Evaluation of Bottom-up and Top-down Strategies for Aggregated Forecasts: State Space Models and ARIMA Applications

Milton Soto-Ferrari[1] [✉], Odette Chams-Anturi[2], Juan P. Escorcia-Caballero[3], Namra Hussain[1], and Muhammad Khan[1]

[1]Indiana State University, Terre Haute IN, USA
milton.soto-ferrari@indstate.edu
{nhussain1,mkhan12}@sycamores.indstate.edu
[2] Universidad de la Costa, Barranquilla, Colombia
ochams@cuc.edu.co
[3] Universidad del Norte, Barranquilla, Colombia
juane@uninorte.edu.co

Abstract. In this research, we consider monthly series from the M4 competition to study the relative performance of top-down and bottom-up strategies by means of implementing forecast automation of state space and ARIMA models. For the bottom-up strategy, the forecast for each series is developed individually and then these are combined to produce a cumulative forecast of the aggregated series. For the top-down strategy, the series or components values are first combined and then a single forecast is determined for the aggregated series. Based on our implementation, state space models showed a higher forecast performance when a top-down strategy is applied. ARIMA models had a higher forecast performance for the bottom-up strategy. For state space models the top-down strategy reduced the overall error significantly. ARIMA models showed to be more accurate when forecasts are first determined individually. As part of the development we also proposed an approach to improve the forecasting procedure of aggregation strategies.

Keywords: Top-down, Bottom-up, Forecast automation, Forecast performance, State space models, ARIMA.

1. Introduction

Selecting an appropriate forecasting method for a number of time series is a major concern when making decisions. At the organizational level, forecasts are required as critical inputs to many activities in various business areas such as inventory management, marketing, sales, finance, and accounting [1]. There is a frequent need in business for completely automatic forecasting methods (i.e., forecast automation) that takes into account series characteristics and other features of the data without the need for human interference [2]. Literature propose various selection rules in order to enhance forecasting accuracy. The simplest approach for model selection when evaluating multiples series, involves the identification of a single method which is applied over a combined series without taking into account the specifications of its own components [3]. The idea behind this approach is known as aggregation, where multiple series are combined into a single series without considering their individual specifications such as trend and seasonality [3–5]. When series are aggregated the overall variability of the combined series is reduced, which may result on a superior forecasting accuracy [5, 6]. In addition, the automation for model selection is simpler with a lower complexity but with the cost of losing the specifications from the individual series [3, 7]. On the other hand, individual selection involves the identification of the best method for each series, but this approach is more computationally intensive [4]. The inquiry in this context is to determine which approach would be more effective, in terms of performance, since small improvements in forecast accuracy can lead to large reductions in inventory and increase in service levels [8–10].

The research about the aggregation level of a forecasting process is referred in the literature as Hierarchical Forecasting [5, 11]. In this setting, two forecasting strategies are typically denoted: The bottom-up strategy (BU) and the top-down strategy (TD). In BU, the forecast is developed for each series individually and then these are combined to generate a cumulative forecast of the aggregated series. This is referred as the cumulative forecast, since it is made up by the combination of the individual forecasts of each series. In TD, series are first aggregated to produce a combined forecast, then the forecast is disaggregated and a derived forecast for each series is established usually by means of proportions. Research about the comparisons between TD and BU is available in [12–21], and the principal objective from the developments is to identify which strategy presents a higher forecast performance. However, the findings about whether TD strategies perform better than BU, or vice-versa, remain debatable. Therefore, improvement of forecast performance using these strategies are contemporary, especially if considerations about forecast automation are part of the analysis [4] since this is substantial when working with a large number of series.

Forecast automation is essential when modelling several time series. Automation methodologies for two of the most broadly forecasting methods, autoregressive integrated moving average (ARIMA) and state space models, are recognized to perform very well with several types of time series [2, 22–26]. More advanced or complex methods of forecasting include machine learning procedures such as: Bayesian neural networks, K-nearest neighbor regression, kernel regression, CART regression trees, and support vector regression [27]. The disadvantage with many machine learning algorithms is that often them appear as black boxes or infinite networks with limited and restricted insights into how the forecasts are produced and which data components are important. These attributes of forecasting are often critical for practitioners [28]. State space and ARIMA models are relatively simple but robust approaches to forecasting that are widely used in business with great success in both academic research, educational competitions, and industrial applications [23, 29, 30].

According to Weller and Crone [31] in a survey of forecasting practices, the exponential smoothing family of models is the most frequently used. Actually, it is implemented almost 1/3 of times (32.1%) in detriment of more advanced forecasting techniques that are only applied in 10% of cases. In general, simpler methods are used 3/4 of times, a result that is consistent with the relative accuracy of such methods in forecasting competitions.

Currently the M-Competitions, now in its four version, have attracted great interest in providing objective evidence of the most appropriate way of forecasting various variables of interest. In this research, the objective is to evaluate performance of BU and TD strategies using only forecast automation of state space and ARIMA models. We selected a set of time series from the M4 competition with the purpose of identify the most accurate forecasting method when implementing state space and ARIMA models in combination with the application of BU and TD strategies. Machine learning methods are not considered in this analysis, since the focus of this research is the implementation of forecast automation of the most widely used methods but in the context of aggregation and cumulative forecasts.

2. Forecast Automation

2.1 State Space Models

Since 1950, exponential smoothing methods have been applied with success to several types of time series [2]. The basic variations of exponential smoothing include: simple exponential smoothing, trend-corrected exponential smoothing or Holt's model, additive damped trend, and Holt-Winters additive and multiplicative methods that might include damped trend errors [32–34]. The usual description of these methods is the component form. Component form of exponential smoothing methods comprise a forecast equation and a smoothing equation for the components [24].
Hyndman et al [2] developed a statistical framework for all exponential smoothing methods. In this statistical structure each model, referred as state space model, consists of a measurement equation that describes the evaluated data, and state or transition equations that describe how the unobserved components or states (level, trend, seasonal) evolve over time. For illustration, let us denote the formulation for the component and state space form of the most common models.

Simple exponential smoothing (1)
Component form

$$\hat{y}_{t+h|t} = l_t$$
$$l_t = \alpha y_t + (1-\alpha) l_{t-1}$$

State space form

$$y_t = l_{t-1} + e_t$$
$$l_t = l_{t-1} + \alpha e_t$$

Where: $e_t = y_t - l_{t-1} = y_t - \hat{y}_{t|t-1}$ for t=1,…., T, the one-step within-sample forecast error at the time t. l_t is an unobserved state.

Additive damped trend (3)
Component form

$$\hat{y}_{t+h|t} = l_t + (\emptyset + \emptyset^2 + \cdots + \emptyset^h) b_t$$
$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + \emptyset b_{t-1})$$
$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)\emptyset b_{t-1}$$

State space form

$$y_t = l_{t-1} + \emptyset b_{t-1} + e_t$$
$$l_t = l_{t-1} + \emptyset b_{t-1} + \alpha e_t$$
$$b_t = \emptyset b_{t-1} + \beta e_t$$

Where: Damping parameter $0<\emptyset<1$.
If $\emptyset = 1$, indical to Holt`s linear trend
As h→∞, $\hat{y}_{T+h|T} \to l_t + \emptyset b_T/(1-\emptyset)$
Short-run forecasts trend, long-run forecasts constant.

Holt´s linear trend (2)
Component form

$$\hat{y}_{t+h|t} = l_t + h b_t$$
$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$$
$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*) b_{t-1}$$

State space form

$$y_t = l_{t-1} + b_{t-1} + e_t$$
$$l_t = l_{t-1} + b_{t-1} + \alpha e_t$$
$$b_t = b_{t-1} + \beta e_t$$

Where:
$\beta = \alpha\beta^*$
$e_t = y_t - (l_{t-1} + b_{t-1}) = y_t - \hat{y}_{t|t-1}$

Holt-winters additive method (4)
Component form

$$\hat{y}_{t+h|t} = l_t + h b_t + s_{t-m+h_m^+}$$
$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1})$$
$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*) b_{t-1}$$
$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma) s_{t-m},$$

State space form

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + e_t$$
$$l_t = l_{t-1} + b_{t-1} + \alpha e_t$$
$$b_t = b_{t-1} + \beta e_t$$
$$s_t = s_{t-m} + \gamma e_t$$

Holt-winters multiplicative (5)

Component form

$$\hat{y}_{t+h|t} = (l_t + hb_t)s_{t-m+h_m^+}$$
$$l_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(l_{t-1} + b_{t-1})$$
$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1}$$
$$s_t = \gamma \frac{y_t}{(l_{t-1} - b_{t-1})} + (1-\gamma)s_{t-m}$$

State space form

$$y_t = (l_{t-1} + b_{t-1})s_{t-m} + e_t$$
$$l_t = l_{t-1} + b_{t-1} + \alpha e_t / s_{t-m}$$
$$b_t = b_{t-1} + \beta e_t / s_{t-m}$$
$$s_t = s_{t-m} + \gamma e_t / (l_{t-1} + b_{t-1})$$

Holt-winters damped method (6)

Component form

$$\hat{y}_{t+h|t} = [l_t + (\emptyset + \emptyset^2 + \cdots + \emptyset^h)b_t]s_{t-m+h_m^+}$$
$$l_t = \alpha(y_t / s_{t-m}) + (1-\alpha)(l_{t-1} + \emptyset b_{t-1})$$
$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)\emptyset b_{t-1}$$
$$s_t = \gamma \frac{y_t}{(l_{t-1} + \emptyset b_{t-1})} + (1-\gamma)s_{t-m}$$

State space form

$$y_t = (l_{t-1} + \emptyset b_{t-1})s_{t-m} + e_t$$
$$l_t = l_{t-1} + \emptyset b_{t-1} + \alpha e_t / s_{t-m}$$
$$b_t = \emptyset b_{t-1} + \beta e_t / s_{t-m}$$
$$s_t = s_{t-m} + \gamma e_t / (l_{t-1} + \emptyset b_{t-1})$$

Where for all cases m denotes the period of seasonality, l_t denotes an estimate of the level of the series at time t, b_t denotes an estimate of the trend of the series at time t, s_t denotes an estimate of the seasonality of the series at time t. The initial states and the smoothing parameters α, β, γ are estimated from the observed data. The smoothing parameters α, β, γ are constrained between 0 and 1 with the purpose of that the equations can be interpreted as weighted averages [35].

For space state models, the letters (E, T, S) denote the three forecast components: “Error”, “Trend” and “Seasonality”. The notation ETS refers to a three-character string form identifying the method used by the framework terminology. For instance, the first letter denotes the error type ("A", "M" or "Z"); the second letter denotes the trend type ("N","A","M" or "Z"); and the third letter denotes the seasonal type ("N","A","M" or "Z"). In all cases, "N"=none, "A"=additive, "M"=multiplicative, "Z"=automatically selected, and "A_d" denotes additive damped. Then, for example, "ANN" is simple exponential smoothing with additive errors, and "MAM" is multiplicative Holt-Winters with multiplicative errors. The letter “Z” refers to forecast automation, where for the given the data, a state space model is identified automatically to optimize or minimize errors.

2.2 ARIMA

The class of ARIMA models is broad. It can represent many different types of stochastic seasonal and non-seasonal time series such as autoregressive (AR), moving average (MA), and mixed AR or MA processes, where the baseline might need to be differenced and integrated (I). Box-Jenkins et al. [36] developed a systematic and practical model building method. Using this process ARIMA follows three sequential phases; i) model identification: create and evaluate the correlograms of the series, their patterns enables the identification of the time series that is represented in the baseline, ii) model estimation: estimation of the parameter values, and iii) model diagnosis: development of preliminary forecasts, these forecasts are used to diagnose the identification and estimation stage.

The Box-Jenkins methodology has been proved as an effective and practical time series modeling approach. ARIMA models considers three parameters p, d, q that are represented as: ARIMA (*p,d,q*) Where: p denotes the model and forecast that are based in part or completely on autoregression. p is the number of autoregressive parameters in the model, d is the number of times the series has been differenced to achieve stationarity, and q is the number of moving average parameters in the model that accounts for random jumps in the time series.

For seasonal ARIMA: ARIMA(*p,d,q*)(P,D,Q), the uppercase letters have the same meaning as the lowercase letters, but these are referred to seasonal parameters. For illustration, let us denote the formulation for the multiplicative seasonal ARIMA model (*p,d,q*) x $(P,D,Q)_m$

$$\emptyset_p(B)\emptyset_P(B^m)(1-B)^d(1-B^m)^D y_t = c + \theta_q(B)\Theta_Q(B^m)\varepsilon_t \tag{7}$$

Where:

$$\emptyset_p(B) = 1 - \emptyset_1 B - \cdots - \emptyset_p B^p, \emptyset_P(B^m) = 1 - \emptyset_1 B^m - \cdots - \emptyset_P B^{Pm}$$
$$\theta_q(B) = 1 + \theta_1 B + \cdots + \theta_q B^q, \Theta_Q(B^m) = 1 + \Theta_1 B^m + \cdots - \Theta_Q B^{Qm}$$

With m as the seasonal frequency, B is the backward shift operator, d is the degree of ordinary differencing, and D is the degree of seasonal differencing, $\emptyset_p(B)$ and $\theta_q(B)$ are the regular autoregressive and moving average polynomials of orders p and q, respectively, $\Theta_Q(B^m)$ and $\Theta_Q B^{Qm}$ are the seasonal autoregressive and moving average polynomials of orders P and Q, respectively, $C = \mu(1 - \emptyset_1 - \cdots - \emptyset_p)(1 - \emptyset_1 - \cdots - \emptyset_p)$where μ is the mean of $(1-B)^d(1-B^m)^D y_t$ process and ε_t is a zero mean Gaussian white noise process with variance σ^2.

For this research, the state space methods are estimated using the forecast package for R statistical software described in [25]. The automatic ETS function (AUTO.ETS) is used to estimate the state space model form parameters. The ARIMA function (AUTO.ARIMA) implemented in the same package is also used to identify and estimate the ARIMA models. The AUTO.ARIMA function conducts a stepwise selection over possible models and returns the best ARIMA model. The algorithms are applicable to both seasonal and non-seasonal data, these are illustrated using series from the M4 competition, with the forecasting strategies for TD and BU.

3. Forecast Strategy Analysis

3.1 Data Selection

The research literature on TD versus BU strategies is generally characterized into two categories. The first category assumes that the statistical properties of the sub aggregated time series components are known perfectly. In this framework, both TD and BU forecasting would perform equally well, only when the components are uncorrelated and have identical stochastic structures [37, 38]. The second category, assumes that the generating process is not known a priori and data is constantly updated. When data is constantly updated TD could be developed with a higher efficacy since forecasting is done simultaneously for several different components [21, 39–41]. For our case, the series that are part of the study do not have identical stochastic structures, which is in accordance with most business processes. Since the structure for the series are different to each other, it is expected for the implementation a variation on performance when both strategies are applied to the selected series.

3.2 Data Description

We consider five monthly series from the M4 competition to study the relative forecast performance of TD and BU strategies when aggregated series are considered. In the BU strategy, the forecast for each series is determined individually and then a cumulative forecast is obtained by adding the individual components forecasts. We first determine the forecasts for each series individually and then we evaluate the errors for the cumulative forecast when compared with the aggregated series using AUTO.ETS and AUTO.ARIMA. In the TD strategy, the series are first combined to obtain the aggregated series and then a single forecast is determined from the aggregated series. Forecast performance is evaluated using the combined testing sets.

All series from the M4 competition are divided into training and testing sets. The selected series for this study are defined in the competition dataset as: M19, M20, M21, M22, and M23. The total number of observations for each of the series is equal to 192 months, 174 for the training and 18 for the testing set respectively. We assumed an ending period for all five series equal to December 2018 (12/2018). In order to evaluate forecast performance for the different strategies, RMSE, MAE, and MAPE are calculated for the testing set. Where:

(8)

$$\text{RMSE} = \sqrt{\frac{1}{n-m}\sum_{t=m+1}^{n}(y_t - \hat{y}_t)^2} \qquad \text{MAE} = \frac{1}{n-m}\sum_{t=m+1}^{n}\left|y_t - \hat{y}_t\right|$$

$$\text{MAPE} = \frac{1}{n-m}\sum_{t=m+1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \text{x100}$$

We used the training set to identify the best model for the series, using BU and TD strategies, with AUTO.ETS and AUTO.ARIMA. All evaluations of forecast performance are implemented over the testing set. Figure 1 presents the graphical representations of the selected series.

3.3 BU Strategy

For the BU strategy, we developed the forecast for each of the series individually with the forecast accuracy calculated for each component. Then, all forecasts are combined and the performance is evaluated with the aggregated values from the series. Performance efficiency is calculated using the aggregated testing set. The methodology is developed using AUTO.ETS and AUTO.ARIMA.

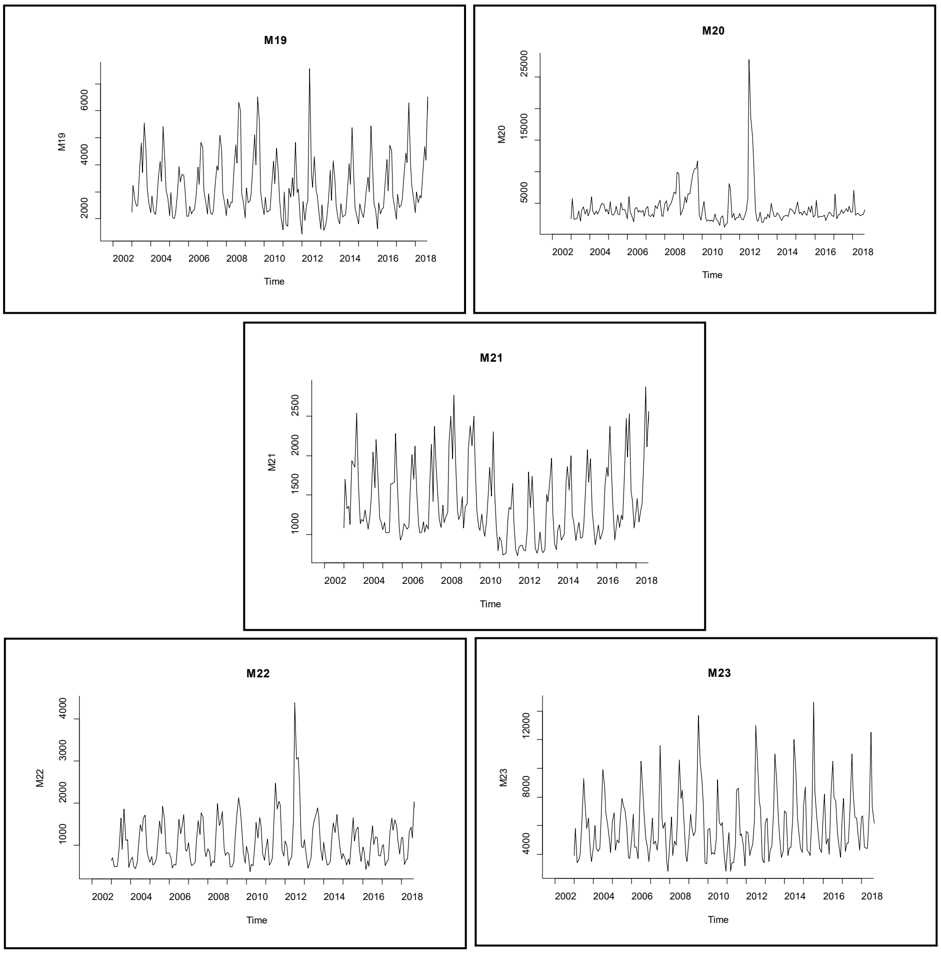


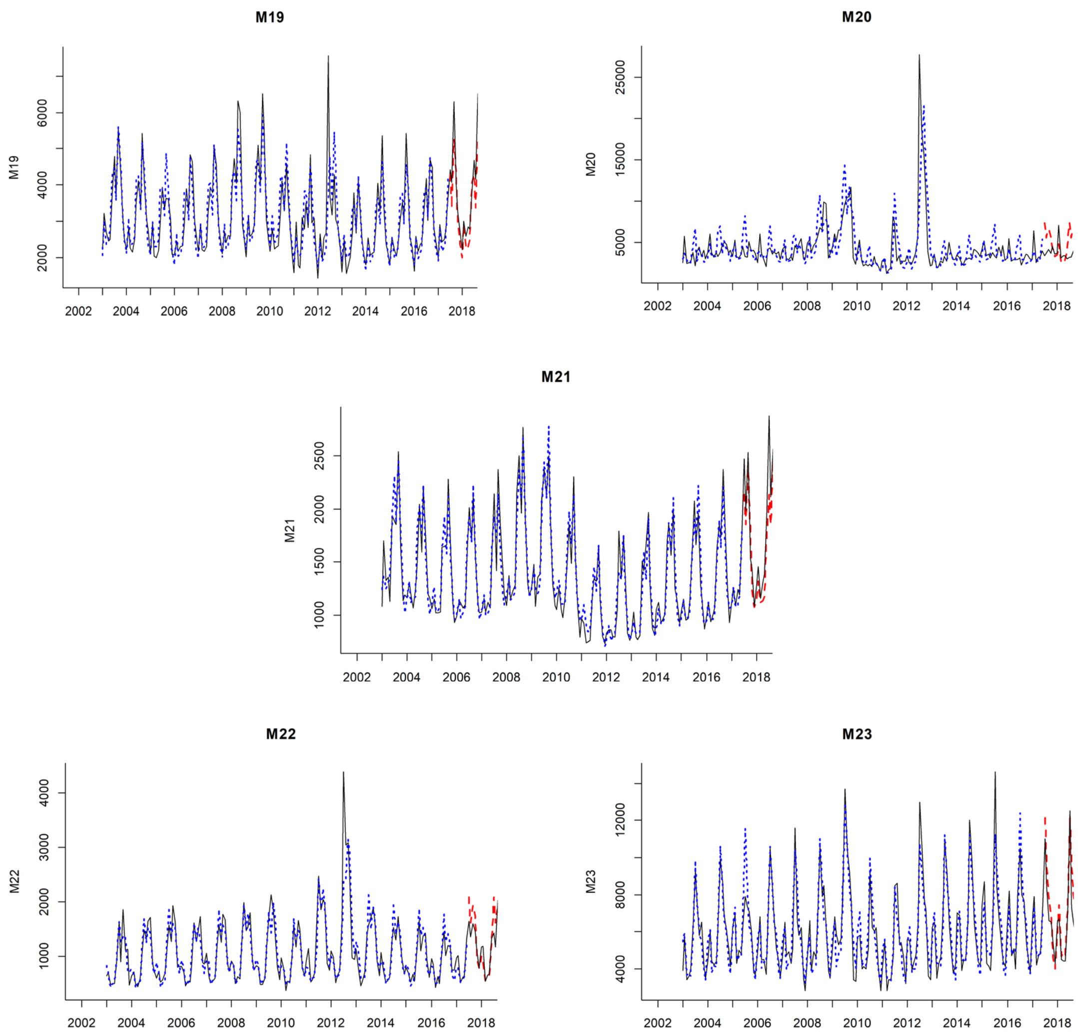
Figure 1: Graphical representation of selected series.

ETS - BU

We developed the forecasts for each of the series and then we combined them to evaluate forecast performance of the aggregated series with AUTO.ETS. Figure 2 shows the forecast and performance for each of the components. Figure 4 shows the forecast performance for the aggregated series.

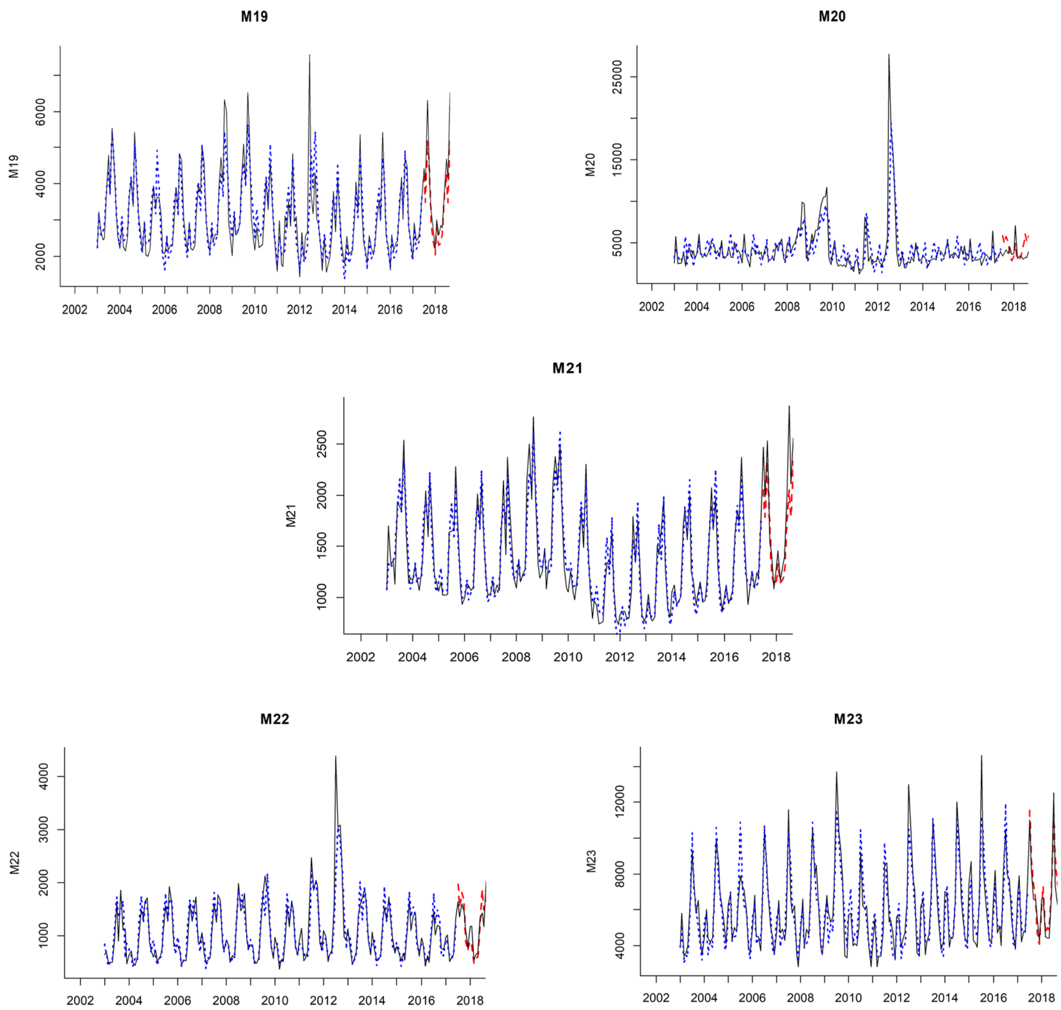
ARIMA - BU

Similar to the ETS-BU approach, we developed the forecasts for each of the series and then we combined them to evaluate forecast performance of the aggregated series with AUTO.ARIMA. Figure 3 shows the forecast and performance for each of the components. Figure 5 shows the forecast performance for the aggregated series.



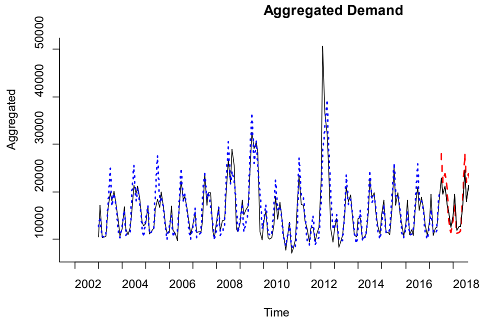
ETS – BU	RMSE	MAE	MAPE	AUTO ETS
M19	543.18	448.30	11.56	ETS(M, N, M)
M20	1979.29	1567.11	42.72	ETS(M, N, M)
M21	239.95	184.54	9.93	ETS(M, Ad, M)
M22	267.81	200.64	15.25	ETS(M, N, M)
M23	865.15	677.19	10.42	ETS(M, N, M)

Figure 2: ETS-BU individual forecasts



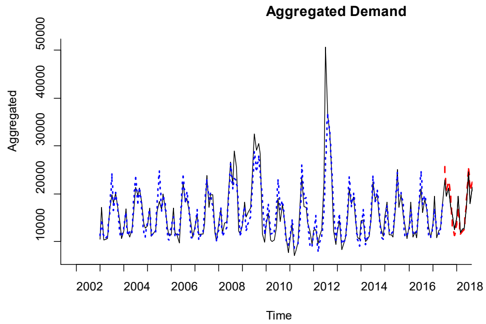
ARIMA – BU	RMSE	MAE	MAPE	AUTO ARIMA
M19	517.56	398.20	9.81	ARIMA(0,1,1)
M20	1443.18	1174.38	31.29	ARIMA(1,0,0) with non-zero mean
M21	272.69	206.89	10.55	ARIMA(1,1,1)
M22	230.56	202.13	16.76	ARIMA(1,0,0) with non-zero mean
M23	716.67	589.12	9.27	ARIMA(1,1,1)

Figure 3: ARIMA-BU individual forecasts



Aggregated ETS – BU	RMSE	MAE	MAPE
Cumulative Forecast	2368.23	2012.99	11.06

Figure 4: ETS- BU aggregated series forecast



Aggregated ARIMA – BU	RMSE	MAE	MAPE
Cumulative Forecast	1581.34	1344.11	7.66

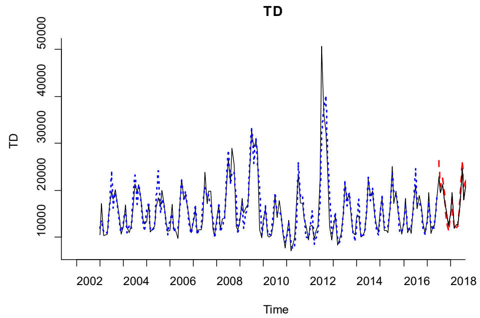
Figure 5: ARIMA- BU aggregated series forecast

3.4 TD Strategy

For the TD strategy, the series or components values are first combined and then a single forecast is determined for the aggregated series. Forecast performance is evaluated using the combined testing set. The methodology is developed using AUTO.ETS and AUTO.ARIMA.

ETS - TD
We first combined the series and then a single forecast is determined with AUTO.ETS. Figure 6 shows the forecast and performance for the aggregated series.

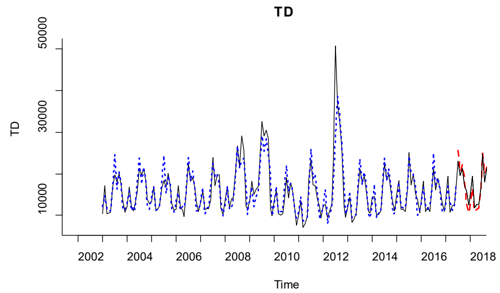
ARIMA - TD
Similar to the ETS-TD approach, we first combined the series and then single a forecast is determined with AUTO.ARIMA. Figure 7 shows the forecast and performance for the aggregated series.



ETS – TD	RMSE	MAE	MAPE
Aggregated Forecast	1548.71	1302.69	7.29

AUTO ETS
ETS(M, Ad, M)

Figure 6: ETS-TD aggregated series forecast



ARIMA – TD	RMSE	MAE	MAPE
Aggregated Forecast	1662.38	1419.60	8.86

AUTO ARIMA
ARIMA(1,0,0)with non-zero mean

Figure 7: ARIMA-TD aggregated series forecast

4. Discussion and Conclusions

AUTO.ETS showed a higher performance when a TD strategy is applied. On the other hand, AUTO.ARIMA had a greater performance for the BU strategy. For the ETS method, the TD strategy improved significantly the forecast when compared to BU. ARIMA proved to be more accurate when forecasts are determined individually. The best forecast for the study, given MAPE, is ETS with a TD strategy. Table 1 shows the overall results for the strategies.

Table 1. Overall results for the strategies

Strategy	MAPE
1) ETS-BU	11.06
2) ARIMA-BU	7.66
3) ETS-TD	7.29
4) ARIMA-TD	8.86

ETS-BU (Figure 2) forecasted significantly higher values for the periods in the testing set of July/2017 to Oct/2017 and July/2018 to Oct/2018 when compared with ARIMA-BU. This inaccuracy of the forecast in these periods increased the error significantly for the ETS-BU strategy. It appears that the forecast strategy ETS-BU does not properly predict the values for this range of months. Given these results it is expected that if the conditions for the series remain constant, this undesirable performance will continue for the following years. For this reason ETS-BU is not recommended as an adequate strategy for the study. On the other hand, the ETS-TD strategy did not show this conduct and it was able to achieve a more accurate forecast when compared with the actual values in the testing set, resulting in a significant lower MAPE. The ETS-TD strategy is the best approach for the study.

Certainly, the most important constrain when applying the proposed forecasting strategies is the arrangement of the data structure in combination with the inherent complexity for the BU strategy of having to calculate each forecast individually. Keeping track of the developed forecasts for each series might prevent researchers of applying the BU strategy. AUTO.ETS and AUTO.ARIMA functions are able to handle as many series as necessary, but the process of aggregating forecasts and evaluating performance is time consuming and in some cases unpractical to apply. If we also include series with different ending and starting times the process becomes more difficult to control. To our knowledge there is not in the literature an automatic procedure to apply the aggregation strategies with several series. The forecast package for R described in [25] details the use of AUTO.ETS and AUTO.ARIMA but before applying these functions with the forecast strategies, the data must be arranged with the appropriate structure in order to be able to perform. In this context, we developed two sets of algorithms presented in Figure 8, to automatically develop the arrangement of the data and to keep track of the forecasts when applying the strategies.

The first algorithm deals with the calculations in regards of different ending or starting periods of the series. The second algorithm evaluates the periods available for all series and arrange the structure for forecast calculation. We created both algorithms to be applied on our developments in R. For the algorithms two data frames denominated *master.data* and *dates.data* are generated. The data frame *master.data* contains all the actual values of the series, arranged each one in columns (n). The second data frame *dates.data* has the same number of columns than *master.data* with two rows for each column, the first row has the information of the starting year, for that specific series, and the second row the information about the starting period. The first algorithm will automatically calculate and save the ending date given the frequency, and the second algorithm will arrange the structure to run the forecasts strategies. The second algorithm evaluates if the size of a series is smaller when compared with the others. If that is the case only the tail values that match the size of the smallest series should be considered to calculate the combined forecast. All series must have an equal starting and ending time in order to apply the strategies.

Algorithm 1

```
start.dates <- matrix(nrow=2, ncol=n)
ending.dates <- matrix(nrow=2, ncol=n)
for (i in 1:n) {
  start.dates[1,i] <- dates.data[1,i]
  start.dates[2,i] <- dates.data[2,i]
  frequen <- frequency
  data <- na.omit(master.data[i])
  observations <- nrow(data)
  cycles <- observations/frequen
  years <- floor(cycles)
  plusperiods <- (cycles - years)* frequen
  ending.dates[1,i] <- start.dates[1,i] + years
  ending.dates [2,i] <- round(start.dates[2,i] + plusperiods - 1, 0)
  if (ending.dates[2,i] > frequen) {
    ending.dates[1,i] <- start.dates[1,i] + years + 1
    ending.dates[2,i] <-  ending.dates[2,i] - frequen
  }
}
```

Algorithm 2

```
observations <- 0
for (i in 1:n) {
  data <- as.ts((master.data[i]))
  observations[i] <- length(na.omit(data))
}
minvalue<- min(observations)
Nmiss <- tail(na.omit(master.data),minvalue)
```

Figure 8: Algorithms for implementation of strategies

After the application of the proposed algorithms we completed the procedure with the implementation of the different formulations detailed in [25] including the functions "*ts*" and "*as.ts*" to run the forecast automation procedure. The presented algorithms in combination with the automatic forecasting functions enhance significantly the efficiency, coordination, and development of the strategies when several time series are available. We recommend the use of a similar approach to apply the forecasting procedure of aggregation strategies. Finally, future research might include the analysis of machine learning approaches but it is necessary to consider the disadvantages of these methods when analyzing the obtained results.

5. References

1. Mentzer, J., Bienstock, C.: Sales Forecasting Management: Understanding the Techniques. Syst. Manag. Sales Forecast. Process. Sage Publ. (1998).
2. Hyndman, R., Koehler, A., Snyder, R., Grose, S.: A state space framework for automatic forecasting using exponential smoothing methods. Int. J. Forecast. 18, 439–454 (2002).
3. Fildes, R., Petropoulos, F.: Simple versus complex selection rules for forecasting many time series. J. Bus. Res. 68, 1692–1701 (2015).
4. Hendry, D., Hubrich, K.: Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate. J. Bus. Econ. Stat. 29, 216–227 (2011).

5. Widiarta, H., Viswanathan, S.: Forecasting family demand with autoregressive item-level demands: Evaluation of top-downand bottom-up forecasting strategies. Work. Pap. (2005).
6. McLeavey, D., Narasimhan, S.: Production Planning and Inventory Control. Allyn Bacon, Inc., Bost. (1985).
7. Lutkepohl, H.: Forecasting With VARMA Processes. Handb. Eco nomic Forecast. eds. G. Elliott, C. W. J. Granger, A. Timmermann, Amsterdam, Netherlands Elsevier. (2006).
8. Seifert, M., Siemsen, E., Hadida, A., Eisingerich, A.: Effective judgmental forecasting in the context of fashion products. J. Oper. Manag. 36, 33–45 (2015).
9. Nenova, Z., May, J.: Determining an optimal hierarchical forecasting model based on the characteristics of the dataset: technical note. J. Oper. Manag. 44, 62–88 (2016).
10. Van der Laan, E., Van Dalen, J., Rohrmoser, M., Simpson, R.: Demand forecasting and order planning for humanitarian logistics: an empirical assessment. J. Oper. Manag. 45, 114–122 (2016).
11. Zotteri, G., Kalchschmidt, M., Caniato, F.: The impact of aggregation level on forecasting performance. Int. J. Prod. Econ. 93, 479–491 (2005).
12. Theil, H.: Linear Aggregation of Economic Relations. North-holl. Publ. Company, Amsterdam. (1954).
13. Y., G., Griliches, Z.: Is Aggregation Necessarily Bad? Rev. Econ. Stat. 42, 1–13 (1960).
14. Schwarzkoph, A., Tersine, R., Morris, J.: Topdown versus bottom-up forecasting strategies. Int. J. Prod. Res. 26, 1833–1843 (1988).
15. Oller, L.: Aggregating problems when forecasting industrial production using business survey data. Discuss. Pap. No. 20, Minist. ofFinance, Econ. Dep. Helsinky. (1989).
16. Ilmakunnas, P.: Aggregation vs. disaggregation in forecasting construction activity. Barker, T.S., Pesaran, H. (Eds.), Disaggregation Econom. Model. Routledge, London. 73–86 (1990).
17. Kahn, K.: Revisiting top-down versus bottom-up forecasting. J. Bus. Forecast. 14–19 (1998).
18. Lapide, L.: New developments in business forecasting. J. Ofbus. Forecast. 28–29 (1998).
19. Orcutt, G., Watts, H., Edwards, J.: Data aggregation and information loss. Am. Econ. Rev. 58, 773–787 (1968).
20. Zellner, A., Tobias, J.: A note on aggregation, disaggregation and forecasting performance. J. Forecast. 19, 457–469 (2000).
21. Weatherford, L., Kimes, S., Scott, D.: Forecasting for hotel revenue management: Testing aggregation against disaggregation. Cornell Hotel Restaur. Adm. Q. 53–64 (2001).
22. Makridakis, S., Hibon, M.: The M3-competition: Results, conclusions and implications. Int. J. Forecast. 16, 451–476 (2000).
23. Chu, C., Zhang, P.: A comparative study of linear and nonlinear models for

aggregate retail sales forecasting. Int. J. Prod. Econ. 86, 217–231 (2003).
24. Ramos, P., Santos, N., Rebelo, R.: Performance of state space and ARIMA models for consumer retail sales forecasting. Robot. Comput. Integr. Manuf. 34, 151–163 (2015).
25. Hyndman, R., Khandakar, Y.: Automatic time series for forecasting: the forecast package for R. Clayt. VIC, Aust. Monash Univ. Dep. Econom. Bus. Stat. 6/7, (2007).
26. Billah, B., King, M., Snyder, R., Koehler, A.: Exponential smoothing model selection for forecasting. Int. J. Forecast. 22, 239–247 (2006).
27. Admed, N., Atiya, A., El gayar, N., El-shishiny, H.: An empirical comparision of machine learning models for time series forecasting. Econmetric Rev. 29, 594–621 (2010).
28. Sagaert, Y., Aghezzaf, E., Kourentzes, N., Desmet, B.: Tactical sales forecasting using a very large set of macroeconomic indicators. Eur. J. Oper. Res. 264, 558–569 (2018).
29. Gardner, E.: Exponential smoothing: The state of the art. J. Forecast. 4, 1– 28 (1985).
30. Alvarado-Valencia, J., Barrero, L., Önkal, D., Dennerlein, J.: Expertise, credibility of system forecasts and integration methods in judgmental demand forecasting. Int. J. Forecast. 33, 298–313 (2017).
31. Weller, M., Crone, S.: Supply Chain Forecasting: Best Practices & Benchmarking Study. Lancaster Cent. Forecast. Tech. Rep. (2012).
32. Brown, R.: Statistical forecasting for inventory control. New York McGraw Hill. (1959).
33. Holt, C.: Forecasting Trends and Seasonal by Exponentially Weighted Averages. Int. J. Forecast. 20, 5–13 (1957).
34. Winters, P.: Forecasting sales by exponentially weighted moving averages. Manage. Sci. 6, 324–342 (1960).
35. Makridakis, S., Wheelwright, S., Hyndman, R.: Forecasting: Methods and applications. John Wiley Sons, New York, 3rd ed. (1998).
36. Box, G., Jenkins, G.: Time Series Analysis: Forecasting and Control. Holden-Day, San Fr. (1976).
37. Rose, D.: Forecasting aggregates of independent ARIMA processes. J. Econom. 5, 323–345 (1977).
38. Lutkepohl, H.: Linear Transformations of Vector ARMA ProcesseS. J. Econom. 26, 283–293 (1984).
39. Barnea, A., Lakonishok, J.: An analysis ofthe usefulness of disaggregated accounting data for forecasts of corporate performance. Decis. Sci. 11, 17–26 (1980).
40. Dangerfield, B., Morris, J.: Top-down or bottom-up: Aggregate versus disaggregate extrapolations. Int. J. Forecast. 8, 233–241 (1992).
41. Fliedner, G.: An investigation ofaggregate variable time series forecast strategies with specific subaggregate time series statistical correlation. Comput. Oper. Res. 26, 1133–1149 (1999).