On the performance of temporal demand aggregation when optimal forecasting is used

Bahman Rostami-Tabar [1], M. Zied Babai [2], Mohammad Ali [1], John E. Boylan [3]

[1] Coventry University, Coventry, United Kingdom
[2] Kedge Business School, Talence, France
[3] Lancaster University, Lancaster, United Kingdom
{ac1055@coventry.ac.uk, mohamed-zied.babai@kedgebs.com, ab5925@coventry.ac.uk, j.boylan@lancaster.ac.uk}

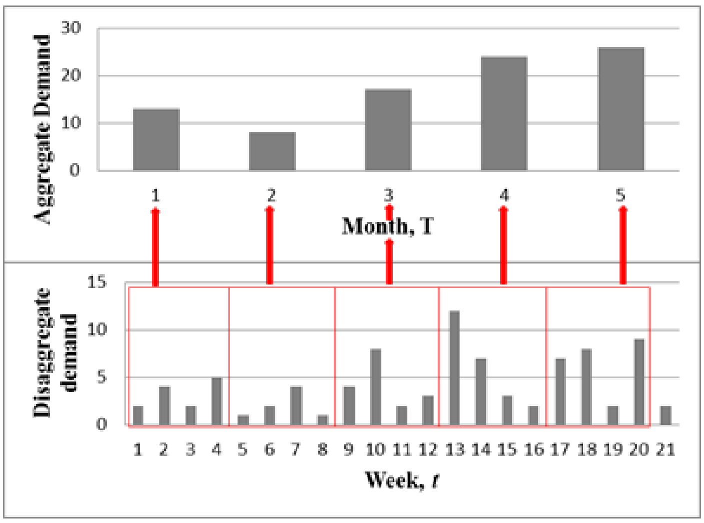
Abstract. Earlier research has shown that non-overlapping temporal aggregation of auto-correlated demand can improve the forecast accuracy of single exponential smoothing, especially for negative or low positive autocorrelation parameter. In this paper, we analyse the impact of non-overlapping temporal aggregation when an optimal forecasting method is used. We consider an AR(1) demand process and a minimum mean square error (MMSE) forecasting method. The expressions of the mean square error (MSE) before and after the aggregation of the demand are derived. The numerical results of the comparison of the MSEs show that by using the optimal MMSE forecasting method, regardless of the aggregation level and the autocorrelation parameter, the non-overlapping temporal aggregation approach is outperformed by the non-aggregation one.

Keywords: Demand aggregation, Forecast accuracy, Mean Square Error.

1 Introduction

Demand uncertainty is among the most important challenges facing modern companies [4]. The existence of high variability in demand for fast moving and intermittent moving items pose considerable difficulties in terms of forecasting and stock control. There are many approaches that may be used to reduce demand uncertainty and thus to improve the forecasting (and inventory control) performance of a company. An intuitively appealing such approach that is known to be effective is demand aggregation [5].
Aggregation across time or temporal aggregation refers to the process by which a low frequency time series (e.g. quarterly) is derived from a high frequency time series (e.g. monthly) [6]. This is achieved through the summation (bucketing) of every m periods of the high frequency data, where m is called the aggregation level. There are two different types of temporal aggregation: non-overlapping and overlapping. In the former case the time series are divided into consecutive non-overlapping buckets of time where the length of the time bucket equals the aggregation level. As shown in Figure 1, the aggregate demand is created by summing up the values inside each bucket. The number of aggregate periods is $[N/m]$, where N is the number of the original periods, m the aggregation level and the $[x]$ operator returns the integer part of x. As a consequence the number of periods in the aggregate demand is less than the original demands.

Figure 1: Non-overlapping temporal aggregation



The overlapping case is similar to a moving window technique where the window's size equals to the aggregation level. At each period, the window is moved one step ahead, so the oldest observation is dropped and the newest is included. It is observed that the number of overlapping aggregate periods is higher than those of the non-overlapping and equals to *N-m+1*. Therefore, the information loss is negligible as compared to the non-overlapping case. This is an important observation in terms of data availability and for the cases where little history of data is available.
The literature that deals with the impact of temporal aggregation on demand forecasting has grown during the last decade. Rostami-Tabar et al. [7] have shown that for an auto-correlated demand of order 1, AR(1), when the Single Exponential Smoothing (SES) is used to forecast the demand, the benefit of using non-overlapping temporal aggregation on the forecast accuracy depends on the autocorrelation parameter, the aggregation level and the smoothing constant. It has also been shown that for high levels of positive autocorrelation in the original series the aggregation approach is outperformed by the non-aggregation one.
In this paper, our objective is to analyse the impact of non-overlapping temporal aggregation when an optimal forecasting method is used. To do so, we consider an AR(1) demand process and we calculate the mean square error (MSE) of aggregate demand by considering the minimum mean square error (MMSE) forecast method. We then compare the mean square error obtained when aggregate demand is used to that of non-aggregate demand.
The rest of this paper is organised as follows. Section 2 briefly reviews the literature that deals with temporal aggregation. Section 3 presents the analytical derivations of the MSEs before and after aggregation. We show some numerical results in Section 4. The paper concludes in Section 6 with a summary of the findings and directions for future research.

2 Research Background

The analysis of temporal aggregation has started with the work of [1]. It has then been followed by a considerable research work ([3], [11], [10], [2]). Most of the literature has considered demands modelled with AutoRegressive Integrated Moving Average (ARIMA) type processes. They have characterised the aggregated demand processes (non-overlapping temporal aggregation in most of the cases) and they have shown that the aggregation approach results generally in an improvement of the forecast accuracy. The main limitations of this literature is that the forecasting methods and the performances measures often considered are theoretically advanced but not attractive from the practical perspective. To overcome these limitations, more recently, [7] and [8], have shown that for auto-correlated demand, AR(1) and ARMA(1,1), when the Single Exponential Smoothing (SES) is used to forecast the demand, the benefit of using non-overlapping temporal aggregation on the forecast accuracy depends on the autocorrelation parameter, the aggregation level and the smoothing constant. In fact, the performance of aggregation was, generally, found to improve as the aggregation level increases. Furthermore, it has been shown that for high levels of positive autocorrelation in the original series the aggregation approach is outperformed by the non-aggregation one. This is an intuitive finding since at any time the most recent demand information is so precious in that case that the disaggregate process works better as it fully exploits such recent information. However, in contrast, for low positive autocorrelation when the recent demand information is not that crucial then a more long term view on demand is preferable, which can be obtained as discussed above by selecting high aggregation levels (or also low smoothing constant). Finally, the literature shows that for optimal values of the smoothing constant of the non-aggregated process and a highly positive auto-correlation parameter, the aggregation doesn't work.

3 Analytical MSE

3.1 Notation and assumptions

For the remainder of the paper, we denote by:

m: Aggregation level, i.e. number of periods considered to build the block of aggregated demand.
n: total number of periods available in the demand history.
t: Time unit in the original non-aggregated time series. $t=1,2,\dots,n$.
T: Time unit in the aggregated time series. $T=1,2,\dots,\lfloor n/m \rfloor$.
d_t: Non-aggregated demand in period t
$D_{T,1}=d_{m,t}$: Aggregated demand in period T
ε_t : Independent random variables for non-aggregated demand in period t, normally distributed with zero mean and variance σ^2

ε'_T : Independent random variables for aggregated demand in period T, normally distributed with zero mean and variance σ'^2
$f_{t,m}$: Forecast of non-aggregated demand in period t, the forecast produced in $t-1$ for the demand over forecast horizon m.
$F_{T,1}$: Forecast of aggregated demand in period T, the forecast produced in T-1 for the demand in T.
MSE_{BA} : Theoretical Mean Squared Error (*MSE*) resulted from nonaggregated data by MMSE method.
MSE_{AA} : Theoretical Mean Squared Error (*MSE*) resulted from aggregate data by MMSE method.
γ_k : Covariance of lag k of non-aggregated demand, $\gamma_k = Cov(d_t, d_{t-k})$
γ'_k : Covariance of lag k of aggregated demand, $\gamma'_k = Cov(D_T, D_{T-k})$
ϕ : Autoregressive parameter of disaggregate demand process, $|\phi| < 1$
ϕ' : Autoregressive parameter of temporally aggregated demand process, $|\phi'| < 1$
C : Constant value of disaggregate demand in any time period
C' : Constant value of aggregated demand in any time period

We assume that the non-aggregated demand series d_t follows a first order autoregressive data generation process, AR(1) that can be mathematically written in period t by (1).

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

The m periods non-overlapping aggregated demand $d_{t,m} = D_{T,1}$ can be expressed as a function of the non-aggregated demand series as follows

$$d_{t,m} = D_{T,1} = \sum_{l=1}^{m} d_{t+(l-1)} \quad . \quad (2)$$

The forecasting method considered in this study is the conditional expectation that provides the Minimum Mean Squared Error (MMSE) unbiased forecast. Using MMSE, the forecast of demand at period t by knowing the demands d_{t-1}, d_{t-2},..., d_1 is:

$$f_{t,m} = E(d_{t,m} | d_{t-1}) \quad (3)$$

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

In this section we derive the *MSE* of the aggregate forecasts by considering the non-aggregated and the aggregated demand. Comparisons are to be performed at the aggregate level; to that end, the aggregation approach works as follows: firstly buckets of aggregated demand are created based on the aggregation level; then MMSE are applied to these aggregated data to produce the aggregate forecasts.
In this study the *MSE* is used as a forecast accuracy measure as it is the only theoretically tractable such measure. For each process under consideration we calculate the MSE based on disaggregate and aggregate demand data by considering MMSE forecasting method.

3.2 MSE before aggregation

We begin the analysis by deriving the *MSE* of aggregate forecast for the AR(1) process by using the disaggregate demand data. It is known that when demand follows an AR(1) process the following properties hold [12]:

$$\gamma_k = \begin{cases} \dfrac{\sigma^2}{1-\phi^2} & k = 0 \\ \phi^k \gamma_0 & k \geq 1 \end{cases}, \quad (4)$$

Now we derive the MSE of aggregate forecast resulted from the disaggregate data for the optimal method:

$$MSE_{BA} = Var(Forecast\ Error) = Var(d_{t,m} - f_{t,m}). \quad (5)$$

The aggregate forecast at period T when the subaggregate item follows an AR(1) process is:

$$f_{t,m} = E\left(d_{t,m} \middle| d_{t-1}\right) = E\left(d_t + d_{t+1} + \ldots + d_{t+m-1} \middle| d_{t-1}\right) =$$
$$E\left(\begin{matrix} C + \phi d_{t-1} + C + C\phi + \phi^2 d_{t-1} + C + C\phi + C\phi^2 + \phi^3 d_{t-1} + \cdots + \\ C + C\phi + C\phi^2 + C\phi^3 + \cdots + C\phi^{m-1} + \phi^m d_{t-1} + +\varepsilon_{t+m-1} + (1+\phi)\varepsilon_{t+m-2} \\ + \cdots + \left(1 + \phi + \phi^2 + \cdots + \phi^{m-1}\right)\varepsilon_t \end{matrix} \middle| d_{t-1}\right)$$
$$= E\left(\begin{matrix} C + C(1+\phi) + C\left(1+\phi+\phi^2+\phi^3\right) + \cdots + \\ C + C\phi + C\phi^2 + C\phi^3 + \cdots + C\phi^{m-1} + \phi d_{t-1} + \phi^2 d_{t-1} + \phi^3 d_{t-1} + \\ \cdots + \phi^m d_{t-1} + \varepsilon_{t+m-1} + (1+\phi)\varepsilon_{t+m-2} + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)\varepsilon_t \end{matrix} \middle| d_{t-1}\right) \tag{6}$$
$$= E\left(\begin{matrix} C + C(1+\phi) + +C\left(1+\phi+\phi^2+\phi^3\right) + \cdots + \\ C\left(1+\phi+\phi^2+\phi^3+\cdots+\phi^{m-1}\right) + \phi d_{t-1}\left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right) + \\ + \varepsilon_{t+m-1} + (1+\phi)\varepsilon_{t+m-2} + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)\varepsilon_t \end{matrix} \middle| d_{t-1}\right)$$
$$= \frac{C}{(1-\phi)}\left(m - \sum_{i=1}^{m}\phi^i\right) + \frac{\phi(1-\phi^m)}{(1-\phi)} d_{t-1}$$

By substituting $d_{t,m} = \sum_{l=1}^{m} d_{t+l-1}$ and (6) into **Erreur ! Source du renvoi introuvable.** and then substituting (4) into that we get the MSE_{BA}:

$$MSE_{BA} = Var\left(d_{t,m} - f_{t,m}\right) = Var\left(\begin{matrix} C + C(1+\phi) + +C\left(1+\phi+\phi^2+\phi^3\right) + \cdots + \\ C\left(1+\phi+\phi^2+\phi^3+\cdots+\phi^{m-1}\right) + \phi d_{t-1}\left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right) + \\ + \varepsilon_{t+m-1} + (1+\phi)\varepsilon_{t+m-2} + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)\varepsilon_t - \frac{C}{(1-\phi)}\left(m - \sum_{i=1}^{m}\phi^i\right) + \frac{\phi(1-\phi^m)}{(1-\phi)} d_{t-1} \end{matrix}\right)$$

$$= Var\left(\begin{matrix} C + C(1+\phi) + +C\left(1+\phi+\phi^2+\phi^3\right) + \cdots + \\ C\left(1+\phi+\phi^2+\phi^3+\cdots+\phi^{m-1}\right) + \frac{\phi(1-\phi^m)}{(1-\phi)} d_{t-1} + \\ + \varepsilon_{t+m-1} + (1+\phi)\varepsilon_{t+m-2} + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)\varepsilon_t - \frac{C}{(1-\phi)}\left(m - \sum_{i=1}^{m}\phi^i\right) - \frac{\phi(1-\phi^m)}{(1-\phi)} d_{t-1} \end{matrix}\right)$$

$$= Var\left(\begin{matrix} C + C(1+\phi) + +C\left(1+\phi+\phi^2+\phi^3\right) + \cdots + \\ C\left(1+\phi+\phi^2+\phi^3+\cdots+\phi^{m-1}\right) + \\ + \varepsilon_{t+m-1} + (1+\phi)\varepsilon_{t+m-2} + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)\varepsilon_t - \frac{C}{(1-\phi)}\left(m - \sum_{i=1}^{m}\phi^i\right) \end{matrix}\right)$$

$$= \left(\sigma^2 + (1+\phi)^2\sigma^2 + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)^2\sigma^2\right) = \sigma^2\left(1 + (1+\phi)^2 + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)^2\right)$$

$$= \sigma^2\left(1 + (1+\phi)^2 + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)^2\right) = \frac{\sigma^2(1-\phi)^2}{(1-\phi)^2}\left(1 + (1+\phi)^2 + \cdots + \left(1+\phi+\phi^2+\cdots+\phi^{m-1}\right)^2\right)$$

$$== \frac{\sigma^2}{(1-\phi)^2}\left((1-\phi)^2 + \left(1-\phi^2\right)^2 + \cdots + \left(1-\phi^m\right)^2\right) = \frac{\sigma^2}{(1-\phi)^2}\sum_{i=1}^{m}\left(1-\phi^i\right)^2$$

Hence, the MSE before aggregation is given by:

$$MSE_{BA} = \frac{\sigma^2}{(1-\phi)^2}\sum_{i=1}^{m}\left(1-\phi^i\right)^2 \tag{7}$$

3.2 MSE after aggregation

In this section we proceed with the derivation of the *MSE* of the aggregate forecasts resulted from the aggregate demand data is defined as

$$MSE_{AA} = Var(D_{T,1} - F_{T,1}) = Var(D_{T,1}) + Var(F_{T,1}) - 2Cov(D_{T,1}, F_{T,1}), \tag{8}$$

Demand is first aggregated to produce high frequency demand, then based on *MMSE* method we provide the aggregate forecasts. By applying MMSE method the aggregate forecast for period *T* is defined as:

$$F_{T,1} = E\left(D_{T,1} \middle| D_{T-1,1}, D_{T-2,1}, \dots, D_{1,1}\right) \tag{9}$$

If the non-aggregated series follows an AR(1) process then the aggregated series follows an ARMA(1,1) process ([1]; [9]). It can be shown that the following properties hold when the process is ARMA(1,1):

$$D_{T,1} = \mu'(1-\phi') + \varepsilon'_{T,1} + \phi' D_{T-1,1} - \theta'\varepsilon'_{T-1,1}, \; where \; |\theta'| < 1, |\phi'| < 1 \;, \tag{10}$$

$$\gamma'_k = \begin{cases} \dfrac{1 - 2\phi'\theta' + \theta'^2}{1-\phi'^2}\sigma'^2 & k = 0 \\ \dfrac{(\phi' - \theta')(1 - \phi'\theta')}{1-\phi'^2}\sigma'^2 & |k| = 1 \;. \\ \phi'\gamma'_{k-1} = \phi'^{k-1}\gamma'_1 & |k| > 1 \end{cases} \tag{11}$$

The aggregate forecast at period *T* by using MMSE method is:

$$F_{T,1} = E\left(D_{T,1} \middle| D_{T-1,1}\right) = \mu'(1-\phi') - \theta'\varepsilon'_{T-1,1} + \phi' D_{T-1,1}, \tag{12}$$

The variance of forecast and the covariance between aggregate demand and its forecast is:

$$Var(F_{T,1}) = Var\left(\mu'(1-\phi') - \theta'\varepsilon'_{T-1} + \phi' D_{T-1}\right) = \theta'^2\sigma'^2 + \phi'^2\gamma'_0 - 2\phi'\theta'\sigma'^2, \tag{13}$$

$$\begin{aligned} Cov(D_{T,1}, F_{,1T}) &= Cov\left(\mu'(1-\phi') + \varepsilon'_{T,1} + \phi' D_{T-1,1} - \theta'\varepsilon'_{T-1,1}, \mu'(1-\phi') - \theta'\varepsilon'_{T-1,1} + \phi' D_{T-1,1}\right) = -\phi'\theta' Cov\left(D_{T-1,1}, \varepsilon'_{T-1,1}\right) \\ &+ \phi'^2\gamma'_0 + \theta'^2\sigma'^2 - \phi'\theta' Cov\left(\varepsilon'_{T-1,1}, D_{T-1,1}\right) = \phi'^2\gamma'_0 + \theta'^2\sigma'^2 - 2\phi'\theta'\sigma'^2, \end{aligned} \tag{14}$$

By substituting $Var(D_{T,1}) = \gamma'_0$,(13), and (14) into (8) we get ;

$$\begin{aligned} MSE_{AA} &= \left(Var(D_{T,1}) + Var(F_{T,1}) - 2Cov(D_{T,1}, F_{T,1})\right) = \gamma'_0 + \theta'^2\sigma'^2 + \phi'^2\gamma'_0 - 2\phi'\theta'\sigma'^2 \\ &- 2\left(\phi'^2\gamma'_0 + \theta'^2\sigma'^2 - 2\phi'\theta'\sigma'^2\right) = \gamma'_0 - \left(\phi'^2\gamma'_0 + \theta'^2\sigma'^2 - 2\phi'\theta'\sigma'^2\right) = \\ &\gamma'_0(1-\phi'^2) - \theta'\sigma'^2(\theta' - 2\phi') = \frac{(1 - 2\phi'\theta' + \theta'^2)\sigma'^2}{1-\phi'^2}(1-\phi'^2) - \theta'^2\sigma'^2 + 2\phi'\theta'\sigma'^2 = \sigma'^2. \end{aligned} \tag{15}$$

Based on [12] we can show that the relationship between the parameters of the non-aggregated and the aggregated demand is

$$\gamma'_0 = \gamma_0\left(m + \sum_{k=1}^{m-1} 2(m-k)\phi^k\right), \tag{16}$$

$$\gamma'_1 = \gamma_0\left(\sum_{k=1}^{m} k\phi^k + \sum_{k=1}^{m-1} k\phi^{2m-k}\right), \tag{17}$$

and

$$\phi' = \phi^m \;. \tag{18}$$

Comparing γ'_0 / γ'_1 using (17), (16) and (11) we get:

$$\frac{\gamma'_0}{\gamma'_1} = \frac{\left(m + \sum_{k=1}^{m-1} 2(m-k)\phi^k\right)}{\left(\sum_{k=1}^{m} k\phi^k + \sum_{k=1}^{m-1} k\phi^{2m-k}\right)} = \frac{\dfrac{1 - 2\phi'\theta' + \theta'^2}{1-\phi'^2}}{\dfrac{(\phi' - \theta')(1 - \phi'\theta')}{1-\phi'^2}} \tag{19}$$

Therefore we get the following quadratic equation where we can calculate the moving average parameter after aggregation:

$$\left(1-\frac{\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)}{\left(\sum_{k=1}^{m}k\phi^k+\sum_{k=1}^{m-1}k\phi^{2m-k}\right)}\phi^m\right)\theta'^2+\left(\frac{\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)}{\left(\sum_{k=1}^{m}k\phi^k+\sum_{k=1}^{m-1}k\phi^{2m-k}\right)}+\frac{\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)}{\left(\sum_{k=1}^{m}k\phi^k+\sum_{k=1}^{m-1}k\phi^{2m-k}\right)}\phi^{2m}-2\phi^m\right)\theta'$$
$$+\left(1-\frac{\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)}{\left(\sum_{k=1}^{m}k\phi^k+\sum_{k=1}^{m-1}k\phi^{2m-k}\right)}\phi^m\right)=0, \tag{20}$$

By solving this equation we get:

$$\theta_1'=\frac{-\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)+\sqrt{\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)^2-4\left(1-\mathrm{X}\phi^m\right)^2}}{2\left(1-\mathrm{X}\phi^m\right)}\quad if\quad \phi>0 \tag{21}$$

$$\theta_2'=\frac{-\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)-\sqrt{\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)^2-4\left(1-\mathrm{X}\phi^m\right)^2}}{2\left(1-\mathrm{X}\phi^m\right)}\quad if\quad \phi<0 \tag{22}$$

where

$$\mathrm{X}=\frac{\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)}{\left(\sum_{k=1}^{m}k\phi^k+\sum_{k=1}^{m-1}k\phi^{2m-k}\right)}$$

By considering (11) and (4) we get:

$$\sigma'^2=\frac{\left(1-\phi^{2m}\right)\gamma_0\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)}{\left(\left(1-2\phi^m\theta'+\theta'^2\right)\right)}, \tag{2}$$

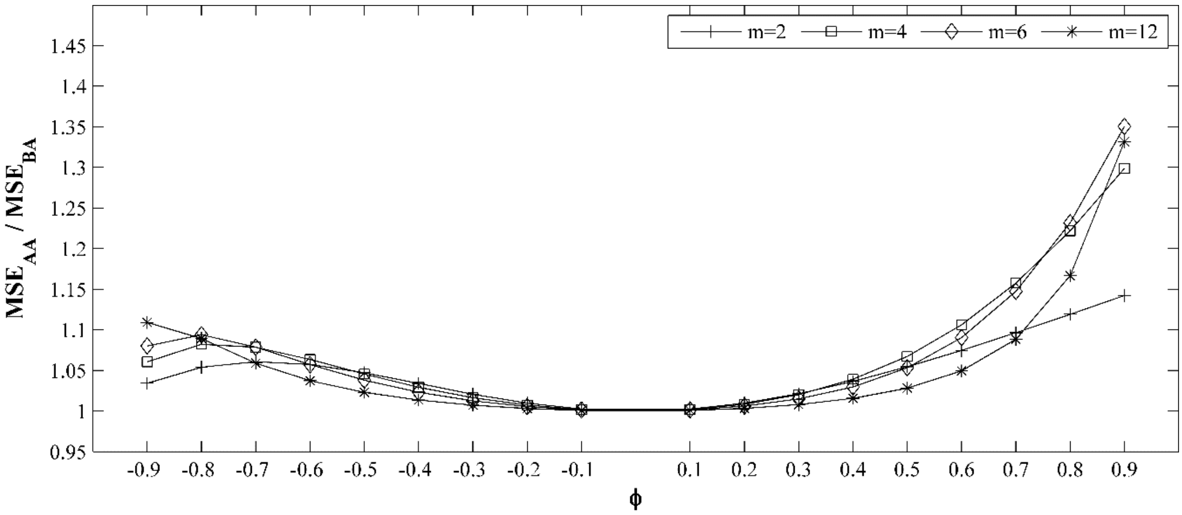
By substituting (21) and (22) in (2) we get:

$$MSE_{AA}=\sigma'^2=\begin{cases}\dfrac{\left(1-\phi^{2m}\right)\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)\sigma^2}{\left(1-\phi^2\right)\left(\left(1-2\phi^m\left(\dfrac{-\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)+\sqrt{\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)^2-4\left(1-\mathrm{X}\phi^m\right)^2}}{2\left(1-\mathrm{X}\phi^m\right)}\right)+\left(\dfrac{-\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)+\sqrt{\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)^2-4\left(1-\mathrm{X}\phi^m\right)^2}}{2\left(1-\mathrm{X}\phi^m\right)}\right)^2\right)\right)} & if \\[2ex] \dfrac{\left(1-\phi^{2m}\right)\left(m+\sum_{k=1}^{m-1}2(m-k)\phi^k\right)\sigma^2}{\left(1-\phi^2\right)\left(\left(1-2\phi^m\left(\dfrac{-\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)-\sqrt{\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)^2-4\left(1-\mathrm{X}\phi^m\right)^2}}{2\left(1-\mathrm{X}\phi^m\right)}\right)+\left(\dfrac{-\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)-\sqrt{\left(\mathrm{X}+\mathrm{X}\phi^{2m}-2\phi^m\right)^2-4\left(1-\mathrm{X}\phi^m\right)^2}}{2\left(1-\mathrm{X}\phi^m\right)}\right)^2\right)\right)} & if\end{cases} \tag{24}$$

4 Numerical results

In this section, we numerically analyse the ratio MSE_{AA} / MSE_{BA} for different values of the aggregation level *m* and the autocorrelation parameter ϕ . Figure 2 shows the ratio MSE_{AA} / MSE_{BA} for *m* = 2, 4, 6, 12 and ϕ between -1 and +1.

Figure 2: Ratio of MSEs after aggregation to before aggregation



According to Figure 2, when the data generating process (DGP) is known, for negative and less positive values of ϕ (negative autocorrelation close to zero and low positive autocorrelation) forecasting disaggregate data is at least as efficient, in terms of mean squared forecast error (MSFE), as directly forecasting the aggregated data. However, for higher values of ϕ (high positive autocorrelation), the differences between two approaches becomes more significant and forecasting disaggregate data is more accurate. It is shown that the aggregation level, m does not play a significant role in terms of MSE forecast error.
Furthermore, the numerical results in Figure 2 show that by using the optimal MMSE forecasting method, regardless of the aggregation level and the autocorrelation parameter, the ratio is always higher than 1. This means that the non-overlapping temporal aggregation doesn't work and the aggregation approach is outperformed by the non-aggregation one.

5 Conclusion

Non-overlapping temporal aggregation has been proven to be an interesting approach to reduce demand uncertainty and thus to improve the forecasting accuracy in the case of auto-correlated demand and single exponential smoothing. However, we have shown in this work that for an AR(1) demand, when the minimum mean square error forecasting method is used, the temporal aggregation approach is always outperformed by the non-aggregation one.
There are some interesting avenues for further research. The first one consists of expanding the analytical work discussed in this paper on higher order stationary processes and more importantly on non-stationary processes, since this is a very important issue both from an academic and practitioner perspective. Another avenue would be to conduct an empirical investigation to validate the findings of this paper.

References

[1] Amemiya , T., Wu, R.Y., 1972. The effect of aggregation on prediction in the autoregressive model. Journal of the American Statistical Association 67, 628-632.
[2] Athanasopoulos, G., Hyndman, R.J., Song, H., Wu, D.C., 2011. The tourism forecasting competition. International Journal of Forecasting 27, 822-844.
[3] Brewer, K.R.W., 1973. Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models. Journal of Econometrics 1, 133-154.
[4] Chen, A., Blue, J., 2010. Performance analysis of demand planning approaches for aggregating, forecasting and disaggregating interrelated demands. International Journal of Production Economics 128, 586-602.

[5] Chen, A., Hsu, C.H., Blue, J., 2007. Demand planning approaches to aggregating and forecasting interrelated demands for safety stock and backup capacity planning. International Journal of Production Research 45, 2269-2294.
[6] Nikolopoulos, K., Syntetos, A.A., Boylan, J., Petropoulos, F., Assimakopoulos, V., 2011. An aggregate-disaggregate intermittent demand approach (ADIDA) to forecasting: an empirical proposition and analysis. Journal of the Operational Research Society 62, 544-554.
[7] Rostami-Tabar, B., Babai, M.Z., Syntetos, A.A., Ducq, Y., (2013). Demand forecasting by temporal aggregation, *Naval Research Logistics*, Vol. 60(6), pp. 479-498.
[8] Rostami-Tabar, B., Babai, M.Z., Syntetos, A., Ducq, Y. (2014). A note on the forecast performance of temporal aggregation, *Naval Research Logistics*, Vol. 61, pp. 489-500.
[9] Sariaslan, N., 2010. The Effect of Temporal Aggregation on Univariate Time Series Analysis, Department of Statistics. Middle East Technical University, p. 150.
[10] Stram, D.O., Wei, W.W.S., 1986. Temporal aggregation in the ARIMA process. Journal of Time Series Analysis 7, 279-292.
[11] Tiao, G.C., 1972. Asymptotic behaviour of temporal aggregates of time series. Biometrika 59, 525-531.
[12] Wei, W.W.S., 2006. Time Series Analysis: Univariate and Multivariate Methods. Addison-Wesley Publishing Company, Boston, USA.