeled by a pair of letters (T,S) defining the type of "Trend" and "Seasonal" components. The possibilities for each component are: $\text{Trend}=\{\text{N},\text{A},\text{A}_\text{d},\text{M},\text{M}_\text{d}\}$ and $\text{Seasonal}=\{\text{N},\text{A},\text{M}\}$. For illustration, denoting the time series by $y_1, y_2, \ldots, y_n$ and the forecast of y_{t+h}, based on all of the data up to time t, by $\hat{y}_{t+h|t}$, the component form for the method (A,A) (additive Holt-Winters' method) is (Hyndman and Athanasopoulos, 2013):

$$\hat{y}_{t+h|t} = l_t + hb_t + s_{t-m+h_m^+} \tag{1}$$

$$l_t = \alpha\left(y_t - s_{t-m}\right) + \left(1-\alpha\right)\left(l_{t-1} + b_{t-1}\right) \tag{2}$$

$$b_t = \beta^*\left(l_t - l_{t-1}\right) + \left(1-\beta^*\right)b_{t-1} \tag{3}$$

$$s_t = \gamma\left(y_t - l_{t-1} - b_{t-1}\right) + \left(1-\gamma\right)s_{t-m}, \tag{4}$$

where m denotes the period of the seasonality, l_t denotes an estimate of the level of the series at time t, b_t denotes an estimate of the trend of the series at time t, s_t denotes an estimate of the seasonality of the series at time t and $\hat{y}_{t+h|t}$ denotes the point forecast for h periods ahead where $h_m^+ = \lfloor (h-1) \bmod m \rfloor + 1$. The initial sates $l_0, b_0, s_{1-m}, \ldots s_0$ and the smoothing parameters α, β^*, γ are estimated from the observed data. The smoothing parameters α, β^*, γ are constrained between 0 and 1 so that the equations can be interpreted as weighted averages. Details about all the other methods may be found in (Hyndman and Athanasopoulos, 2013).

To be able to generate forecast intervals and other properties, Hyndman et al. (2008a) (amongst others) developed a statistical framework for all exponential smoothing methods. In this statistical framework each stochastic model, referred as an innovations state space model, consists of a measurement equation that describes the observed data, and state equations that describe how the unobserved components or states (level, trend, seasonal) change over time. For each exponential smoothing method, Hyndman et al. (2008a) described two possible innovations state space models, one corresponding to a model with additive random errors and other corresponding to a model with multiplicative random errors, giving a total of 30 potential models. To distinguish the models with additive and multiplicative errors, an extra letter E was added: the triplet of letters (E,T,S) refers to the three components: "Error", "Trend" and "Seasonality". The notation ETS(,,) helps in remembering the order in which the components are specified. For illustration, the equations of the model ETS(A,A,A) (additive Holt-Winters' method with additive errors) are (Hyndman and Athanasopoulos, 2013):

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t \tag{5}$$

$$l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t \tag{6}$$

$$b_t = b_{t-1} + \beta\varepsilon_t \tag{7}$$

$$s_t = s_{t-m} + \gamma\varepsilon_t \tag{8}$$

and the equations of the model ETS(M,A,A) (additive Holt-Winters' method with multiplicative errors) are:

$$y_t = (l_{t-1} + b_{t-1} + s_{t-m})(1 + \varepsilon_t) \tag{9}$$

$$l_t = l_{t-1} + b_{t-1} + \alpha(l_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \tag{10}$$

$$b_t = b_{t-1} + \beta(l_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \tag{11}$$

$$s_t = s_{t-m} + \gamma(l_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \tag{12}$$

where

$$\beta = \alpha\beta^*, \qquad 0<\alpha<1, \qquad 0<\beta<\alpha, \qquad 0<\gamma<1-\alpha \tag{13}$$

and ε_t is a zero mean Gaussian white noise process with variance σ^2. Equations (5) and (9) are the measurement equation and Equations (6)-(8) and (10)-(12) are the state equations. The term "innovations" comes from the fact that all equations in this type of specification use the same random error process ε_t. The measurement equation shows the relationship between the observations and the unobserved states. The transition equation shows the evolution of the state through time. It should be emphasized that these models generate optimal forecasts for all exponential smoothing methods and provide an easy way to obtain maximum likelihood estimates of the model parameters (for more details see Hyndman and Khandakar (2008b)).

ARIMA models

ARIMA is one of the most versatile linear models for forecasting seasonal and non-seasonal time series. It has enjoyed great success in both academic research and industrial applications during the last three decades. The class of ARIMA models is broad. It can represent many different types of stochastic seasonal and non-seasonal time series such as pure autoregressive (AR), pure moving average (MA) and mixed AR and MA processes (Brockwell and Davis, 1991). The theory of ARIMA models has been developed by many researchers and its wide application was due to the work by Box and Jenkins (1994) who developed a systematic and practical model building method.

The multiplicative seasonal ARIMA model, denoted as ARIMA $(p,d,q)\times(P,D,Q)_m$, has the following form (Wei, 2005):

$$\phi_p(B)\Phi_P(B^m)(1-B)^d(1-B^m)^D y_t = c+\theta_q(B)\Theta_Q(B^m)\varepsilon_t \tag{14}$$

where

$$\phi_p(B)=1-\phi_1 B-\cdots-\phi_p B^p \qquad \Phi_P(B^m)=1-\Phi_1 B^m-\cdots-\Phi_P B^{Pm}$$

$$\theta_q(B)=1+\theta_1 B+\cdots+\theta_q B^q \qquad \Theta_Q(B^m)=1+\Theta_1 B^m+\cdots+\Theta_Q B^{Qm}$$

and m is the seasonal frequency, B is the backward shift operator, d is the degree of regular differencing, and D is the degree of seasonal differencing, $\phi_p(B)$ and $\theta_q(B)$ are the regular autoregressive and moving average polynomials of orders p and q respectively, $\Phi_P(B^m)$ and $\Theta_Q(B^m)$ are the seasonal autoregressive and moving average polynomials of orders P and Q respectively, $c=\mu(1-\phi_1-\cdots-\phi_P)(1-\Phi_1-\cdots-\Phi_P)$ where μ is the mean of $(1-B)^d(1-B^m)^D y_t$ process and ε_t is a zero mean Gaussian white noise process with variance σ^2. The roots of the polynomials $\phi_p(B)$, $\Phi_P(B^m)$, $\theta_q(B)$ and $\Theta_Q(B^m)$ should lie outside a unit circle to ensure causality and invertibility (Shumway and Stoffer, 2011). For $d+D\geq 2$, $c=0$ is usually assumed because a quadratic or a higher order trend in the forecast function is particularly dangerous.

Empirical study

Estimation results

In order to use ETS models for forecasting the values of initial states and smoothing parameters need to be known. It is easy to compute the likelihood of ETS models and so maximum likelihood estimates are usually preferred. A great advantage of the ETS statistical framework is that information criteria can be used for model selection, namely the Akaike's Information Criteria (AIC). For ETS models, AIC is defined as (Hyndman and Athanasopoulos, 2013):

$$\text{AIC} = -2\log(L) + 2k \tag{15}$$

where L is the likelihood of the model and k is the total number of parameters and initial states that have been estimated. Some of the combinations of (Error, Trend, Seasonal) can lead to numerical difficulties. Specifically, the models that can cause such instabilities are: ETS(M,M,A), ETS(M,Md,A), ETS(A,N,M), ETS(A,A,M), ETS(A,Ad,M), ETS(A,M,N), ETS(A,M,A), ETS(A,M,M), ETS(A,Md,N), ETS(A,Md,A), and ETS(A,Md,M) (Hyndman and Athanasopoulos, 2013). Usually these particular combinations are not considered when selecting a model.

The time series analysis was carried using the statistical software R programming language and the specialized package forecast (Hyndman and Khandakar, 2008b; R Development Core Team, 2013). For each time series of monthly sales all admissible ETS models were applied using the in-sample data between January 2007 and July 2011 (first 55 observations). The initial states and the parameters were estimated by maximizing the likelihood of each model. The ETS model with the minimum value of the AIC was selected to produce forecasts and forecast intervals on the out-of-sample period (August 2011 to July 2012, last 12 observations).

The main task in ARIMA forecasting is selecting an appropriate model order, that is the values of p, q, P, Q, d and D (the seasonal period is 12, $m = 12$). We use the automatic model selection algorithm that was proposed by Hyndman and Khandakar (2008b). We start by choosing the values of d and D by applying unit-root tests. It is recommended that seasonal differencing be done first because sometimes the resulting series will be stationary and there will be no need for a further regular differencing. $D = 0$ or $D = 1$ depending on the OCSB test (Osborn et al., 1988). Once the value of D is selected, d is chosen by applying successive KPSS unit-root tests (Kwiatkowski et al., 1992). Once d and D are known, the orders of p, q, P and Q are selected via Akaike's Information Criteria:

$$\text{AIC} = -2\log(L) + 2(p + q + P + Q + k + 1) \tag{16}$$

where $k = 2$ if $c \neq 0$ and 1 otherwise (the other parameter being σ^2), and L is the maximized likelihood of the model fitted to the differenced data $(1-B)^d (1-B^m)^D y_t$. Rather than considering every possible combination of p, q, P and Q, the algorithm uses a stepwise search to traverse the model space:

(a) The best model (with smallest AIC) is selected from the following four:

- ARIMA $(2,d,2)\times(1,D,1)_{12}$
- ARIMA $(0,d,0)\times(0,D,0)_{12}$,
- ARIMA $(1,d,0)\times(1,D,0)_{12}$,
- ARIMA $(0,d,1)\times(0,D,1)_{12}$.

If $d+D \le 1$, these models are fitted with $c \ne 0$, otherwise $c=0$. This is called the "current" model.

(b) Thirteen variations on the current model are considered where:

- one of p,q,P and Q is allowed to vary from the current model by ± 1;
- p and q both vary from the current model by ± 1;
- P and Q both vary from the current model by ± 1;
- the constant c is included if the current model has $c=0$ and excluded otherwise.

Whenever a model with a lower AIC is found, it becomes the new "current" model and the procedure is repeated. This process finishes when we cannot find a model close to the current model with a lower AIC. For each time series of monthly sales the step-wise algorithm described above was applied using the in-sample data between January 2007 and July 2011 (first 55 observations) to find an appropriate ARIMA model. The parameters of the models are estimated by maximizing the likelihood. The selected model was used to produce forecasts and forecast intervals on the out-of-sample period (August 2011 to July 2012, last 12 observations).

Forecast evaluation results

For each retail series both selected models (ETS and ARIMA) were used to forecast on the out-of-sample period from August 2011 to July 2012 (12 observations). Both one-step and multiple-step forecasts were produced. Using each model fitted for the in-sample period, point forecasts of the next 12 months (one-step forecasts) and the forecast accuracy measures based on the errors obtained were computed. The values of the average MAPE (mean absolute percentage error) of one-step forecasts obtained are presented in Table 1. Supposing T is the total number of observations, N is the in-sample size and h is the step-ahead, multi-step forecasts were obtained using the following algorithm:

For $h=1$ *to* $T-N$

  For $i=1$ *to* $T-N-h+1$

    Select the observation at time $N+h+i-1$ *as out-of-sample*

    Use the observations until time $N+i-1$ *to estimate the model*

    Compute the h*-step error on the forecast for time* $N+h+i-1$

  Compute the forecast accuracy measures based on the errors obtained for step-ahead h

Compute the mean of the forecast accuracy measures

In our case study $T=67$ and $N=55$. It should be emphasized that in multi-step forecasts the model is estimated recursively in each step i using the observations until time $N+i-1$. Both one-step and multi-step forecasts are important in facilitating a short and long planning and decision making. They simulate the real-world forecasting environment in which data need to be projected for short and long periods (Alon et al., 2001). The values of the average MAPE of multi-step forecasts obtained are also presented in Table 1.

The results from Table 1 show that ARIMA models outperform state space models on both one-step and multi-step forecasts judged by MAPE. ARIMA consistently forecasts more accurately than ETS on one-step forecasts and on all steps-ahead of multi-step forecasts.

Improvements on one-step forecasts are of 51%. On multi-step forecasts improvements increase with the increasing of the step until step-ahead 6 and then decrease until step-ahead 12. The improvements are 0%, 12%, 27%, 34%, 37%, 40%, 36%, 31%, 29%, 20%, 17% and 13% respectively. In both models the value of MAPE tends to increase with the increasing of the step until step-ahead 6 and then tends to decrease until step-ahead 12. The results of our analysis also show that multi-step forecasts are more accurate than one-step forecasts (with exception to steps-ahead 4, 5 and 6 in the ARIMA model). This is not surprising since multi-step forecasts incorporate information that is more updated.

Producing estimates of uncertainty is an essential aspect of forecasting which is often ignored. We also evaluated the performance of both forecasting methodologies in producing forecast intervals that provide coverages which are close to the nominal rates. Table 2 and Table 3 show the mean percentage of times that the nominal 95% and 80% forecast intervals contain the true observations for both one-step and multiple-step forecasts, respectively. The results indicate that both ETS and ARIMA produce coverage probabilities that are very close to the nominal rates. ETS produces better coverage probabilities for both 80% and 90% forecast intervals on both one-step and multiple-step forecasts. It can also be observed that both methods slightly underestimate the coverage probabilities for the nominal 80% forecast intervals.

Conclusions

Accurate retail sales forecasting can have a great impact on effective management of retail operations. Retail sales time series often exhibit strong trend and seasonal variations presenting challenges in developing effective forecasting models. How to effectively model these series and how to improve the quality of forecasts are still outstanding questions. Despite the investigator's efforts, the several existing studies have not led to a consensus about the relative forecasting performances of ETS and ARIMA modeling frameworks when they are applied to retail sales data. The purpose of this work was to compare the forecasting performance of state space models and ARIMA models when applied to a very broad collection of retail sales of four different categories of consumer goods from the Portuguese retailer Jerónimo Martins.

For each time series of monthly sales all admissible ETS models were applied using the in-sample period. The ETS model with the minimum value of the AIC was selected to produce forecasts and forecast intervals on the out-of-sample period. The automatic model selection algorithm proposed by Hyndman and Khandakar (2008b) was used to select an appropriate ARIMA model for each time series of monthly sales. The model selected by the step-wise algorithm for each time series was then used to produce forecasts and forecast intervals on the out-of-sample period. Both one-step and multiple-step forecasts were produced. The results indicate that ARIMA models outperform state space models in out-of-sample forecasting judged by MAPE. On both modeling approaches multi-step forecasts are generally better than one-step forecasts which is not surprising because multi-step forecasts incorporate information that is more updated. The performance of both forecasting methodologies in producing forecast intervals that provide coverages which are close to the nominal rates was also evaluated. The results indicate that both ETS and ARIMA produce coverage probabilities that are very close to the nominal rates. However, ETS produces better coverage probabilities for both 80% and 90% forecast intervals on both one-step and multiple-step forecasts. We could also observe that both methods slightly underestimate the coverage probabilities for the nominal 80% forecast intervals.

Table 1 – MAPE (%) for out-of-sample period forecasts (January 2007 to July 2012).

Model	One-step forecasts	Step-ahead of multi-step forecasts											
		1	2	3	4	5	6	7	8	9	10	11	12
ARIMA	**64.83**	**44.31**	**53.20**	**60.96**	**67.74**	**67.42**	**68.26**	**59.38**	**50.65**	**49.29**	**41.78**	**44.35**	**45.76**
ETS	132.93	44.53	60.69	83.59	102.31	107.51	113.90	92.81	73.32	69.21	52.40	53.67	52.56

Table 2 – Forecast 80% interval coverage for out-of-sample period forecasts (January 2007 to July 2012).

Model	One-step forecasts	Step-ahead of multi-step forecasts											
		1	2	3	4	5	6	7	8	9	10	11	12
ARIMA	75	74	75	75	74	73	75	75	77	76	77	76	66
ETS	**80**	**75**	**78**	**78**	**78**	**80**	**80**	**82**	**84**	**82**	**81**	**86**	**78**

Table 3 – Forecast 95% interval coverage for out-of-sample period forecasts (January 2007 to July 2012).

Model	One-step forecasts	Step-ahead of multi-step forecasts											
		1	2	3	4	5	6	7	8	9	10	11	12
ARIMA	92	89	91	91	90	89	92	91	90	90	89	89	90
ETS	**94**	**90**	91	**90**	90	**91**	**93**	**94**	**93**	**93**	**92**	**95**	**97**

Acknowledgments

Project "NORTE-07-0124-FEDER-000057" is financed by the North Portugal Regional Operational Programme (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF), and by national funds, through the Portuguese funding agency, Fundação para a Ciência e a Tecnologia (FCT).This work was also co-supported by Jerónimo Martins SGPS, SA.

References

Agrawal, D. and Schorling, C. (1996), "Market share forecasting: An empirical comparison of artificial neural networks and multinomial logit model", *Journal of Retailing*, 72 (4), 383-407.

Alon, I., Min, Q. and Sadowski, R.J. (2001), "Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional method", *Journal of Retailing and Consumer Services*, 8 (3), 147-156.

Arlot, S. and Alain, C. (2010), "A survey of cross-validation procedures for model selection", *Statistics Surveys*, 4, 40-79.

Box G.E.P., Jenkins G.M. and Reinsel G.C. (1994), *Time series analysis: forecasting and control*, Prentice-Hall, Englewood Cliffs, New Jersey.

Brockwell, P.J. and Davis, R.A. (1991), *Time series: theory and methods*, Springer-Verlag, New York.

Chopra, S. and Meindl, P. (2007), *Supply Chain Management: Strategy, Planning, and Operation*, Prentice Hall, Upper Saddle River, New Jersey.

Chu, C.W. and Zhang, P.G.Q. (2003), "A comparative study of linear and nonlinear models for aggregate retail sales forecasting", *International Journal of Production Economics*, 86, 217-231.

Cryer, J.D. and Chan, K.S. (2009), *Time series analysis with applications in R*, Springer, New York.

Gardner, E.S. (1985), "Exponential smoothing: The state of the art", *Journal of Forecasting*, 4 (1), 1-28.

Gardner, E.S. (2006), "Exponential smoothing: The state of the art-Part II", *International Journal of Forecasting*, 22 (4), 637-666.

Hyndman, R.J., Koehler, A.B., Ord J.K. and Snyder R.D. (2008a), *Forecasting with Exponential Smoothing: The State Space Approach*, Springer-Verlag.

Hyndman, R.J. and Khandakar, Y. (2008b), "Automatic time series forecasting: The forecast package for R", *Journal of Statistical Software*, 26, 1-22.

Hyndman, R.J. and Athanasopoulos, G. (2013), *Forecasting: principles and practice*, http://otexts.com/fpp/.

Kwiatkowski, D., Phillips, P.C., Schmidt, P. and Shin, Y. (1992), "Testing the null hypothesis of stationariry against the alternative of a unit root", *Journal of Econometrics*, 54, 159-178.

Kuvulmaz, J., Usanmaz, S. and Engin, S.N. (2005), "Time-Series Forecasting by Means of Linear and Nonlinear Models", *Advances in Artificial Intelligence*, Springer Berlin, Heidelberg.

Ljung, G.M. and Box, G.E.P. (1978), "On a measure of lack of fit in time series models", *Biometrika*, 65, 297-303.

Osborn, D.R., Chui, A.P.L., Smith, J. and Birchenhall, C.R. (1988) "Seasonality and the order of integration for consumption", *Oxford Bulletin of Economics and Statistics*, 50 (4), 361-377.

Pan, Y., Pohlen, T. and Manago, S. (2013), "Hybrid Neural Network Model in Forecasting Aggregate U.S. Retail Sales", In Kenneth D. Lawrence and Ronald K. Klimberg (ed.), *Advances in Business and Management Forecasting*, 9, 153-170, Emerald Group Publishing Limited.

R Development Core Team (2013), R: A language and environment for statistical computing. http://www.R-project.org/ R version 3.0.0.

Shumway, R.H. and Stoffer, D.S. (2011), *Time series analysis and its applications: with R examples*, Springer, New York.

Wei, W.S. (2005), *Time series analysis: univariate and multivariate methods*, Addison Wesley.

Wong, W.K. and Guo, Z.X. (2010), "A hybrid intelligent model for medium-term sales forecasting in fashion retail supply chains using extreme learning machine and harmony search algorithm", *International Journal of Production Economics*, 128, 614-624.

Zhang, G. and Qi, M. (2005), "Neural network forecasting for seasonal and trend time series", *European Journal of Operational Research*, 160, 501-514.

Osborn-Chui-Smith-Birchenhall (1988).