Int. J. Production Economics 146 (2013) 185–198

Contents lists available at ScienceDirect

Int. J. Production Economics

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

Forecasting aggregate demand: Analytical comparison of top-down and bottom-up approaches in a multivariate exponential smoothing framework

Giacomo Sbrana [a,*], Andrea Silvestrini [b]

[a] Rouen Business School, 1 Rue du Maréchal Juin, 76130 Mont Saint Aignan, France
[b] Bank of Italy, Economics, Research and International Relations, Via Nazionale 91, 00184 Roma, Italy

ARTICLE INFO

Article history:
Received 25 October 2012
Accepted 23 June 2013
Available online 5 July 2013

Keywords:
Top-down forecasting
Bottom-up forecasting
Multivariate exponential smoothing

ABSTRACT

Forecasting aggregate demand represents a crucial aspect in all industrial sectors. In this paper, we provide the analytical prediction properties of top-down (TD) and bottom-up (BU) approaches when forecasting the aggregate demand using a multivariate exponential smoothing as demand planning framework. We extend and generalize the results achieved by Widiarta et al. (2009) by employing an unrestricted multivariate framework allowing for interdependency between its variables. Moreover, we establish the necessary and sufficient condition for the equality of mean squared errors (MSEs) of the two approaches. We show that the condition for the equality of MSEs holds even when the moving average parameters of the individual components are not identical. In addition, we show that the relative forecasting accuracy of TD and BU depends on the parametric structure of the underlying framework. Simulation results confirm our theoretical findings. Indeed, the ranking of TD and BU forecasts is led by the parametric structure of the underlying data generation process, regardless of possible misspecification issues.

1. Introduction

Forecasting aggregate demand represents a crucial aspect in all industrial sectors (see Zotteri and Kalchschmidt, 2007; Kalchschmidt et al., 2006). Consider, for instance, a retail company which offers a broad range of items to its customers. In order to reduce inventory costs and to manage efficiently the supply chain planning process, the company has to rely on accurate predictions for each demand segment and for the whole aggregate demand (see Kerkkänen et al., 2009, for a discussion on the impacts that sales forecast errors have on the supply chain). In the field of industrial maintenance, a related issue is faced when forecasting future spare parts demand, which is needed in order to keep equipment operating properly. This problem is very relevant, for instance, in military logistics, which represents one of the largest outlays of military budgets (see Moon et al., 2012, 2013, for an analysis of demand for spare parts in the South Korean Navy). A similar problem occurs in the automobile industry (Fliedner and Lawrence, 1995).

In the context of aggregate demand forecasting, one of the most important issues faced by both theoretical and empirical literature can be summarized as in the abstract of Dunn et al. (1976, p. 68): "*Should statistical forecasts be constructed by aggregating data to each level for which forecasts are required or aggregating the forecasts from the lower levels? The relevant literature suggests no general answer*". Despite that this paper dates back to the seventies, the question raised by the authors remains an open issue.

In general, the aggregate demand can be forecasted using different procedures. In this paper, we compare the forecasting performance of top-down (TD) and bottom-up (BU) approaches whose definition can be found, for example, in Zotteri et al. (2005, p. 480): "*In the bottom-up approach, individual forecasts for each demand segment (e.g., single stock-keeping unit, single day, or single store) are combined to produce a forecast of aggregate demand (e.g., group of products, week or group of stores). This is referred to as the cumulative forecast since it is the sum of individual lower level forecasts. In the top-down process, aggregate demand data are used to forecast aggregate demand, etc.*".

The goal of this paper is to provide explicit analytical expressions for the TD and BU approaches when forecasting the aggregate demand. We assume as production planning framework the multivariate version of the simple exponential smoothing. The simple

* Corresponding author. Tel.: +33 232 824 673.
E-mail addresses: giacsbrana@hotmail.com, gsb@rouenbs.fr (G. Sbrana), andrea.silvestrini@bancaditalia.it (A. Silvestrini).

0925-5273/$ - see front matter
http://dx.doi.org/10.1016/j.ijpe.2013.06.022

exponential smoothing, also known as exponentially weighted moving average (EWMA), has a long tradition in forecasting economic time series (Muth, 1960).

Technically, an EWMA smoothing recursion leads to the same forecasts produced by an IMA(1,1) model, which is the reduced form of a random walk plus noise structural time series model (Harvey, 1989). Regarding the IMA(1,1), it is worth quoting the Nobel prize winner Clive Granger: "*This model provides a very good representation of a wide range of economic time series […] we do not advocate the adoption of this model in all occasions. However, if some simple specific model is to be assumed on a priori grounds, we feel that the first-order integrated moving average process is a serious candidate for economic time series in general.*" (the quotation is taken from Granger and Newbold, 1977, p. 203). Indeed, despite its simplicity, exponential smoothing represents a strong candidate compared to other complex models as also discussed by several authors (see for example Dekker et al., 2004; Fliedner and Lawrence, 1995; Fliedner, 1999; Moon et al., 2012, 2013).

As a distinctive feature, we adopt a multivariate framework in order to allow for interdependencies in the demand environment, while most previous studies featured univariate demand models (with few exceptions such as Chen and Blue, 2010; Kremer et al., 2012). Therefore, our results generalize the results achieved by Widiarta et al. (2009) by avoiding coefficient restrictions and allowing for interdependency among the individual components. This is relevant since in empirical applications the concept of interdependency is usually neglected in order to avoid complications. Furthermore, we derive the necessary and sufficient condition for the equality of mean squared errors (MSEs) of the TD and BU approaches. In particular, our results shed light on the analytical properties of TD and BU when assuming a first order vector integrated moving average model, which corresponds to the multivariate exponential smoothing with no restrictions on parameters. To our knowledge, such a framework has never been used to compare TD and BU approaches.

Recently, other papers have compared the forecasting properties of alternative demand planning approaches based on MSEs. For instance, Chen and Blue (2010) consider a bivariate first order vector autoregressive framework. Moreover, Widiarta et al. (2009) assume simple exponential smoothing as the forecasting technique for both TD and BU. They prove that TD and BU are equally efficient in terms of MSE if the individual components follow univariate MA(1) processes with identical MA coefficients and if the smoothing constants used for forecasting the individual components are equal (p. 91). Whereas these authors give a condition for equal efficiency, relying on rather strong restrictions, our results are valid in general, regardless of any restrictions. Indeed, contrary to Widiarta et al. (2009), we relax the assumption of identical parameters of the single subaggregate components. This is clearly more realistic since, in a standard production planning context, the parametric structure of the components is hardly ever identical.

The remainder of the paper is structured as follows. After a brief literature review in Sections 2, we present in 3 the methodological framework, which is based on the multivariate exponential weighted moving average. In Section 4 we derive the parameters and the MSE of the TD approach, while in Section 5 we focus on the BU approach. In Section 6, we give the necessary and sufficient condition for the equality of MSEs. Then, in Section 7, using a simple bivariate model, we show that the mentioned condition can be achieved even when the single components differ in dynamics. To this purpose, we provide conditions under which the equality of MSEs holds while the equality of predictors does not. In Section 8 we present results of a simulation study to compare the out-of sample MSEs and mean absolute errors (MAEs) of TD and BU. Section 9 concludes. All proofs are relegated to Appendix.

2. Literature review

The literature on TD and BU approaches is wide and considers different frameworks. For this reason, we do not attempt to survey all the contributions, but rather to give the main references related to our paper.

The first reference literature is the time series literature where TD and BU are often referred to as aggregate and disaggregate specifications. When the data generation process is assumed to be a vector ARMA model, the consequences of contemporaneous (cross-sectional) aggregation have been discussed since the original contributions of Granger and Morris (1976) and Box and Jenkins (1976). Most of the theoretical results on the aggregation of ARMA are collected in Lütkepohl (1987, Chapter 4), Lütkepohl (2006, Section 2.4) and Lütkepohl (2009). Other prominent contributions are those of Rose (1977), Tiao and Guttman (1980), Wei and Abraham (1981), Kohn (1982), and Lütkepohl (1984a, 1984b, 1987).

The second reference literature encompasses the demand planning studies focusing on the effectiveness of TD and BU approaches. Fliedner (1999) argues that forecast performance is dependent upon the statistical nature of the disaggregate items comprising the aggregate series: In particular, higher positive/negative correlation leads to improved forecast performance at the aggregate level. Weatherford et al. (2001) show that a fully disaggregated forecasting method outperforms aggregated forecasts. More recently, Dekker et al. (2004) focus on seasonal demand forecasts, while Moon et al. (2012, 2013) consider in detail alternative forecasting methods for predicting the demand for spare parts by the South Korean Navy. In general, the choice between competitive specifications seems to be based on the specific framework employed and case study analyzed. Indeed, as argued by Zotteri et al. (2005) and Zotteri and Kalchschmidt (2007), the choice of the appropriate aggregation level depends on the underlying data generation process.

Our contribution is related to these two types of literature focusing on the issue of forecasting aggregated variables. One of the major gap left by the previous literature is the lack of analytical results establishing conditions under which the TD approach outperforms the BU and vice versa. This paper fills this gap by shedding light on the algebraic conditions determining the forecasting performance of TD and BU assuming the multivariate exponential smoothing as the framework. Clearly, this is only one of the possible production planning framework and therefore it can be considered as the starting point to be further developed as future research.

3. The demand planning framework

In this section we present our assumptions on the demand planning framework, the forecasting problem and the methodology adopted.

We consider a system describing the demand estimates for $i=1,2,\ldots,N$ products at time $t=1,2,\ldots,T$. We assume that demand estimates follow an unrestricted multivariate exponential weighted moving average (EWMA) process where the vector x_t is N-variate and