A systemic view of the ADIDA framework

Georgios P. SPITHOURAKIS*, Fotios PETROPOULOS
Forecasting & Strategy Unit, School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece
Konstantinos NIKOLOPOULOS
Bangor Business School, Bangor University, Bangor LL57 2DG, UK
&
Vassilios ASSIMAKOPOULOS
Forecasting & Strategy Unit, School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece

*Corresponding author: giorgos@fsu.gr

This paper is an attempt to gain mathematical insight into the ADIDA (Aggregate-Disaggregate Intermittent Demand Approach) forecasting framework, by formulating it as a multi-rate signal processing system. After a brief synopsis of the framework's background, an alternative way to perceive ADIDA from a systemic viewpoint is derived by breaking down its managerial steps into fundamental, well-studied components. Mathematical properties stemming from each separate system block are then thoroughly explored and their practical effects are exemplified through simulated paradigms of common time series patterns. Subsequently, theoretical and practical evidence are combined to draw useful conclusions about the framework's performance and make suggestions on its application. Finally, guidelines for further research are proposed.

Keywords: disaggregation; forecasting framework; multi-rate processing; systemic view; temporal aggregation.

1. Introduction

Aggregate-Disaggregate Intermittent Demand Approach (ADIDA; Nikolopoulos et al, 2011) is a methodology originally proposed for forecasting intermittent demand data; though it can be invariably used with non-intermittent demand. This paper is an attempt to demonstrate a systemic view of the ADIDA framework by its mathematical decomposition into fundamental sub-procedures, so as to give us understanding of the essential underlying processes that account for the properties and effectiveness of the ADIDA framework.

In brief, the ADIDA methodology can be broken down into four basic stages, as follows:

1. Gathering the original data at a lowermost observation level;

2. Aggregating at an aggregation level, A;

3. Forecasting with an extrapolation method, F;

4. Disaggregating the forecasts according to a disaggregation algorithm, D.

In more detail, the first stage refers to the formation of a time series to forecast by collection of relevant data. Subsequently, the observations are serially allocated into non-overlapping time buckets of an agreed length, A, starting from the end of the series and moving backwards. In case the length of the time series is not a multiple of the length of the buckets, the remaining $N \bmod A$ first samples are discarded. At the end of the second stage, the contents of each time bucket are summed to yield an aggregate time series, which is then extrapolated at the third stage. Finally, the aggregate forecasts are reverted back to the original observation level through a process of disaggregation, typically by weighting the forecasts using an appropriate set of weights.

The methodology is briefly symbolized as *ADIDA(A, F, D)*, where the aggregation level, forecasting method and disaggregation algorithm represent the framework's parameters. The framework is generic in the sense it is not connected to a unique combination of parameters. Furthermore, it can describe any forecasting method F as *ADIDA(A, F, -)*. However, the optimal choice of parameters has been a topic of recent research and discussion.

In the next section, the relative background literature of the ADIDA framework will be presented. The proposed systemic view of the existing framework is demonstrated in Section 3 and important mathematical insights derived from this formulation are exhibited in Section 4. These results are showcased by graphic simulated examples in Section 5. Finally, important conclusions are drawn and extensions for future research are proposed in Section 6.

2. Background literature

ADIDA framework revolves around the concepts of temporal non-overlapping aggregation and disaggregation to enhance the mathematical and managerial tools available to forecasters and, consequently, the accuracy of the forecasts themselves.

The instigation for the ADIDA framework stems from the zero values of intermittent time-series, which further complicate the series' processing (forecasting included) by rendering common mathematical manipulations impractical or ineffective. Intermittency naturally arises in many applications, such as inventory and spare parts management, which pose special requirements (Boylan & Syntetos, 2010). To tackle with such problems, the framework aims at reducing the series intermittency, so that standard tools can be applied efficiently. Apart from the obvious fact that the framework can reduce the intermittency of a time series to facilitate

forecasting and the empirical results that have shown the potential of such a practice, the mathematical properties of ADIDA have not yet been investigated in depth.

The empirical application of ADIDA yielded satisfactory results by improving forecasting accuracy and acting as a self-improving mechanism for popular forecasting methods. In particular, the framework was successfully tested on 5000 monthly demand inventory time series from stock keeping units of the Royal Air Force (Nikolopoulos *et al.*, 2011). In the same study, the framework was tested with the SBA (Syntetos & Boylan, 2001) and Naïve forecasting estimators, the former being an improvement of Croston's method for intermittent demand (Croston, 1972) and the latter producing the highest forecasting accuracy between the two. A follow up research (Babai *et al.*, 2012) verified the initial results and exhibited the robustness of the methodology with inventory metrics as well. Furthermore, the impact of ADIDA as a self-improving mechanism has been empirically verified for non-intermittent demand too. Specifically, the framework was tested on the monthly data of the M3-competition (Makridakis & Hibon, 2000) and the results portrayed a significant reduction of forecasting errors with simpler forecasting methods (Spithourakis *et al.*, 2011). The same paper demonstrated the importance of the level of temporal aggregation and proposed selection of an appropriate level upon minimization of the Mean Square Error (MSE), Bayesian Information Criterion (BIC; Schwarz, 1978) or Akaike Information Criterion (AIC; Akaike, 1974). The study utilized a deseasonalize-forecast-reseasonalize scheme to deal with time series' seasonality.

Apparently, the framework is inextricably connected to temporal aggregation. Major research work in the subject has been collectively described by Silvestrini & Veredas (2008). Proper selection of an appropriate aggregation level seems crucial to

the performance of ADIDA and research has revealed the remarkably good results of the managerial heuristic approach of setting the aggregation level to a product's lead time plus a review period (Nikolopoulos *et al.*, 2011; Babai *et al.*, 2012). However this approach is feasible only in the case of intermittent demand and when the lead time for each product is available. Temporal aggregation is also related to the much investigated cross-product aggregation, which is routinely implemented in hierarchical forecasting frameworks (Strijbosch *et al.*, 2008; Strijbosch & Moors, 2010; Viswanathan & Widiarta, 2008; Widiarta & Viswanathan, 2008). Temporal aggregation can be reversed through disaggregation schemes (e.g. Boot *et al.*, 1967; Chan, 1993; Feijoó *et al.*, 2003; Lisman & Sandee, 1964; Wei & Stram, 1990). Nikolopoulos *et al.* (2011) investigated the use of constant weights according to an equal (EQW – all weights equal to the inverse of the aggregation level), previous (PRW – weights equal to the original-aggregate observations' ratio of the last time bucket) and average (AVW – weights equal to the mean original-aggregate observations' ratio of all time buckets) weights schemes and found that, for their dataset, EQW yielded the best results in terms of forecasting accuracy.

3. A systemic view

Even though the procedural description of the ADIDA framework is easily and efficiently implementable from a managerial point of view, it is rather obscure from a mathematical standpoint. In fact, aggregation and disaggregation are composite procedures whose properties are not readily available for further mathematical analysis. An alternative formulation for the ADIDA framework can be reached by identifying the following two simple facts. First, temporal aggregation can be broken down into a simple moving average followed by down-sampling. Secondly,

disaggregation by a set of constant weights is equivalent to up-sampling and subsequently applying a weighted moving average on the data. Without loss of generality, we will consider the case that any remaining samples are discarded beforehand, so that the numbering for the series' indices begins from the first used sample.

Indeed, the definition of temporal aggregation at an aggregation level A is given by (3.1), which can be viewed as a sequential application of the formulas (3.2) and (3.3). The former equation stands for moving summation of A observations and is equivalent to a simple moving average (SMA) multiplied by the aggregation level. The latter one is the definition for down-sampling by A. This operation is executed by simply retaining every A-th observation and discarding all the rest. In all following equations, the following naming conventions are used:

$x(n)$: initial time series

$d(n)$: averaged time series (i.e. series after SMA filtering)

$y(n)$: aggregate time series

$\hat{y}(n)$: forecast model for aggregate series

$\hat{d}(n)$: downsampled aggregate forecast model

$\hat{x}(n)$: forecast model for initial series

Moreover, series are treated as zero-based, that is the first sample is obtained at point n=0.

$$y[n] = \sum_{k=0}^{A-1} x[nA + k] \tag{3.1}$$

$$d[n] = \sum_{k=0}^{A-1} x[n+k] \qquad (3.2)$$

$$y[n] = d[nA] \qquad (3.3)$$

According to the ADIDA framework, aggregation is followed by extrapolation of the aggregate series by use of a forecasting method, F. Such manipulation is analogous to data processing to derive a forecast series from an initial series by application of a forecast operator (3.4).

$$\hat{y} = F\{y\} \qquad (3.4)$$

Lastly, disaggregation under the simple scheme of constant weights can be performed in two steps described by (3.5) and (3.6). The first step refers in fact to a process of up-sampling by A, i.e. stretching the series by adding $A-1$ zeros after each real sample. The second equation represents a weighted moving average (WMA) of length A, where $\{w_k, \ k=0,\dots,A-1\}$ is the set of constant weights by which disaggregation is carried out.

$$\hat{d}[n] = \begin{cases} \hat{y}[n/A] & (n/A) \in Z \\ 0 & otherwise \end{cases} \qquad (3.5)$$

$$\hat{x}[n] = \sum_{k=0}^{A-1} w_k \hat{d}[n-k] \qquad (3.6)$$

By taking into consideration the aforementioned break down of aggregation and disaggregation into fundamental components, we end up with a systemic formulation for the ADIDA framework (Fig. 1), which allows for consistent mathematical handling of the framework from the beginning to the end. The properties of each of the fundamental components, i.e. moving averages and sample rate conversions, have been extensively explored. Moreover, this formulation reveals the astounding similarities of the ADIDA framework with multi-rate signal processing techniques (Crochiere & Rabiner, 1983; Oppenheim & Schafer, 1999). Such a comparison can give useful insight into the function of each individual component, as well as the whole process. Specifically, ADIDA methodology is shown to be made up of a reduction of data rate, signal processing and increase of the data rate back to the initial value. Data rate reduction is achieved by a decimation process, comprising in turn down-sampling preceded by a simple moving average, which effectively acts as a low-pass anti-aliasing filter. Data processing essentially consists in forecasting the series via an extrapolation method. Finally, the aggregate series is interpolated to restore the initial data rate, in which case the rate conversion is executed by up-sampling and a weighted moving average performs reconstruction filtering. In a multi-rate signal processing system, the reconstruction filter is typically a low-pass filter that suppresses any unwanted spectral duplicates.
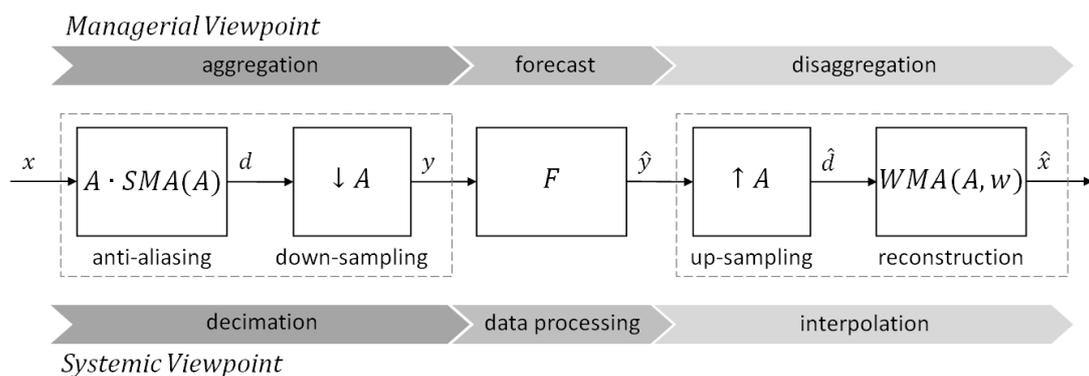
FIG. 1. Managerial and Systemic Viewpoint of the ADIDA framework.

4. Mathematical Insights

4.1. Aggregation

Aggregation is the first stage of the ADIDA framework and has been shown to be mathematically equivalent to decimation by the aggregation level, A.

In the time domain, SMA(A) is described by (3.1) with an extra division by A, the Z transform of which is shown by (4.1). There, z^{-1}, is the lag (backshift) operator, whereas its inverse, z, represents the lead operator. Thus, z^A represents a buffer of length A. It is clear that SMA can be handled as a linear time-invariant filter. Therefore, we derive its frequency response (4.2) by substituting $z=e^{j\omega}$ in (4.1). The frequency response is plotted in Fig. 2 for various aggregation levels.

$$D(z) = \frac{1}{A}\sum_{k=0}^{A-1} z^k X(z) = z^A \frac{1}{A}\sum_{k=1}^{A} z^{-k} X(z) \tag{4.1}$$

$$H(e^{j\omega}) = \frac{D(e^{j\omega})}{X(e^{j\omega})} = \frac{\sin\left(\dfrac{\omega A}{2}\right)}{A\sin\left(\dfrac{\omega}{2}\right)} e^{\frac{j\omega(A-1)}{2}} \tag{4.2}$$
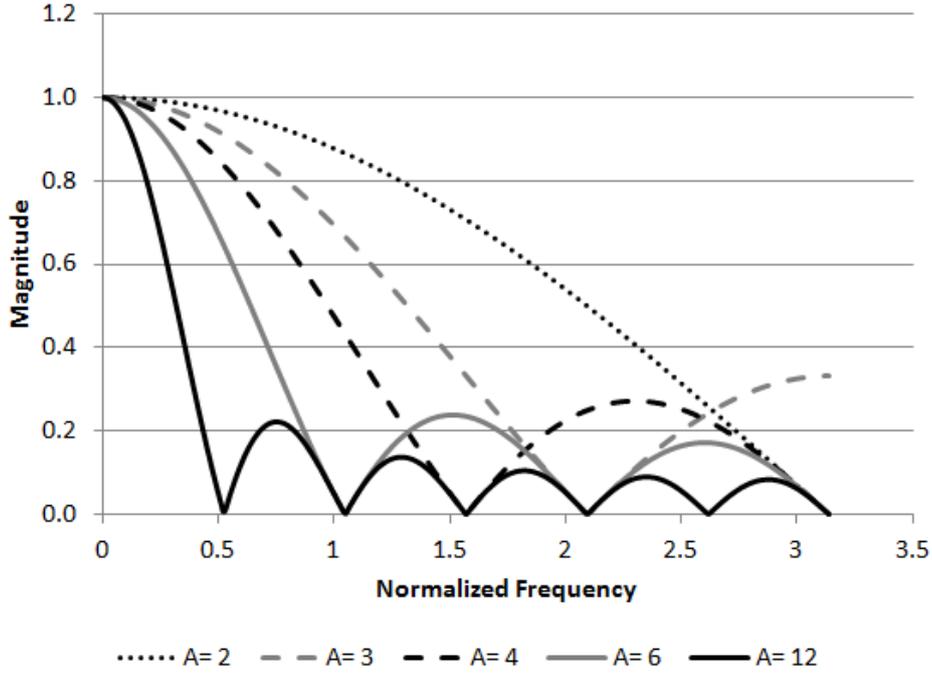
FIG. 2. Frequency response of the SMA(A) filter

The SMA(A) filter presents a low-pass behavior that is reinforced as A increases. However, the filter's response essentially differs from an ideal low-pass filter, because the *(A-1) div 2* sidelobes also let part of the higher frequencies pass, although with increasing attenuation. Maxima for these lobes occur at frequencies that are solutions to (4.3).

$$\tan\left(\frac{\omega A}{2}\right) = A\tan\left(\frac{\omega}{2}\right) \qquad (4.3)$$

There are *A-1* null points occurring at frequencies $\omega = \frac{2k\pi}{A}$, $k=1,\ldots,A$-1. The first null point, i.e. at $\omega = \frac{2\pi}{A}$, gives a rough estimate of the filters bandwidth. Another approximation can be made about the lobes' maxima, which approximately

occur in the middle between two consecutive nulls, i.e. at $\omega = \dfrac{(2k+1)\pi}{A}$. Therefore,

the worst case (at a lobe's maximum) gain in the stop-band area is approximately

$\left| A \sin\left(\dfrac{(2k+1)\pi}{2A} \right) \right|^{-1}$. The gain for zero frequency is unitary, that is to say the SMA(A)

does not interfere with the series mean level. Furthermore, the filter is linear-phase,

$\left(\dfrac{j\omega(A-1)}{2} \right)$, so the shape of the initial series is retained.

To complete aggregation, data must be down-sampled by A according to (3.3). If T_s is the sampling period for the initial series, the sampling period for the down-sampled series is $T'_s = AT_s$. To acquire the corresponding Z transform expression (A.1), one might consider using the train impulse or shah function (A.2), which equals zero everywhere except for where $\dfrac{n}{A} \in Z$ holds. Derivation of the relevant equations is presented in the Appendix.

Once again, substituting $z = e^{j\omega}$ produces the Fourier Transform of down-sampling (4.4). This reveals that down-sampling scales (stretches) the frequency axis by A, weights the spectrum by $1/A$ and creates spectral duplicates shifted by $\dfrac{2k\pi}{A}$. Accidental overlapping of these duplicates is called aliasing and is a major problem that can interfere with the information content of the data by inserting aliasing noise, a special case of which is insertion of false frequency patterns. Such noise effectively blocks accurate reconstruction of the initial series. Aliasing can be avoiding by cutting off content up to a critical frequency, i.e. applying a low-pass filter, before down-sampling. This critical frequency is found according to the Nyquist-Shannon sampling theorem. Assuming baseband series, the new sampling period must satisfy inequality (4.5), where B is the single-sided bandwidth of the filtered for anti-aliasing series.

$$Y(e^{j\omega}) = \frac{1}{A}\sum_{k=0}^{A-1} D\left(e^{j\frac{\omega-2k\pi}{A}}\right)$$

(4.4)

$$T_s' \le \frac{1}{2B}$$

(4.5)

As far as the effect of the aggregation level is concerned, important conclusions can be reached by identifying it with the length of the SMA. An increase of the aggregation level leads to boosting of the moving average's smoothing and noise rejection capabilities. Noise is an inherent component of many time series and it manifests as random fluctuations. In the special case of additive white Gaussian noise, the noise filtered with an SMA(A) has a mean value and standard deviation given by (4.6) and (4.7), respectively; nevertheless the sample independency is lost due to the averaging (in fact, the output noise is an MA(A) process). The latter equation reveals that the standard deviation of noise is reduced by a factor equal to $\frac{1}{\sqrt{A}} \le 1$.

$$\mu_d = \sum_{k=1}^{A} \frac{1}{A}\mu_x = \mu_x$$

(4.6)

$$\sigma_d = \frac{1}{\sqrt{A}}\sigma_x$$

(4.7)

On the other hand, the higher the aggregation level, the heavier the smoothing the data are subjected to. However, excess smoothing may suppress non-noise features of the series. Furthermore, high aggregation levels introduce substantial

lagging to the filtered series, which makes timely identification of fast changes, such as level shifts, harder.

Alternatively, the choice of an appropriate aggregation level can be viewed as selection of the lower sampling rate induced by down-sampling. It is crucial to avoid aliasing phenomena. That can be achieved by choosing A according to (4.8). The bandwidth B depends on the initial series bandwidth, as well as on the level of smoothing introduced by the SMA(A) which functions as an anti-aliasing low-pass filter.

$$A \leq \frac{1}{2BT_s} \tag{4.8}$$

4.2. Forecasting

ADIDA framework does not set any restrictions on the extrapolation method used. Even though any forecasting method can be applied, time series forecasting techniques, such as moving averages, exponential smoothing approaches (Makridakis *et al.*, 1998) or Theta method (Assimakopoulos & Nikolopoulos, 2000), are more direct and easier to apply. Another important consideration relates to the reduction of the number of available observations for model fitting due to aggregation. Poor model fitting may arise when one chooses forecast models with many parameters with a small number of observations that do not suffice for proper model training.

4.3. Disaggregation

Disaggregation is the last stage of the ADIDA framework, bearing strong resemblance to the mathematical concept of interpolation.

Up-sampling by *A* is described by (3.5) in the time domain. The frequency domain representation (4.9) is derived by use of the Z transform (A.3). If T'_s is the sampling period of a time series, the sampling period becomes $T_s = \dfrac{T'_s}{A}$ after up-sampling by *A*, which reverses the change to the sampling rate caused by down-sampling at the aggregation stage. The equations reveal that up-sampling by *A* is equivalent to scaling (shrinking) the series' spectrum by 1/*A*. Up-sampling does not relate to the adverse effect of aliasing.

$$\hat{D}(e^{j\omega}) = \hat{Y}(e^{j\omega A}) \tag{4.9}$$

The final computational step is the reconstruction of the up-sampled series to complete the disaggregation. The simplest way to do so is to use a set of constant weights to perform a WMA(*A*, w). The time domain representation is given by (4.10), while the disaggregation weights must satisfy a normalization condition and sum up to one. The Z transform of WMA(*A*, w) (4.11) makes use of the lag operator z^{-1} and perfectly portrays the operation of the moving average as a time-invariant linear filter. We observe that there is no need for a buffer. Finally, the frequency response is presented in (4.12) and is connected to the filter's impulse response (4.13) ($\delta[n-k]$ stands for the unit impulse (discrete delta) function centered at n=k), which is composed by the disaggregation weights.

$$\hat{x}[n] = \sum_{k=0}^{A-1} w_k \hat{d}[n-k] \tag{4.10}$$

$$\hat{X}(z) = \sum_{k=0}^{A-1} w_k z^{-k} \hat{D}(z) \tag{4.11}$$

$$\hat{H}(e^{j\omega}) = \frac{\hat{X}(e^{j\omega})}{\hat{D}(e^{j\omega})} = \sum_{k=0}^{A-1} w_k e^{-j\omega k} \tag{4.12}$$

$$\hat{h}[n] = \sum_{k=0}^{A-1} w_k \delta[n-k] = \begin{cases} w_n & n = 0,...,A-1 \\ 0 & otherwise \end{cases} \tag{4.13}$$

The disaggregation weights essentially shape the form and performance of the WMA filter. The weights must add up to one (normalization constraint) and should take nonnegative values. However, negative values can arise as a result of the disaggregation process, when we end up with negative values in the aggregate series. As a result of the normalization constraint, the framework will preserve the mean level of the original data, as long as the forecasting process accurately estimates the mean level of the aggregate series (which is usually required from a forecasting technique to have practical application). In other words, a shift by a constant in the original data will be carried over to the forecast model, save for any level shift induced by the forecasting step.

As a special case, we can consider the EQW disaggregation algorithm, that is to say every weight is equal to $1/A$; hence the WMA falls back to an SMA. Because of this similarity, the magnitude of the frequency response of the disaggregation SMA is again that of a linear-phase low-pass filter, shown by Fig. 2. The usual aim of the reconstruction filter is to cut off the processed spectral duplicates induced by the down-sampling in the aggregation stage. To be consistent with the multi-rate processing theory, the filter must be low-pass, a condition guaranteed by the EQW

disaggregation algorithm. On the contrary, the PRW and AVW algorithms fail to guarantee the low-pass nature of the filter.

5. Simulated Examples

In this section, a few simulated examples will be given to highlight the properties discussed in the previous section. In more detail, we will explore the behavior of ADIDA with three control signal types, ideally representing typical components of a time series. For each case, the investigated component and its corresponding forecast model will be plotted, as well as the corresponding in-sample mean square error (MSE) for various aggregation levels.

Firstly, a random series of white Gaussian noise is considered as a model for the series' randomness. The noise series and its *ADIDA(A, Naïve, EQW)* forecast model are presented in Fig. 3.a. Negative values naturally arise, despite the positive disaggregation weights, when the corresponding values of the aggregate forecast model are negative. For the sake of simplicity, we have selected a zero-mean Gaussian noise, yet the results hold for any mean, as ADIDA was shown to preserve the mean level of the initial data also in the forecast model. As a result of this property, the forecast model has the same mean as the initial noise signal. On the other hand, the model has a lower variance than the original input. Variance is further reduced for higher aggregation levels, due to increased smoothing. The output noise is stepwise, which reveals that observations are no longer independent. Moreover, the in-sample MSE error is plotted in Fig. 3.b. A Monte Carlo simulation scheme calling for multiple random noise series with the same properties was used to achieve a smoother graph. The in-sample error decreases at a declining rate and finally reaches convergence, when the forecast model becomes completely flatted out.
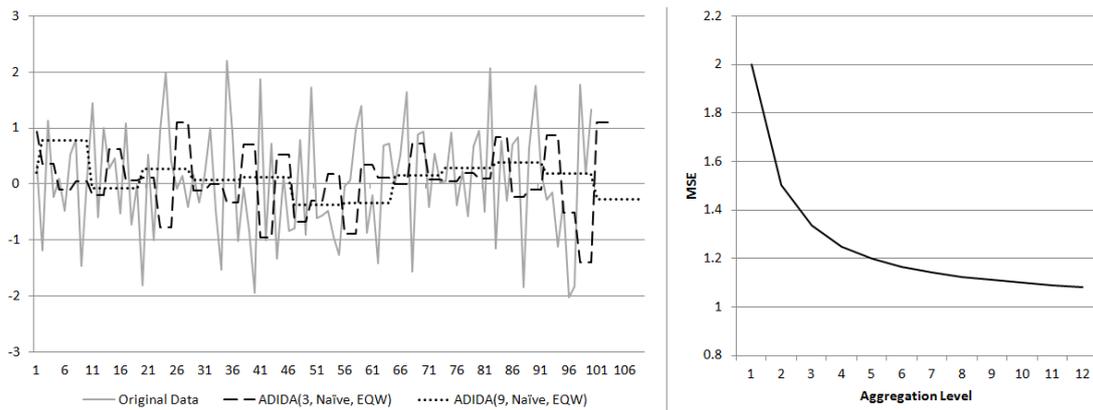
FIG. 3. (a) White Gaussian noise and its ADIDA models; (b) Corresponding in-sample MSE for various aggregation levels

Next, the ramp function portrayed in Fig. 4.a is selected to stand for a series' constant trend component. The *ADIDA(A, Naïve, EQW)* forecast model becomes a staircase function that exhibits lagging with respect to the actual trend line. The stepwise form results from the equal weights used for disaggregation. Lagging becomes an increasingly significant factor leading to poor accuracy, as higher aggregation levels are selected. This adverse effect can be observed in the rapidly increasing in-sample MSE (Fig. 4.b).
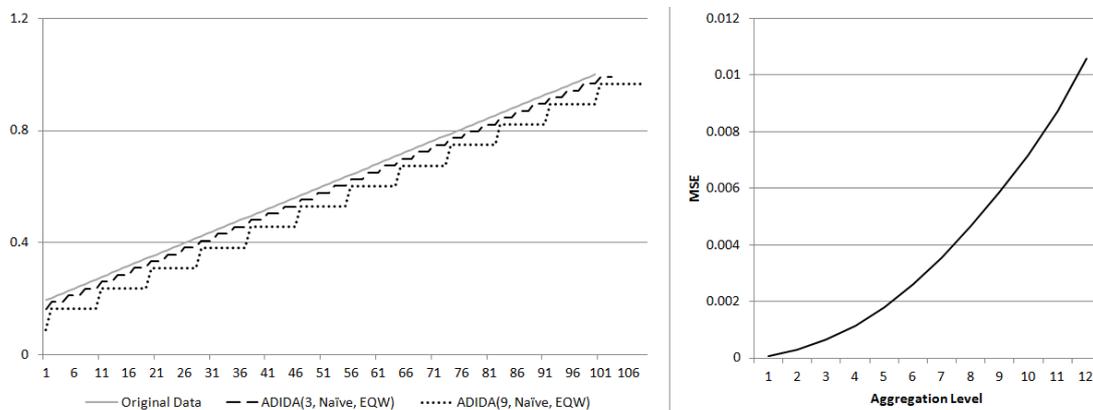
FIG. 4. (a) Linear trend and its ADIDA models; (b) Corresponding in-sample MSE for various aggregation levels

Last but not least, we observe the performance of ADIDA on periodic signals. A sine wave (12 points period) is selected to represent an ideal seasonality pattern. In this case, the ADIDA model takes a variety of forms according to the selected aggregation level. For lower aggregation levels ($A<6$), the ADIDA model is a lagging stepwise rendition of the sine. At $A=6$, which according to (4.8) is the highest possible value the aggregation level can obtain without aliasing occurring, this lagging virtually leads to the inversion of the models phase. For larger levels, aliasing occurs and ADIDA yields a false estimation of the real seasonality length (sine period). Furthermore, excess smoothing tends to suppress the model's extrema. These behaviors are graphically summarized in Fig. 5.b. Initially, lagging results to an increasing in-sample MSE that peaks at $A=6$, because of the phase inversion. As the output series gets smoothed out, the error reduces, albeit seasonality information would not be correct anymore.
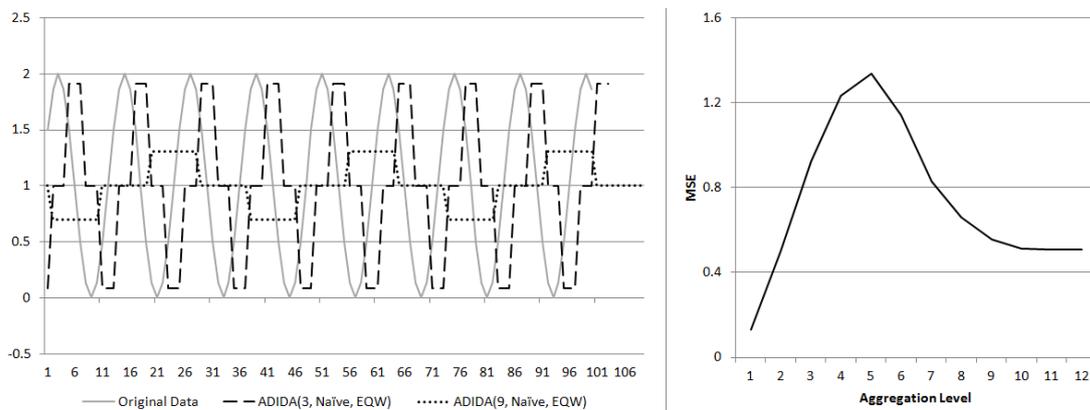

FIG. 5. (a) Periodic sine and its ADIDA models; (b) Corresponding in-sample MSE for various aggregation levels

6. Conclusions

In this paper a systemic view of the ADIDA framework was derived. The framework was broken down into fundamental sub-procedures and described as a multi-rate analogous process. This alternative formulation has been a starting point for a number of significant conclusions about the framework's implications that will be summarized in this section.

The aggregation step of ADIDA was shown to have smoothing properties, because of the SMA(A). Higher aggregation levels induce a greater degree of noise reduction. However, the same process introduces lagging that delays identification of fast signal changes. Besides, excess smoothing may suppress information rich variations of the time series. Moreover, the possible aliasing effect caused by down-sampling with an increased aggregation level may alter the input signal's frequency content. These considerations should be brought in mind when selecting an appropriate aggregation level.

No specific forecasting method is required by the ADIDA framework. Forecasters can choose any forecasting method, based on statistical or judgmental criteria of their own. It should be noted, however, that ADIDA reduces the number of available data points, which may rule out models that require large training datasets.

Disaggregation of the aggregate forecasts by a set of *A* constant weights can be implemented through up-sampling and filtering with those weights. This process does not relate to the adverse effect of aliasing. The simplest disaggregation algorithm, EQW, is equivalent to piecewise constant interpolation.

More practical implications of the abovementioned points were shown through simulated examples, considering the effect of aggregation level on the in-sample MSE

of the forecast models. In particular, ADIDA reduces the variance of the output noise for white Gaussian noise as input, increasingly with the aggregation level. The beneficial effect of higher aggregation level is displayed also by the falling in-sample MSE. However, the output noise observations are no longer independent. This example exhibited that ADIDA with higher aggregation levels blocks out the series' noise component.

On the other hand, ADIDA does not perform as well with trending or periodic data. In both cases, large aggregation levels may kill the input signal due to excess smoothing. Specifically for the ramp input, the in-sample MSE grows rapidly due to the lagging induced by the moving average. Therefore, lower aggregation levels should be considered. In the case of periodic input, aliasing may lead to false identification of the series' periodicity and lead to increased MSE values. Selection of an appropriate aggregation level according to (4.8) is essential for avoiding aliasing. Even so, the effect of lagging is still problematic, as it induces undesirable phase shifts. Consequently, seasonality should be handled before the ADIDA framework is applied. The original series should be deseasonalized so as to remove periodic variations. Afterwards, ADIDA may be used to produce forecasts which must then be reseasonalized.

More light should be shed into the mathematical properties of ADIDA and their managerial implications. The impact of the aggregation level on the methods performance should be further investigated and methods for selection of appropriate aggregation levels be proposed. In particular, the existence of optimal aggregation levels should be theoretically examined, as well the effect of the framework on specific data models. In addition, extrapolation with various forecasting methods under ADIDA can be investigated on large datasets, so as to acquire more empirical

data. Finally, alternative disaggregation algorithms should be tested and evidence be given about the appropriateness of each scheme. For instance, determination of the disaggregation weights according to the given time series, WMA of alternative lengths to allow for linear, cubic, etc. interpolation, as well as the possibility of doing without the reconstruction filter's time-invariance are only some of the possible scenarios that are worth exploring.

Appendix: derivation of down- and up-sampling frequency equations

This short appendix presents the derivation of the down-sampling and up-sampling equations, (4.4) and (4.9), respectively.

Down-sampling:

$$Y(z) = \sum_{n=-\infty}^{\infty} y[n]z^{-n} = \sum_{n=-\infty}^{\infty} x[nL]z^{-n} = \sum_{m=\ldots,-A,0,A,\ldots} x[m]z^{-\frac{m}{L}} \Leftrightarrow$$

$$Y(z) = \sum_{m=-\infty}^{\infty} x[m]II_A[m]z^{-\frac{m}{A}} = \sum_{m=-\infty}^{\infty} x[m]\left(\frac{1}{A} \sum_{k=0}^{A-1} e^{-\frac{j2k\pi m}{A}} \right) z^{-\frac{m}{A}} \Leftrightarrow \qquad (A.1)$$

$$Y(z) = \frac{1}{A} \sum_{k=0}^{A-1} \left[\sum_{m=-\infty}^{\infty} x[m]z^{-\frac{m}{A}} e^{-\frac{j2k\pi m}{A}} \right] = \frac{1}{A} \sum_{k=0}^{A-1} X\left(z^{\frac{1}{A}} e^{-\frac{j2k\pi}{A}} \right)$$

$$II_A[n] = \frac{1}{A} \sum_{k=0}^{A-1} e^{-\frac{j2k\pi n}{A}} = \begin{cases} 1 & \frac{n}{A} \in Z \\ 0 & otherwise \end{cases} \qquad (A.2)$$

Up-sampling:

$$Y(z) = \sum_{n=-\infty}^{\infty} y[n]z^{-n} = \sum_{n=\ldots,-A,0,A,\ldots} x\left[\frac{n}{A} \right] z^{-n} = \sum_{m=-\infty}^{\infty} x[m]z^{-mA} = X(z^A) \qquad (A.3)$$

REFERENCES

AKAIKE, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, pp. 716-723.

ASSIMAKOPOULOS, V. & NIKOLOPOULOS, N. (2000) The theta model: a decomposition approach to forecasting. *International Journal of Forecasting*, **16**, pp. 521-530.

BABAI, M. Z., ALI, M. M. & NIKOLOPOULOS, K. (2012) Impact of Temporal Aggregation on Stock Control Performance of Intermittent Demand Estimators: Empirical Analysis. *OMEGA – International Journal of Management Science, 40 (6)*, pp. 713-721.

BOOT, J. C. G., FEIBES, W. & LISMAN, J. H. C. (1967) Further Methods of Derivation of Quarterly Figures from Annual Data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **6**, pp. 67-75.

BOYLAN, J. E., & SYNTETOS, A. A. (2010) Spare parts management: a review of forecasting research and extensions. *IMA Journal of Management Mathematics*, **21**, pp. 227-237.

CHAN, W. S. (1993) Disaggregation of Annual Time-series Data to Quarterly Figures: A Comparative Study. *Journal of Forecasting*, **12**, pp. 677-688.

CROCHIERE, R. E., & RABINER, L. R. (1983) *Multirate Digital Signal Processing*, Englewood Cliffs, N.J.: Prentice-Hall.

CROSTON, J. D. (1972) Forecasting and Stock Control for Intermittent Demands. *Operational Research Quarterly*, **23**, pp. 289-303.

FEIJOÓ, S. R., CARO, A. R. & QUINTANA, D. D. (2003) Methods for quarterly disaggregation without indicators; a comparative study using simulation. *Computational Statistics & Data Analysis*, **43**, pp. 63-78.

LISMAN, J. H. C. & SANDEE, J. (1964) Derivation of Quarterly Figures from Annual Data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **13**, pp. 87-90.

MAKRIDAKIS, S. & HIBON, M. (2000) The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, **16**, pp. 451-476.

MAKRIDAKIS, S., WHEELWRIGHT, S.C., & HYNDMAN, R.J. (1998) *Forecasting: Methods and Applications* (3rd ed.), New York: John Wiley and Sons.

NIKOLOPOULOS, K., SYNTETOS, A. A., BOYLAN, J. E., PETROPOULOS, F. & ASSIMAKOPOULOS, V. (2011) An aggregate-disaggregate intermittent demand approach (ADIDA) to forecasting: an empirical proposition and analysis. *Journal of Operational Research Society*, **62**, pp. 544-554.

OPPENHEIM, A. V., & SCHAFER, R. W. (1999) *Discrete-Time Signal Processing* (2nd ed.), Englewood Cliffs, N.J.: Prentice-Hall.

SCHWARZ, G. E. (1978) Estimating the dimension of a model. *Annals of Statistics*, **6**, pp. 461–464.

SILVESTRINI, A. & VEREDAS, D. (2008) *Temporal aggregation of univariate and multivariate time series models: a survey*. Temi di discussione (Working papers) 685 (August 2008), Banca d'Italia.

SPITHOURAKIS, G. P., PETROPOULOS, F., BABAI, M. Z., NIKOLOPOULOS, K. & ASSIMAKOPOULOS, V. (2011) Improving the performance of popular supply chain forecasting techniques: an empirical investigation. *Supply Chain Forum: an International Journal*, **12** (4), pp. 16-25.

STRIJBOSCH, L. W. G., HEUTS, R. M. J. & MOORS, J. J. A. (2008) Hierarchical estimation as a basis for hierarchical forecasting. *IMA Journal of Management Mathematics*, **19**, pp. 193-205.

STRIJBOSCH, L. W. G. & MOORS, J. J. A. (2010) Calculating the accuracy of hierarchical estimation. *IMA Journal of Management Mathematics*, **21**, pp. 303-315.

SYNTETOS, A. A. & BOYLAN, J. E. (2001) On the bias of intermittent demand estimates. *International Journal of Production Economics*, **71**, pp. 457-466.

VISWANATHAN, S. & WIDIARTA, H. (2008) Forecasting aggregate time series with intermittent subaggregate components: top-down versus bottom-up forecasting. *IMA Journal of Management Mathematics*, **19**, pp. 275-287.

WEI, W. W. S. & STRAM, D. O. (1990) Disaggregation of time series models. *Journal of the Royal Statistical Society. Series B3*, **52**, pp. 453-467.

WIDIARTA, H. & VISWANATHAN, S. (2008) Forecasting item-level demands: an analytical evaluation of top-down versus bottom-up forecasting in a production-planning framework. *IMA Journal of Management Mathematics*, **19**, pp. 207-218.