Proceedings of the 2015 INFORMS Workshop on Data Mining and Analytics
M. G. Baydogan, S. Huang, A. Oztekin, eds.

EFFICIENTLY GENERATING THOUSANDS OF TIME-SERIES PRODUCT DEMAND FORECASTS USING R

Matthew A. Lanham
Virginia Polytechnic Institute and State University
Department of Business Information Technology (0235)
1007 Pamplin Hall, Blacksburg, VA 24061
lanham@vt.edu

Linda S. Schumacher
Raleigh, NC 27615
https://www.linkedin.com/pub/linda-schumacher/41/b21/9b1
schumachers@bellsouth.net

Abstract

This study identifies and tests a promising open-source framework for efficiently creating thousands of univariate time-series demand forecasts and reports interesting insights that could help improve other product demand forecasting initiatives. The investigation is motivated by the fact that retailers routinely require forecasts to be generated for planning purposes. Moreover, most retailers have expensive commercial inventory planning systems that automatically generate forecasts that a practitioner may use. While such systems exist, practitioners will often create their own forecasts, as well as, try and gain better insight about their products which can be a daunting task. Using a national retailer's data we implement this forecasting solution in R. We show that many of the forecasts are reliable, but also discuss issues we have found generating forecasts. We discuss our current efforts to identify and account for these issues and to improve the predictive performance of our models.

Keywords: Forecasting, Analytics, Product Planning

Introduction

Retailers routinely require forecasts to be generated for planning purposes. Accurate forecasts are important because additional uncertainty can have negative consequences to the retailer's assortment plan and inventory stocking levels, which in turn can impact other key performance indicators (e.g. reduced sales, increased costs, etc.). Most retailers have expensive commercial inventory planning systems that provide generated forecasts automatically that a practitioner may use, but the transparency to understand how they work and to foster better customization is severely lacking.

According to Hyndman and Athanasopoulos (2014), "*Forecasting, is about predicting the future as accurately as possible, given all the information available, including historical data and* *knowledge of any future events that might impact the forecasts*." Forecasting is an important component of business planning. Regardless of the business context or time horizon, without forecasts planning is very inefficient and costly because of the lack of guidance to support the decision. Planning in such a fashion would likely lead to poor decision-making and in turn poor business performance.

Accurate forecasting is difficult, but firms that achieve the feat better than their competitors may realize advantages over them. To predict a phenomenon accurately really depends on how well we understand the factors that contribute to it and how much data is available (Hyndman & Athanasopoulos, 2014). A quality forecast can capture relationships and patterns from the data. Moreover, quality forecasts have the ability to discriminate among random fluctuations, or noise, that should be ignored and an actual pattern or shift that should be modeled and possibly extrapolated (Hyndman & Athanasopoulos, 2014).

Many practitioners will create their own forecasts to try and gain better insight about their products. The main reason to perform this daunting task is to verify the quality of decision-support that is being provided. In addition, practitioners have the ability to incorporate feedback about their products into the models being developed.

We structure this paper by turning the business problem into an analytics problem, describe the technology we selected to support the problems, describe the analytics methodologies we employed, and provide some initial results. Lastly, we discuss the positive impacts of the solution implemented and identify areas for improvement.

Business Problem to Analytics Problem

Category managers are regularly faced with having to determine which products to incorporate in their assortment plan. Depending on the retailer, this could be a short (e.g. grocery), medium (e.g. clothing), or long term (e.g. service parts) planning decision. The assortment decision asks what are the optimal products to offer (Kök, Fisher, & Vaidyanathan, 2015).

The most important decision-support used by category mangers in facilitating their decision of which products to add/remove is the ability to gauge future demand for their next selling season. Such demand metrics often come in various forms based on the context. For example, propensity to purchase measures can gauge the overall likelihood of selling in a particular location, substitution-based choice models (e.g. Multinomial Logit) provide likelihood to purchase (or substitute) based on a specified substitutable set of products, and total unit demand quantity forecasts often provide a measure of total expected units sells, which might be at a stock keeping unit (SKU), store, region, or company-wide level.

We have discovered that these decision-makers often want:

- more than one demand forecast model for verification and reconciliation.
- the ability to generate such forecasts at scale in a timely and efficient manner.
- the ability to gauge the accuracy and reliability of the provided forecasts.
- the ability to identify what is causing a poor forecast and correct such issues in the future.

The justification for such requirements are the decision-makers need to build confidence in the decision-support metrics they are being provided. Commercial systems will regularly generate a forecast that can vary day to day, and such systems lack empirical evidence of their quality, or any transparency into what new modeling tweaks have been incorporated to make the demand forecasts more accurate over time. Having the ability to see and communicate such insights directly with the decision-facilitator builds this confidence and allows the decision-support to improve over time.

Methodology

In this study, we focus on generating long term planning horizon forecasts to support the types of products under investigation. The data measured and used to understand the future are time-series, which contain sequentially equally spaced measurements over time (i.e. time-series) from the previous season. Interestingly, the prediction is based solely on the historical pattern of the phenomena being forecast, thus the cause-effect relationship specified for the univariate model is in the form of legacy. Inherently, the only "factors" that must be identified for univariate models are trend and seasonality.

The most popular methods employed in this area are exponential smoothing and Autoregressive Integrated Moving Average (ARIMA) models. The objective is to estimate the next point(s) based on the pattern of previously measured observations. A typical retail example is the prediction of the number of SKU unit sales in the next period or several periods. In our context of long term planning, we are more interested in the time aggregation of predicted time series sales rather than the distinct pattern over the entire selling season.

Exponential smoothing methods have been used since the 1950s, but it has not been until the last decade that a framework has been built to incorporate such methods in R so that can easily identify a time-series' characteristics and perform model specification automatically. For example, Table 1 shows a modified table from (Hyndman & Khandakar, 2007; Taylor, 2003) that breaks out the fifteen possible exponential smoothing models that could be identified. In addition models may have an additive or multiplicative error type.

	Seasonal Component		
Trend Component	N (None)	A (Additive)	M (Mulitplicative)
N (None)	NN	NA	NM
A (Additive)	AN	AA	AM
A_d (Additive damped)	A_dN	A_dA	A_dM
M (Mulitplicative)	MN	MA	MM
M_d (Multiplicative damped)	M_dN	M_dA	M_dM

Table 1: Fifteen possible exponential smoothing methods

Using this table, an exponential smoothing model can be specified/identified using three model characteristics in order: Error, Trend, and Seasonality (E,T,S).

Technology

For this investigation we used the R programming language. R is a software environment for data analysis, computing, and graphics that is open-source and freely accessible under the GNU *General Public License, version 2* (Ihaka & Gentleman, 1996). Today, R has over 2 million users worldwide and is taught in universities and used in enterprise environments by data analysts, data scientists, IT, and analytics professionals (Hornick & Plunkett, 2013). According to Siegel (2013), author of Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die *"[R is] the leading free, open-source software tool for Predictive Analytics*."

The R system is composed of two components: the base R system which contains fundamental functions and graphics capabilities. The other and much more substantive component is the extensive repository of freely developed and contributed packages by the R community that can be easily incorporated via downloaded installations. These can be found at the Comprehensive R Archive Network (https://cran.r-project.org/).

The packages we used in this project are the **forecast** package (Hyndman & Khandakar, 2007), and the **for each** package (Weston, 2014). According to Hyndman and Athanasopoulos (2014), the increases in forecast incidence (i.e. volume, frequency) necessitate the need for more automation over manual processing. We found this to be true in our case, as well as the need to generate predictions at scale in a reasonable amount of time.

Data & Modeling Building

The data set investigated consists of 160,222 unique SKU time-series from a national retailer. Each time-series measures 13 sequentially spaced values over multiple selling seasons. SKUs have varying degrees of historical point of sales data ranging from less than a year to more than five years.

The forecast package provides exponential time-series procedures that automatically perform model identification of error, trend, and seasonality. You can easily tune the provided parameters, but we found a user-specified additive error performed significantly better than in the case of multiplicative error, which led to forecasts over predicting by several factors. Aside from setting the error to additive, all other possibilities were identified automatically in the model selection procedure.

Results

We generated the forecasts on a machine with a 64-bit Windows 7 Enterprise operating system, having an Intel(R) Xenon(R) CPU E3-1270 v3 @ 3.50GHz processor, with 16 GB of RAM. The version of R used was 3.1.2. Among the fifteen possible ETS models, only six were identified and selected. Where the SKU history was insufficient, either the mean or naïve (i.e. latest data point) was used to forecast the future. Since the forecasts generated were for long term planning purposes, it is most important to the decision-maker to know the quality of total forecast units rather than the predicted 13 point time-series forecast. For this reason, the 13 point forecasts were aggregated to obtain a total predicted amount for the upcoming selling season. Figure **1**Figure **1** shows the aggregated hold out versus the actual SKU total number of units sold.

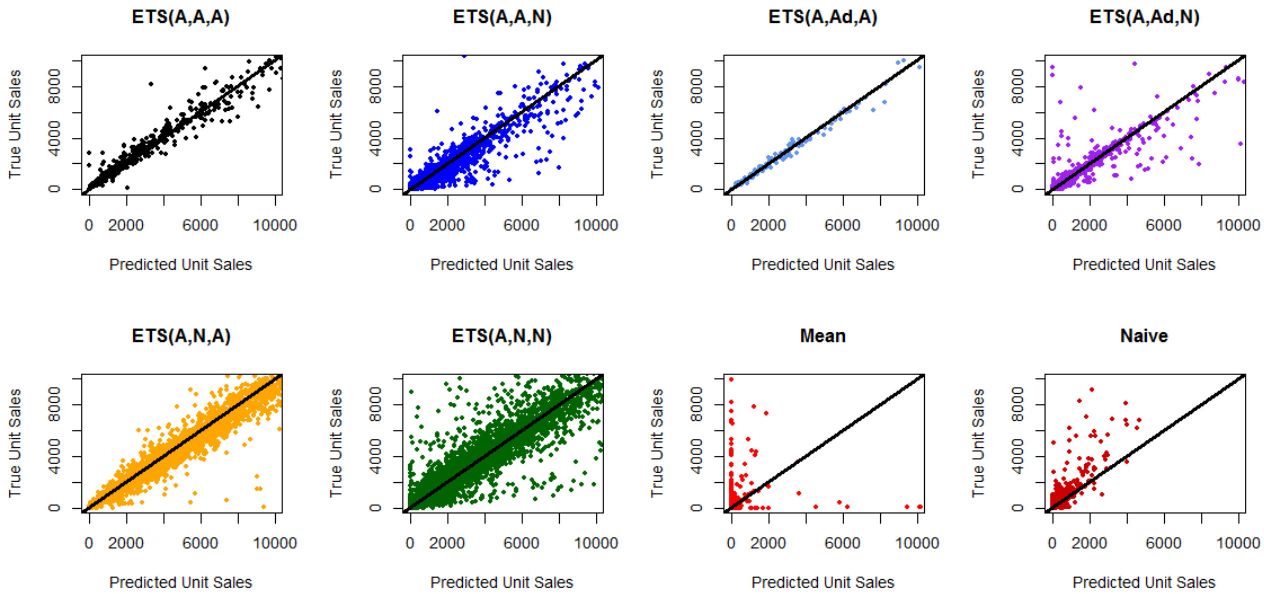

Figure 1: Predictions versus actual unit sales by model type.

Overall, it took approximately 36 hours of runtime to produce these forecasts using the 64-bit version of R, and the forecast and foreach packages. The **for each** package provides sequential or parallel processing over collections of objects such as SKUs. Earlier versions of the project used the 32-bit version of R and execution time was two to three times longer than the 64 bit version.

Interestingly, using this free R package does surprisingly well for different levels of sales. Figure **1** shows that the mean and naïve method perform poorly at all levels of sales, but this was expected. Table **1** shows assessment statistics for each model grouped by sales history. While, ETS(A,N,N) was specified the most in our time-series it had the greatest median absolution deviation in all cases with sufficient sales history.

We were not able to compare these results to those generated by a commercial ERP system because the predictions were not being stored for benchmarking purposes. We have begun extracting and storing them as a way to gauge their predictive ability.

Method	Sales History	Num Obs	MAD	MAPE	MSE	RMSE
Mean	t < 1 year	4,188	1,553,026	371	13,563,087	3,683
Naive method	t < 1 year	3,802	776,313	204	3,184,203	1,784
ETS(A,A,A)	1 <= t < 3	9	54	6	97	10
ETS(A,A,N)	1 <= t < 3	1,089	334,162	307	2,380,830	1,543
ETS(A,Ad,A)	1 <= t < 3	2	12	6	43	7
ETS(A,Ad,N)	1 <= t < 3	134	20,358	152	363,932	603
ETS(A,N,A)	1 <= t < 3	50	296	6	111	11
ETS(A,N,N)	1 <= t < 3	12,978	1,046,129	81	587,983	767
Mean	1 <= t < 3	36	5,043	140	134,355	367
Naive method	1 <= t < 3	19	670	35	1,870	43
ETS(A,A,A)	t >= 3	640	659,183	1,030	6,503,875	2,550

ETS(A,A,N)	t >= 3	17,488	3,099,280	177	9,883,566	3,144
ETS(A,Ad,A)	t >= 3	156	55,438	355	1,426,486	1,194
ETS(A,Ad,N)	t >= 3	7,156	1,066,807	149	1,733,464	1,317
ETS(A,N,A)	t >= 3	10,197	8,090,775	793	17,014,585	4,125
ETS(A,N,N)	t >= 3	98,762	12,764,790	129	2,212,977	1,488

Table 2: Forecasting results: exponential smoothing, mean and naïve methods.

Conclusions & Future Research

Decision-makers are becoming more aware of the importance of the parameters that they are using to support their decision-making. We show that using free and open-source modeling tools such as R and its relevant packages can be very powerful at generating univariate time-series forecasts for hundreds of thousands of products.

While there is uncertainty in these models, we have identified step level shifts in sales as a primary cause of poor performance in these automatic time-series predictions and have been testing various heuristics to identify optimal intervals to mitigate these shifts when fitting the model. The development costs and time have been relatively minor because R is free and the packages we have used have been robust enough to support a major retailer's required decision-support. The forecast package is a powerful package to perform model identification automatically, and the foreach package provides the parallel runtime performance gains that make this platform efficient enough to generate the predictions required at scale.

References

Hornick, M. F., & Plunkett, T. (2013). *Using R to unlock the value of big data*: McGraw Hill.

Hyndman, R. J., & Athanasopoulos, G. (2014). *Forecasting: principles and practice*: OTexts.

Hyndman, R. J., & Khandakar, Y. (2007). *Automatic time series for forecasting: the forecast package for R*. Retrieved from

Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, *5*(3), 299-314.

Kök, A. G., Fisher, M. L., & Vaidyanathan, R. (2015). Assortment planning: Review of literature and industry practice *Retail Supply Chain Management* (pp. 175-236): Springer.

Siegel, E. (2013). *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*: John Wiley & Sons.

Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, *19*(4), 715-725.

Weston, S. (2014). *Using The foreach Package*. Retrieved from Comprehensive R Archive Network (CRAN): https://cran.r-project.org/web/packages/foreach/vignettes/foreach.pdf