



FORECASTING AUTOREGRESSIVE DEMANDS WITH TEMPORAL AGGREGATION

Bahman Rostami Tabar, Yves Ducq, Mohamed Zied Babai, Aris Syntetos

► To cite this version:

Bahman Rostami Tabar, Yves Ducq, Mohamed Zied Babai, Aris Syntetos. FORECASTING AUTOREGRESSIVE DEMANDS WITH TEMPORAL AGGREGATION. 9th International Conference on Modeling, Optimization & SIMulation, Jun 2012, Bordeaux, France. hal-00728574

HAL Id: hal-00728574

<https://hal.science/hal-00728574>

Submitted on 30 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FORECASTING AUTOREGRESSIVE DEMANDS WITH TEMPORAL AGGREGATION

B. Rostami Tabar^{1,2}, Y. Ducq¹

¹IMS / University of Bordeaux
351, Cours de la Liberation
33405 Talence Cedex, France
bahman.rostamitabar@u-bordeaux1.fr, yves.ducq@ims-bordeaux.fr

M.Z. Babai²

²BEM-Bordeaux Management
School, 680 cours de la libération,
Talence, France
mohamed-zied.babai@bem.edu

A.A. Syntetos³

³Salford Business School, University of Salford, Maxwell Building,
Manchester M5 4WT, UK,
a.syntetos@salford.ac.uk

ABSTRACT: Demand forecasting performance is subject to the variability and uncertainty underlying the time series related to the Stock Keeping Units (SKUs) an organisation is dealing with. Different strategies may be used to reduce the demand variability and forecast error, thus improve the forecasting performance. One such strategy is to aggregate demand in lower-frequency ‘time buckets’, thereby reducing the demand variability. In this paper, we investigate the impact of temporal aggregation on forecasting performance. We assume that the underlying demand follows a first-order autoregressive process [AR(1)] and a Single Exponential Smoothing (SES) procedure is used to estimate the level of demand. The theoretical forecast errors are derived for the aggregated and non-aggregated demand in order to contrast the relevant forecasting performances. A theoretical analysis indicates that performance improvements achieved through the aggregation strategy are a function of the aggregation level, the smoothing constant value and the autoregressive parameter. The results are also empirically validated by means of experimentation with the sales data related to 555 AR(1) series from a major European superstore.

KEYWORDS: Temporal Aggregation, Forecasting Performance, AR(1) process, Exponential Smoothing.

1 INTRODUCTION

Demand variability and uncertainty are among the most important management challenges facing modern industries and companies. These issues have been addressed in the academic literature for many years.

The existence of high variability in demand for fast moving and slow/intermittent moving items (items with a high ratio of zero observations), pose considerable difficulties in terms of forecasting and stock control. Deviations from the degree of variability accommodated by the Normal distribution often render standard forecasting and inventory theory inappropriate.

There are many strategies that may be used to reduce the demand variability and thus to improve the forecasting (and inventory control) performance. An intuitively appealing strategy is to aggregate demand in lower-frequency ‘time buckets’, thereby reducing the presence of potential zero observations (in case of intermittent demand) or generally reduce variability in case of fast demand.

Such an aggregation strategy is often referred to in the academic literature as Temporal Aggregation. Another aggregation strategy discussed in the literature refers to Cross-Sectional Aggregation, which involves aggregat-

ing different time series to reduce variability across a group of items.

Aggregation has been used in economics and econometrics for the last 40 years and therefore the idea of aggregation is not new. In a production planning framework, many researchers have focused on the effectiveness of cross-sectional aggregation and especially the bottom-up and top-down approaches. The former involves the aggregation of individual SKU forecasts to the group level whereas the latter relates to forecasting directly at the group level (followed by the disaggregation of forecasts at the individual level). However there are comparatively less studies focusing on the effects of temporal aggregation in production planning and inventory management frameworks.

In this paper we assume that the underlying series follows an Autoregressive process of order one, AR(1) and the forecasting method is the Single Exponential Smoothing (SES). Both assumptions are very realistic as demand very often follows an AR(1) process and SES has often been discussed to be a most popular forecasting method. In our previous work (Rostami et al., 2012), we have evaluated by analytical approach the impact of temporal aggregation on forecasting when the demand process follows MA(1) and AR(1). In this paper in addition to theoretical analysis, we investigate the empirical analysis for AR(1) demand process. We use a real data-

set from a European Grocery Store in order to analyze the performance of the temporal aggregation forecasting approach as compared to the classical approach.

The remainder of this paper is structured as follows. In Section 2 we provide a literature review on demand aggregation and forecasting. Section 3 presents the notations and assumptions. In Section 4 a mathematical evaluation of the MSE for non-aggregated and aggregated demand is provided. Section 5 presents details related to the empirical investigation and the results. The paper concludes in Section 6 with the implications of our work along with some natural next steps of research.

2 LITERATURE REVIEW

In this section we briefly present a literature review on demand aggregation and forecasting. Demand variability may affect the different systems in the organizations such as demand forecasting, production planning and inventory management systems.

It has been shown by Theil (1954), Yehuda and Zvi (1960) and Aigner and Goldfeld (1974) that demand uncertainty can be effectively reduced through appropriate demand aggregation and forecasting. In the literature of supply chain and demand planning, demand aggregation is known as a 'risk-pooling' strategy to reduce demand fluctuation for more effective material/capacity planning (Chen and Blue (2010)). In some papers, the authors have shown that aggregation of data improves the accuracy of the forecasts (Armstrong (2001); Zotteri and Verganti (2001)). This improvement can be explained in terms of the lower variability of the aggregated demand as compared to the original demand streams.

There are two types of aggregation discussed in the academic literature.

In *temporal aggregation*, the aggregated time series is obtained through the sum of every m periods of the high frequency data, where m is called the aggregation level.

When considering temporal aggregation, there are two different ways of aggregation: non-overlapping and overlapping. In the former case the time series are divided into consecutive non-overlapping buckets of time; that is if the number of observations in the basic series is n and the aggregation level is m , then the number of observations in the aggregated series is N , where N is the integer part of n divided by m .

In the overlapping case, at each period the oldest observation is dropped and the newest is included, therefore the number of aggregated series is $n-m+1$. In both the non-overlapping and overlapping cases, the length of the time bucket equals the aggregation level. In this paper the case of temporal non-overlapping aggregation will be considered.

Another type of aggregation refers to the *cross-sectional* or *contemporaneous* aggregation and this is further discussed below.

2.1 Cross-sectional aggregation

Most of the forecasting literature in this area has looked at the comparative performance of the top-down (TD) and the bottom-up (BU) methods.

In the top-down aggregate forecasting approach, demand is first aggregated, then the forecast is made by applying a forecasting method at the aggregate level and finally the forecast is disaggregated back to the original level by applying an appropriate disaggregation method. In the subaggregate approach, called bottom-up approach, the aggregate demand forecast is produced by combining individual forecasts for each demand segment, i.e. a separate forecasting model is developed for each item in the family (Zotteri et al. (2005)).

In the literature of cross-sectional aggregation, when dealing with the impact of aggregation on the forecasting performance, there are different findings. Lütkepohl (1984) and Fliedner (1999) have shown that the forecast performance at the aggregate level for a derived (top-down) strategy is better. On the other hand, Orcutt et al. (1968), Edwards and Orcutt (1969) argue that information loss is substantial in aggregation and therefore the bottom-up method gives more accurate forecasts. Some authors take a contingent approach and analyze the conditions under which one approach produces more accurate forecasts than the other. Lapide (1998) and Fliedner (2001) summarized some practical guidelines for the situations where top-down is preferable to bottom-up and vice versa.

Widiarta et al. (2007) and Widiarta et al. (2009) have studied analytically and by means of simulation the relative performance of TD and BU strategies when the item demands follow a first-order univariate autoregressive [AR(1)] and first-order moving average [MA(1)] process. They showed that when the time series for all of the subaggregate components follow an MA(1) process with identical coefficients of the serial correlation term, there is no difference in the relative performance of TD and BU forecasting strategies. However, when the subaggregate components follow an AR(1) process with the same autoregressive parameters for all the items, the superiority of each strategy is a function of the autoregressive parameter; in particular, they showed that when the autoregressive parameter is smaller than or equal to $1/3$, the maximum difference in the performance of the two forecasting strategies is only 1%. However, if the autoregressive parameter is greater than $1/3$, then the bottom-up strategy consistently outperforms the top-down one.

2.2 Temporal aggregation

Temporal aggregation refers to aggregation in which a low frequency time series (e.g. quarterly) is derived from a high frequency time series (e.g. monthly) (Nikolopoulos et al. ,2011). Most of the literature that deals with temporal aggregation may be found in the Economics discipline.

Amemiya and Wu (1972) have shown using simulation that forecasts generated from aggregated series in the

case of ARIMA models perform remarkably well compared to those from the original series. Souza and Smith (2004) have also shown that for Auto-Regressive Fractionally IMA (ARFIMA) models, temporal aggregation results in a bias reduction. On the other hand, Rose (1977) compared forecasts from an aggregate ARIMA model and an underlying disaggregate ARIMA model and showed that forecasts from the aggregate model are in general less efficient

Willemain et al. (1994) have empirically explored the effects of temporal aggregation on forecasting intermittent demand. In this study Croston's method has been applied as a forecasting approach. They have compared the results of the forecasting approach, using 16 real data sets with 905 daily observations. Daily and weekly data have been considered. In this study, the daily data form corresponds to the basic series and the weekly data form corresponds to the aggregated series with an aggregation level that is equal to seven. Results based on the mean absolute percentage error (MAPE) have been obtained for both forms and the researchers showed a significant reduction in forecasting errors when weekly demand aggregate data were used instead of daily data.

Athanasopoulos et al. (2009) have recently looked at the effects of temporal aggregation on forecasting accuracy in the tourism industry. They have conducted an empirical investigation using 366 monthly series and three forecasting methods tested in the M3 competition data, namely the exponential smoothing, the ARIMA method, and a commercial software, Forecast Pro. They have compared forecast errors from forecasting yearly data directly with errors from temporally aggregating the forecasts produced for monthly and quarterly data. They have considered one and two forecast-ahead periods and three statistical measures have been used to compare the results: MAPE, Mean Absolute Scaled Error (MASE) and Median Absolute Scaled Error (MdASE). The aggregated forecasts (whether produced from monthly or quarterly data) were more accurate than forecasts produced directly from yearly data.

Nikolopoulos et al. (2011) have empirically analysed the effects of temporal aggregation on forecasting intermittent demand requirements. The proposed approach called aggregate-disaggregate intermittent demand approach (ADIDA) works as follows. Firstly the original time series is converted to a high frequency time series according to the level of aggregation, then a forecasting method is applied to this new series to obtain a cumulative forecast and finally the cumulative forecast is broken down (disaggregated) to the original form. In this research 5,000 SKUs over 7 years (84 monthly demand observations) from the Royal Air Force(RAF, UK) have been used and 3 methods (Naïve, Croston and Syntetos-Boylan Approximation) have been applied to basic series and aggregated series based on an aggregation level varied from 2 to 24. The results show that the proposed ADIDA methodology may indeed offer considerable improvements in terms of forecast accuracy. Four different accuracy measures have been used including: Bias (Mean Error, ME and Median Error, MdE), Scaled Er-

rors (MAE, MdAE), Mean Squared Errors (MSE) and a relative MSE error (Relative Geometric Root MSE, RGRMSE) Various disaggregation approaches were also considered and the best one was found to be that related to an equal weight split of the aggregated forecasts to the disaggregated components.

Luna and Ballini (2011) have used an aggregation approach to predict daily time series of cash money withdrawals in the neural forecasting competition. Each time series consists of 735 samples, from March 18, 1996, to March 22, 1998 and have been used to forecast 56 daily steps ahead. Daily samples are aggregated to give weekly time series and then an adaptive fuzzy rule-based system is applied to predict the 8- step-ahead, thus aggregation reduces the forecast horizon from 56 to 8 steps. Two different aggregated approaches are evaluated for this purpose: the historical top-down (TD-H) approach and the daily top-down (TD-DM) approach, where the main difference between the two is the disaggregation procedure. The statistical measure of symmetric mean absolute percentage error (sMAPE) and mean absolute error (MAE) have been used to compare the results. The researchers show that the aggregated base models perform well for both TD-H and TD-DM regarding the basic daily forecasts.

Babai et al. (2012) have compared various forecasting methods with and without aggregation based on stock control measures. They have considered ADIDA as a temporal aggregation approach along with a periodic order-up-to-level stock control policy. Three forecasting methods, SES, Croston and SBA have been used. They have used the same dataset used by Nikolopoulos et al. (2011) which consists of 5000 SKUs from the RAF, and the demand was assumed to be Negative binomially distributed. Performance was reported through the inventory holding and backlog volumes and costs, for three possible targets Cycle Service Level (CSL): 90%, 95% and 99%. For high CSLs, the aggregation approach has been shown to be more efficient and for low CSLs, the aggregation approach was outperformed by the classical one when Croston's method was used. For SES the aggregation approach outperforms the classical approach even for low CSLs.

3 NOTATION AND ASSUMPTIONS

Let d_t be the basic demand series that follows a first order autoregressive process, AR(1), that can be mathematically written in period t as follows:

$$d_t = \mu(1 - \phi) + \varepsilon_t + \phi d_{t-1}, \text{ where } |\phi| < 1. \quad (1)$$

Assume D_T is the m period non-overlapping aggregate of d_t defined as

$$D_T = \sum_{i=0}^{m-1} d_{m^*T-i} = (1 + B + \dots + B^{m-1}) d_{m^*T}. \quad (2)$$

d_t	Basic demand in period t
D_T	Aggregated demand in period T
m	Aggregation level
N	Total number of observations
ε_t	Random term for basic demand in period t , normally distributed with zero mean, variance σ^2 and $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$ for all $k \neq 0$
ε'_T	Random term for aggregated demand in period T , normally distributed with zero mean, variance σ^2 and $Cov(\varepsilon'_T, \varepsilon'_{T-k}) = 0$ for all $k \neq 0$
f_t	Forecast of basic demand in period t
\bar{f}	The expected forecast of basic demand in any time period
F_T	Forecast of aggregated demand in period T
α	Smoothing constant used in exponential smoothing method, $0 < \alpha \leq 1$
γ_k	Covariance of lag k of basic demand
γ'_k	Covariance of lag k of aggregated demand
MSE_{BA}	MSE before aggregation
MSE_{AA}	MSE after aggregation
ϕ	Autoregressive parameter before aggregation, $ \phi < 1$
ϕ'	Autoregressive parameter after aggregation, $ \phi' < 1$
B	Backshift Operator
d	Integrated parameter
$\mu = \bar{d}$	Expected value of basic demand in any time period
μ'	Expected value of aggregated demand in any time period
0_n	$1 \times n$ vector of zeros
C	$1 \times [2(m-1)+1]$ vectors of C_i which is the coefficient of B^i in the polynomial $(1 + B + \dots + B^{m-1})^2$

Table 1 Notation

All the notations used in this paper are provided in Table 1.

The forecasting method considered in this study is Single Exponential Smoothing (SES); this method has been applied in many companies and most of the managers use this method in a production planning environment because of its simplicity.

Using SES to forecast the demand in period t means that

$$f_t = \alpha d_{t-1} + (1 - \alpha) f_{t-1}, \quad (3)$$

which, through recursive substitutions, can also be written in an infinite summation as

$$f_t = \sum_{k=1}^{\infty} \alpha (1 - \alpha)^{k-1} d_{t-k}. \quad (4)$$

We further assume that the standard deviation of the demand is significantly smaller than the expected value of the demand, so that the probability of a negative value of the generated values is negligible.

Constraining ϕ to lie between -1 and 1 in (1) means that the process is stationary.

4 MATHEMATICAL EVALUATION

In this section we analytically compare the mean square error (MSE) of the forecast before and after aggregation. The comparisons will result in the derivation of theoretical rules that indicate under which conditions forecasting of aggregated demand is theoretically expected to perform better than forecasting of the basic demand. These theoretical rules will be based on the aggregation level and control parameters. The cut-off values to be assigned to both criteria will be the outcome of a numerical analysis to be conducted on the theoretical results. Having obtained the cut-off values, we can then specify regions of superior performance of the aggregated approach over the non-aggregate approach. In this study MSE is used to compare the results of basic and aggregated forecasts.

The MSE is similar to the statistical measure of the variance of forecast errors (which consists of the variance of the estimates produced by the forecasting method under concern and the variance of the actual demand) but not quite the same since bias can also be explicitly considered.

$$\begin{aligned} Var(\text{Forecast Error}) &= Var(d_t - f_t) \\ &= E\{(d_t - f_t) - (\bar{d} - \bar{f})\}^2 \\ &= \{E(d_t - \bar{d})^2 + E(f_t - \bar{f})^2 - 2E(d_t - \bar{d})E(f_t - \bar{f})\}, \end{aligned}$$

Assuming a stationary model and then independence of terms we have:

$$\begin{aligned} Var(\text{Forecast Error}) &= E(f_t - \bar{f})^2 + E(d_t - \bar{d})^2 \\ &= Var(\text{Estimate}) + Var(\text{Actual Demand}) \\ &= MSE - Bias^2, \end{aligned} \quad (5)$$

We can show easily that the SES provides the unbiased estimations

$$Bias = E(d_t - f_t) = E(d_t) - E(f_t) = \mu - \mu = 0, \quad (6)$$

The bias of the forecasting method is zero, so the variance of forecast error is equal to the MSE, i.e. $MSE = Var(\text{Forecast Error})$.

4.1 MSE of basic demand

We begin the analysis by deriving the MSE_{BA} . In the basic series, the MSE is defined as follows

$$MSE_{BA} = Var(Forecast\ Error) = Var(d_t - f_t) \\ = (Var(d_t) + Var(f_t) - 2Cov(d_t, f_t)), \quad (7)$$

When the time series demand follows an AR(1) the following properties hold as long as $|\phi| < 1$ (William W.S. Wei, 2006, p34)

$$\gamma_k = \begin{cases} \sigma^2 & k=0 \\ 1-\phi^2 & k \geq 1 \end{cases}, \quad (8)$$

So

$$\gamma_1 = \phi\gamma_0. \quad (9)$$

We first start by deriving the covariance between the forecast and the demand in period t , by substituting (1) and (4) into (10) we have

$$Cov(d_t, f_t) = Cov(\mu(1-\phi) + \varepsilon_t + \phi d_{t-1}, \\ \sum_{k=1}^{\infty} \alpha(1-\alpha)^{k-1} d_{t-k}), \quad (10)$$

Then using the fact that $Cov(d_t, d_t) = Var(d_t)$ and $Cov(d_t, \varepsilon_t) = 0$ for all $k \geq 1$, and substituting (8) into (10), we get

$$Cov(d_t, f_t) = \frac{\alpha\phi\sigma^2}{(1-\phi^2)(1-\phi+\alpha\phi)} = \frac{\alpha\gamma_1}{1-\phi+\alpha\phi}, \quad (11)$$

Next, the variance of forecast in period t is derived as follows:

$$Var(f_t) = Var(\alpha d_{t-1} + (1-\alpha)f_{t-1}) = \\ \alpha^2 Var(d_{t-1}) + (1-\alpha)^2 Var(f_{t-1}) \\ + 2\alpha(1-\alpha)Cov(d_{t-1}, f_{t-1}). \quad (12)$$

Using the fact that $Var(f_t) = Var(f_{t-1})$ and substituting (8) and (11) into (12), we get:

$$Var(f_t) = \frac{\alpha\gamma_0(1+\phi-\alpha\phi)}{(2-\alpha)(1-\phi+\alpha\phi)}, \quad (13)$$

Now we can calculate the mean square error of the forecast before aggregation. By substituting (8), (11) and (13) into (7), we have

$$MSE_{BA} = \gamma_0 + \frac{\alpha\gamma_0(1+\phi-\alpha\phi)}{(2-\alpha)(1-\phi+\alpha\phi)} - \frac{2\alpha\gamma_1}{1-\phi+\alpha\phi}, \quad (14)$$

And finally by substituting (9) into (14), we get:

$$MSE_{BA} = \gamma_0 \left[\frac{1-\phi}{(1-0.5\alpha)(1-\phi+\alpha\phi)} \right]. \quad (15)$$

4.2 MSE of aggregated demand

Let us consider a non-overlapping aggregation of a first-order time series process, AR(1). Amemiya and Wu (1972) have shown that when the basic series follows an AR(p), the temporal aggregation of this series follows an ARMA(p,q*) process where

$$q^* = \left[\frac{(m-1)(p+1)}{m} \right]. \quad (16)$$

According to (16), we can show that when the basic series follows an AR(1), the aggregated series follows an ARMA(1,1) process with different parameters; the demand process after aggregation can be defined as follows:

$$D_T = \mu'(1-\phi') + \varepsilon'_T + \phi'D_{T-1} \text{ where } |\phi'| < 1. \quad (17)$$

By applying SES as a forecasting method, the forecast in period T is defined as

$$F_T = \sum_{k=1}^{\infty} \alpha(1-\alpha)^{k-1} d_{T-k}, \quad (18)$$

Now the MSE of the forecast after aggregation is

$$MSE_{AA} = Var(D_T - F_T) = Var(D_T) + Var(F_T) - \\ 2Cov(D_T, F_T). \quad (19)$$

In order to obtain the MSE_{AA} , we need to calculate the variance of the demand, the variance of the forecast and the covariance between the demand and the forecast.

If the time series follows an ARMA(1,1), we can show that the following properties hold:

$$\gamma'_k = \begin{cases} \frac{1-2\phi'\theta'+\theta'^2}{1-\phi'^2} \sigma'^2 & k=0 \\ \frac{(\phi'-\theta')(1-\phi'\theta')}{1-\phi'^2} & |k|=1 \\ \phi'\gamma'_{k-1} & |k|>1 \end{cases}. \quad (20)$$

It is obvious from (20) that

$$\gamma'_0 = Var(D_T) = \frac{1-2\phi'\theta'+\theta'^2}{1-\phi'^2} \sigma'^2. \quad (21)$$

We can derive the covariance between the demand and forecast by substituting (17) and (18) into (22) as follows:

$$\text{Cov}(D_T, F_T) = \text{Cov}\left(D_T, \sum_{k=1}^{\infty} \alpha(1-\alpha)^{k-1} D_{T-k}\right), \quad (22)$$

Then by substituting (20) into (22), we can get

$$\text{Cov}(D_T, F_T) = \frac{\alpha\gamma'_1}{1-\phi'+\alpha\phi'}, \quad (23)$$

And finally the variance of the forecast is

$$\begin{aligned} \text{Var}(F_T) &= \text{Var}(\alpha D_{T-1} + (1-\alpha)F_{T-1}) \\ &= \alpha^2 \text{Var}(D_{T-1}) + (1-\alpha)^2 \text{Var}(F_{T-1}) \\ &\quad + 2\alpha(1-\alpha)\text{Cov}(D_{T-1}, F_{T-1}), \end{aligned} \quad (24)$$

Using the fact that $\text{Var}(F_T) = \text{Var}(F_{T-1})$ and by substituting (20) and (23) into (24), we get

$$\text{Var}(F_T) = \frac{\alpha\gamma'_0}{2-\alpha} + \frac{2\alpha(1-\alpha)\gamma'_1}{(2-\alpha)(1-\phi'+\alpha\phi')}. \quad (25)$$

Now the MSE of the forecast after aggregation can be calculated by substituting (21), (22) and (25) into (19)

$$\text{MSE}_{AA} = \frac{1}{(1-0.5\alpha)} \left(\gamma'_0 - \frac{\alpha\gamma'_1}{1-\phi'+\alpha\phi'} \right). \quad (26)$$

4.3 The relationship of autocovariance between basic and aggregated demand

It has been shown that the autocovariance function on basic and aggregated series are related as follows (Wei, 2006)

$$\gamma'_k = (1+B+B^2+\dots+B^{m-1})^{2(d+1)} \times \gamma(mk+(d+1)(m-1)) \quad (27)$$

This can be transformed to a matrix form as follows:

$$\begin{bmatrix} \gamma'_0 \\ \gamma'_1 \\ \vdots \\ \gamma'_k \end{bmatrix} = A \times \begin{bmatrix} \gamma_{(-(d+1)(m-1))} \\ \vdots \\ \gamma_0 \\ \vdots \\ \gamma_{(mk+(d+1)(m-1))} \end{bmatrix},$$

where

$$A = \begin{bmatrix} C & 0_{mk} & \vdots & \vdots \\ 0_m & C & 0_{m(k-1)} & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0_{mk} & \vdots & \vdots & C \end{bmatrix}.$$

since $\gamma_k = \gamma_{-k}$ for all k, it's unnecessary to put autocovariance for negative values of k in the above matrix, so by removing these values from the above matrix, we need also to adjust the values of matrix A, in this matrix we delete the first $(d+1)(m-1)$ column corresponding to $\gamma_{-(m-1)(d+1)}, \dots$ and $\gamma_{(-1)}$ by adding them to column corresponding to $\gamma_{(m-1)(d+1)}, \dots, \gamma_1$ respectively; thus the above equation can be reduced to (28):

$$\begin{bmatrix} \gamma'_0 \\ \gamma'_1 \\ \vdots \\ \gamma'_k \end{bmatrix} = A_m^d \times \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{(mk+(d+1)(m-1))} \end{bmatrix}. \quad (28)$$

Where A_m^d is a modified matrix A, after deletion of the first $(d+1)(m+1)$ columns and adding them to the corresponding remaining columns of the matrix A.

Based on this assumption that the basic series follows an AR(1) and by calculating the matrix A_m^d using (27) and by substituting (8) into (28) we get:

$$\gamma'_0 = \gamma_0 \left(m + \sum_{k=1}^{m-1} 2(m-k)\phi^k \right), \quad (29)$$

$$\gamma'_1 = \gamma_0 \left(\sum_{k=1}^m k\phi^k + \sum_{k=1}^{m-1} k\phi^{2m-k} \right), \quad (30)$$

And for all $k > 1$, we have :

$$\gamma'_k = \gamma_0 (\phi^{(k-1)m+1} + 2\phi^{(k-1)m+2} + \dots + m\phi^{km} + (m-1)\phi^{km+1} + \dots + \phi^{(k+1)m-1}). \quad (31)$$

Finally, by using the relationship between successive autocovariance we can calculate the autoregressive parameter after aggregation as follows:

In equation (20) we have $\gamma'_k = \phi'\gamma'_{k-1}$, so

$$\phi' = \frac{\gamma'_k}{\gamma'_{k-1}}. \quad (32)$$

Now, by using (31) we calculate both the autocovariance of lag k and lag(k-1), then by substituting γ'_k and γ'_{k-1} into (32) we get

$$\phi' = \phi^m. \quad (33)$$

4.4 Comparison results

The performance and effectiveness of temporal aggregation versus non aggregation on forecasting can be compared by analyzing the ratio of their mean square errors. The relative efficiency of aggregation approach to classic

approach is defined as MSE_{BA}/MSE_{AA} . If this relative efficiency is greater than 1, we would conclude that aggregation is more efficient than classic approach, in the sense that it has a smaller mean square error. From (15) and (26)

$$Ratio = \frac{MSE_{BA}}{MSE_{AA}} = \frac{\gamma_0 \left(\frac{1-\phi}{(1-0.5\alpha)(1-\phi+\alpha\phi)} \right)}{\frac{1}{1-0.5\alpha} \left(\gamma'_0 - \frac{\alpha\gamma'_1}{1-\phi'+\alpha\phi'} \right)} \quad (34)$$

To be able to compare the result of the aggregated and the basic forecast, we have to multiply the basic demands and forecasts by the aggregation level m (or to divide the aggregated demands and their forecasts by m), which converts the above equation to

$$\frac{MSE_{BA}}{MSE_{AA}} = \frac{m^2 \gamma_0 \left(\frac{1-\phi}{(1-0.5\alpha)(1-\phi+\alpha\phi)} \right)}{\frac{1}{1-0.5\alpha} \left(\gamma'_0 - \frac{\alpha\gamma'_1}{1-\phi'+\alpha\phi'} \right)} \quad (35)$$

Now by substituting (29), (30) and (33) into (35)

$$\frac{MSE_{BA}}{MSE_{AA}} = \frac{m^2 \left(\frac{1-\phi}{(1-\phi+\alpha\phi)} \right)}{\left(m + \frac{\sum_{k=1}^{m-1} 2(m-k)\phi^k - \frac{\alpha(\sum_{k=1}^m k\phi^k + \sum_{k=1}^{m-1} k\phi^{2m-k})}{1-\phi^m + \alpha\phi^m}}{1-\phi^m + \alpha\phi^m} \right)} \quad (36)$$

It is obvious from (36) that the ratio is a function of the aggregation level m , the autoregressive parameter ϕ and the smoothing parameter α .

The analysis shows that when the value of the autoregressive parameter is $-1 < \phi \leq 0.5$, the ratio is always greater than 1. It means that the MSE of the aggregated forecast is less than MSE of the basic forecast, in other words, the forecast of aggregated demands are more accurate than the forecast of basic demands, but when the autoregressive parameter is $0.5 < \phi < 1$, the superiority of each strategy is a function of three parameters: the aggregation level m , the autoregressive parameter ϕ and the smoothing constant α .

In the following figures the MSE ratios have been shown for aggregation levels $m=2$, $m=7$ and $m=12$. It is easy to see the superiority of each strategy based on the value of the autoregressive parameter and the smoothing constant. The results show that for a highly positive value of the autoregressive parameter, the classical approach is more accurate than the aggregated approach, and consequently the aggregation approach is not recommended in this region. However, when the aggregation level increases, the parameters' regions in which the basic forecasts are superior disappear and when $m \rightarrow \infty$, for all autoregres-

sive values and smoothing constants, the aggregated forecasts are more accurate.

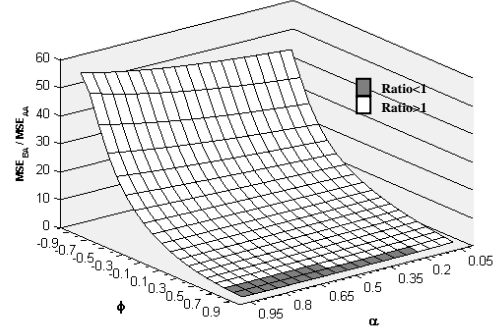


Figure 1. Impact of ϕ and α on the analytical ratio of MSE, $m=2$

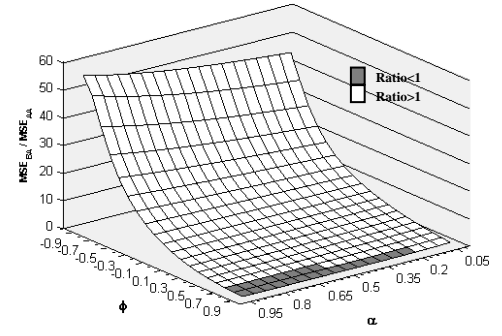


Figure 2. Impact of ϕ and α on the ratio of MSE, $m=7$

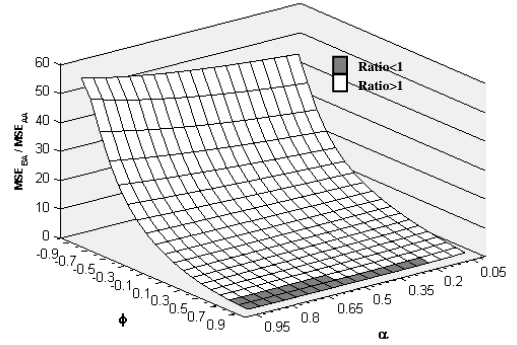


Figure 3. Impact of ϕ and α on the ratio of MSE, $m=12$

When ϕ is positive, the series is 'slowly changing' or can be considered as a positively autocorrelated process, in that the differences between successive observations tend to be relatively small. A negative ϕ or negatively autocorrelated process leads to a more irregular pattern, since observations tend to oscillate successively, in other words high values of d_t tend to be followed by low values and vice versa (Harvey, 1993). Thus, when the time series is very smooth, aggregation doesn't add any value and by moving from a positive ϕ toward negative values of ϕ , the forecasts of aggregated series are more accurate than those related to the basic series.

5 EMPIRICAL ANALYSIS

The objective of our empirical analysis is to assess the validity of the main theoretical findings of this research. In the following sections we describe the experimental structure and we provide details of the empirical data available for the purposes of our investigation. The demand dataset available for the purposes of our investigation consists of weekly sales data over a period of two years for 1798 stock keeping units (SKUs) and comes from a European grocery store. The Time Series Expert Modeling function of SPSS (version 19) has been used to identify the ARIMA demand process for each series and estimate the relevant parameters. It was found that more than 30% of the series may be represented by the processes considered in our research. All SKUs have been categorized to different intervals according to the value of the process parameter ϕ , this categorization allows us to compare the empirical results with theoretical results. We see that the ϕ parameter estimations are almost positive, so the data set presented in this study don't cover whole range of ϕ which is between -1 and +1.

Autoregressive parameter interval	Average of autoregressive parameter	number of SKUs
$[-.1,0[$	-0.20	1
$[0,.1[$	0.07	3
$[.1,.2[$	0.16	12
$[.2,.3[$	0.26	43
$[.3,.4[$	0.36	120
$[.4,.5[$	0.45	149
$[.5,.6[$	0.55	142
$[.6,.7[$	0.63	60
$[.7,.8[$	0.75	22
$[.8,.9[$	0.82	3

Table 2 Empirical Data

5.1 Simulation details

The available data series are divided into two parts. The first part consists of 41 time periods and is used in order to initialize the SES estimates. The second part consists of the remaining 62 time periods and is used for the evaluation of the performance (out-of-sample).

The smoothing parameters for the SES method are not optimized and the values of 0.05 to 0.95 have been considered with a step increase of 0.05. In the classical approach, we first calculate 62 one-step ahead forecasts for each series and then we calculate the variance of the forecast error.

In order to get forecasts via the aggregation approach we start creating buckets of aggregated data based on the

aggregation level and then we apply SES to these aggregated data.

Aggregation level = 2: Starting from the 62th monthly observation, we sum observations backwards in buckets of two (2), resulting in a bi-monthly series consisting of 31 aggregate observations. Subsequently, we create a one-step- ahead bi-monthly forecast using the SES method for each series.

Aggregation level = 3 . . . 12: Similarly, we continue with time buckets of up to 12 periods. At this level there are only 5 aggregate yearly observations, so only 2 monthly observations at the start of the series remain unused.

After getting all aggregated SKUs and their forecasts, we calculate the variance of the forecast error for the aggregated data.

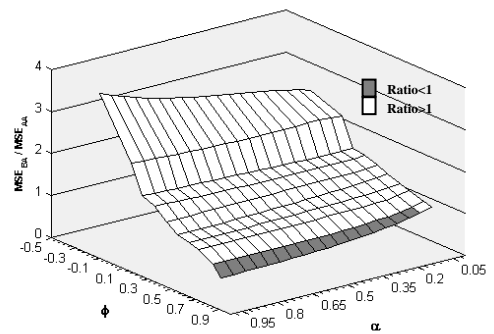
In order to estimate the m period demand level at each time period, the basic forecast is simply multiplied by m . Similarly the variance of the forecast error before aggregation is multiplied by m^2 .

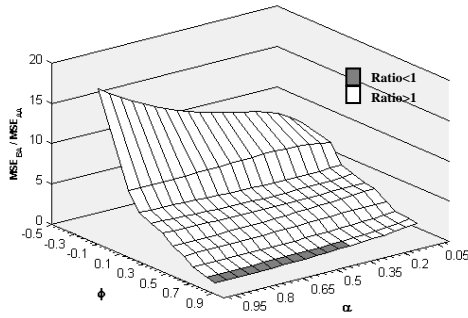
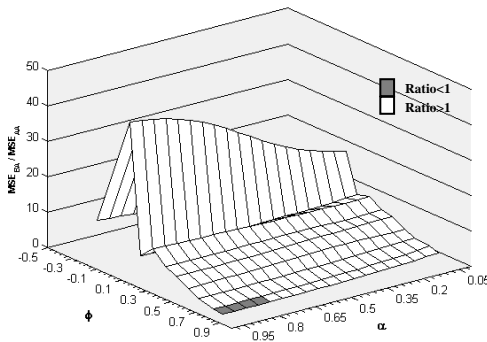
Finally the value of the variance of forecast error before aggregation (multiplied by m^2) is divided by the variance of the forecast error after aggregation, to get the ratio of MSE_{BA} to MSE_{AA} .

5.2 Empirical results

In section 4.4 we analytically showed the conditions under which aggregated forecasts are more accurate than non-aggregate or basic forecasts using the ratio of MSE_{BA} to MSE_{AA} . Now in the following figures we show the results of the empirical investigation.

These results indicate the superiority of each approach according to the control parameters and aggregation level. The results show that basic forecasts are associated with a better performance than the temporally aggregated forecasts when the autoregressive parameter gets high positive values, $0.6 \leq \phi < 1$, but the superiority of each approach is a function of ϕ and α . In other words when the autoregressive parameter takes a value $-1 < \phi < 0.6$, aggregated forecasts is preferable.

Figure 4. Empirical results for the ratio of MSE, $m=2$

Figure 5. Empirical results for the ratio of MSE, $m=7$ Figure 6. Empirical results for the ratio of MSE $m=12$

Similarly when $m \rightarrow \infty$ the ratio of MSE_{BA} to MSE_{AA} increases and it's always greater than one, which indicates that aggregated forecasts are more accurate. Comparing these empirical results with the analytical results reveals the same thing: aggregation doesn't add value for highly positive ϕ and then we observe that by increasing aggregation level, the area in which aggregated approach works better becomes larger as compared to classical approach, showing that aggregated approach is preferable to classical one as shown in Figure 4, Figure 5 and Figure 6.

6 CONCLUSIONS AND FUTURE RESEARCH

Aggregation is an appealing strategy to reduce the variability in both fast and slow moving items. Moreover, most inventory forecasting software packages support aggregation of data. Although this would typically cover cross-sectional aggregation (i.e. aggregation across series), the consideration of temporal aggregation has been neglected in these packages. In this paper we have analytically evaluated the effectiveness of the temporal aggregation strategy on forecasting when the basic series follow a first-order univariate autoregressive [AR(1)] process.

This paper focuses on the identification of the conditions under which the mean square error of the aggregation forecasting strategy would be lower than the non-aggregation one.

Single exponential smoothing is used as the forecasting technique under both basic and aggregated strategies. We analytically obtained the ratio of the mean square error before aggregation to that after aggregation, i.e. MSE_{BA}/MSE_{AA} and then based on certain values of the aggregation level, the autoregressive parameter and the smoothing constant, the superiority of each approach has been determined. We have then performed an investigation to empirically validate our theoretical results. We have done so by considering 555 SKUs from a European grocery store. The analytical and empirical results indicate that overall the aggregation approach improves forecast accuracy. A forecast variance reduction has been demonstrated through the consideration of the MSE metric. The main findings of this study are the following:

- The ratio, MSE_{BA}/MSE_{AA} , is a function of m , ϕ and α .
- For the autoregressive parameter $-1 < \phi \leq 0.5$, aggregation is preferable.
- For autoregressive parameter, $0.5 < \phi < 1$, the superiority of each approach is a function of m and ϕ . For a very high positive ϕ and high smoothing constants α , the classical approach is preferable.
- By increasing the aggregation level, the region of superiority of the aggregation approach as compared to the classical one becomes larger, and when $m \rightarrow \infty$, the aggregation approach is always preferable.

The future research should focus on theoretical analysis of temporal aggregation when the basic series follow other stationary and non-stationary ARMA and ARIMA processes. The interaction between temporal and cross-sectional forecasting is also an exciting area of research. In this study we have considered the case of non-overlapping aggregation; another interesting area is to consider the analytical and empirical evaluation of overlapping aggregation cases.

REFERENCE

- Aigner, D.J and Goldfeld, S.M ,1974. 'Estimation and Prediction from Aggregate Data when Aggregates are Measured More Accurately than Their Components', *Econometrica*, 42 (1), p.113-34.
- Amemiya T. and Wu R.Y, 1972. 'The effect of aggregation on prediction in the autoregressive model', *Journal of the American Statistical Association*, 67 (339), p.628-32.
- Armstrong, J., 2001. *Principles of forecasting: a handbook for researchers and practitioners*, ed. USA Springer Science and Business Media Inc.
- Athanasopoulos, G., Hyndman R. J., Song H., Wu D. C., 2009. 'The tourism forecasting competition',

- Working Paper 10/08 (Monash University,Australia).', Journal of the Operational Research Society, 62 (3), p.544-54.
- Babai, M.Z., Ali, M.M., and Nikolopoulos, K., 2012. 'Impact of Temporal Aggregation on Stock Control Performance of Intermittent Demand Estimators: Empirical Analysis', OMEGA(paper submitted for peer review).
- Chen, A., and Blue, J., 2010. 'Performance analysis of demand planning approaches for aggregating, forecasting and disaggregating interrelated demands', International Journal of Production Economics, 128 (2), p.586-602.
- Flidner, G., 1999. 'An Investigation of Aggregate Variable Time Series Forecast Strategies with Specific Subaggregate Time Series Statistical Correlation', Computers and Operations Research, 26, p.1133-49.
- Flidner, G., 2001. 'Hierarchical forecasting: issues and use guidelines', Industrial Management & Data Systems, 101 (1), p.5-12.
- Harvey, A.C., 1993, Time Series Models, ed. 2nd Edition (MIT Press).
- Lapide, L., 1998. 'New developments in business forecasting', Journal of Business Forecasting, p.28-29.
- Luna, I. and Ballini, R., 2011. 'Top-down strategies based on adaptive fuzzy rule-based systems for daily time series forecasting', International Journal of Forecasting, 27 (3), p.708-24.
- Lütkepohl, H., 1984. 'Forecasting contemporaneously aggregated vector ARMA processes', Journal of Business & Economic Statistics, 2 (3), p.201-14.
- Nikolopoulos, K., A.A. Syntetos, J.E. Boylan, F. Petropoulos, V. Assimakopoulos., 2011. 'An aggregate-disaggregate intermittent demand approach (ADIDA) to forecasting : an empirical proposition and analysis', Journal of the Operational Research Society, 62 (3), p.544-54.
- Rose, D.E., 1977. 'Forecasting aggregates of independent Arima processes', Journal of Econometrics, 5 (3), p.323-45.
- Rostami Tabar B., M.Z. Babai, A.A. Syntetos, Y. Ducq., 2012(Februry). ' The impact of temporal aggregation on demand forecasting' Seventeenth International Working Seminar on Production Economics,Innsbruck, Austria.
- Souza, L.R. and Smith, J., 2004. 'Effects of temporal aggregation on estimates and forecasts of fractionally integrated processes: a Monte-Carlo study', International Journal of Forecasting, 20 (3), p.487-502.
- Theil, H., 1954. *Linear Aggregation of Economic Relations* (Amsterdam: North-Holland Publishing Company).
- Wei, William W. S., 2006. *Time Series Analysis: Univariate and Multivariate Methods*, ed. Second Edition (Addison-Wesley Publishing Company).
- Widiarta, H., Viswanathan, S., and Piplani, R., 2007. 'On the effectiveness of top-down strategy for forecasting autoregressive demands', Naval Research Logistics, 54 (2), p.176-88.
- Widiarta, H., Viswanathan, S., and Piplani, R., 2009. 'Forecasting aggregate demand: An analytical evaluation of top-down versus bottom-up forecasting in a production planning framework', International Journal of Production Economics, 118 (1), 87-94.
- Willemain, T.R., Smart, C. N., Shockor, J. H., DeSautels, P. A. 1994. 'Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method', International Journal of Forecasting, 10 (4), p.529-38.
- Yehuda, G. and Griliches, Z., 1960. 'Is aggregation necessarily bad?', Review of Economics and Statistics, 42 (1), p.1-13.
- Zotteri, G. and Verganti, R., 2001. 'Multi-level approaches to demand management in complex environments: an analytical model', International Journal of Production Economics, 71 (1-3), p.221-33.
- Zotteri, G., Kalchschmidt, M., and Caniato, F., 2005. 'The impact of aggregation level on forecasting performance', International Journal of Production Economics, 93-94, p.479-91.