Method Description

General Information

Type of Entry (*Academic, Practitioner, Researcher, Student*)	**Student**
First Name	**Tugce Ekiz**
Last Name	**Yilmaz**
Country	**Turkey**
Type of Affiliation (*University, Company-Organization, Individual*)	**--**
Affiliation	**--**

Information about the method utilized

Name of Method	ATA-2
Type of Method (*Statistical, Machine Learning, Combination, Other*)	Statistical
Short Description (up to 200 words)	ATA method is an innovative new forecasting technique where the forms of the models are similar to exponential smoothing models but the smoothing parameters depend on the sample size, are optimized on a discrete space and initialization is both easier as it is done simultaneously when the parameters are optimized and is less influential since the weights assigned to initial values approach zero quickly. ATA can be applied to all time series settings easily and provides better forecasting performance due to its flexibility. ATA-2 utilizes a damped version of ATA and a simple combination that is calculated by first performing a model selection between the linearly and multiplicatively trended models and then combining forecasts from this selection and the simple model.

Extended Description:

The Model:

For a time series $\{y_1, \ldots y_n\}$ *ATA method can be given in additive form as below:*

$$l_t = \left(\frac{p}{t}\right) y_t + \left(\frac{t-p}{t}\right)(l_{t-1} + \emptyset b_{t-1}), \tag{1.1}$$

$$b_t = \left(\frac{q}{t}\right)(l_t - l_{t-1}) + \left(\frac{t-q}{t}\right)(\emptyset b_{t-1}), \tag{1.2}$$

where p *is the smoothing parameter for level,* q *is the smoothing parameter for trend,* $\emptyset$ *is the dampening parameter and* $l_t = y_t$ *for* $t \le p$, $b_t = y_t - y_{t-1}$ *for* $t \le q$, $b_1 = 0$, $p \in \{1,2,\ldots,n\}$, $q \in \{0,1,2,\ldots,p\}$, $\emptyset \in (0,1]$. *Then, the h step ahead forecasts can be obtained by:*

$$\hat{y}_{t+h|t} = l_t + (\emptyset + \emptyset^2 + \cdots + \emptyset^h) b_t. \tag{1.3}$$

Similarly for a time series $\{y_1, \ldots y_n\}$ *ATA method can be given in multiplicative form as below:*

$$l_t = \left(\frac{p}{t}\right) y_t + \left(\frac{t-p}{t}\right)(l_{t-1} b_{t-1}^{\emptyset}), \tag{2.1}$$

$$b_t = \left(\frac{q}{t}\right)\left(\frac{l_t}{l_{t-1}}\right) + \left(\frac{t-q}{t}\right)(b_{t-1}^{\emptyset}), \tag{2.2}$$

where again p *is the smoothing parameter for level,* q *is the smoothing parameter for trend,* $\emptyset$ *is the dampening parameter and* $l_t = y_t$ *for* $t \le p$, $b_t = \frac{y_t}{y_{t-1}}$ *for* $t \le q$, $b_1 = 1$, $p \in \{1,2,\ldots,n\}$, $q \in \{0,1,2,\ldots,p\}, \emptyset \in (0,1]$. *Then, the h step ahead forecasts can be obtained by:*

$$\hat{y}_{t+h|t} = l_t + b_t^{(\emptyset + \emptyset^2 + \cdots + \emptyset^h)}. \tag{2.3}$$

Since both versions of the method require three parameters we will distinguish between them by using the notation $ATA_{add}(p,q,\emptyset)$ *for the additive form and* $ATA_{mult}(p,q,\emptyset)$ *for the multiplicative form.*

Notice that when $q = 0$ *both forms of ATA are reduced to the simple form* $ATA(p,0,\emptyset)$ *which can be written as:*

$$l_t = \left(\frac{p}{t}\right) y_t + \left(\frac{t-p}{t}\right) l_{t-1}, \tag{3.1}$$

where $p \in \{1,2,\ldots,n\}$ *and* $l_t = y_t$ *for* $t \le p$. *Forecasts then can be obtained by* $\hat{y}_{t+h|t} = l_t$.

When $q \ne 0$ *and* $\emptyset = 1$ *the additive and multiplicative forms of ATA are reduced to the trended versions* $ATA_{add}(p,q,1)$ *and* $ATA_{mult}(p,q,1)$ *which are given below respectively:*

$$l_t = \left(\frac{p}{t}\right) y_t + \left(\frac{t-p}{t}\right)(l_{t-1} + b_{t-1}), \tag{4.1}$$

$$b_t = \left(\frac{q}{t}\right)(l_t - l_{t-1}) + \left(\frac{t-q}{t}\right)(b_{t-1}), \tag{4.2}$$

$$\hat{y}_{t+h|t} = l_t + h b_t\ , \tag{4.3}$$

and

$$l_t = \left(\frac{p}{t}\right) y_t + \left(\frac{t-p}{t}\right)(l_{t-1} b_{t-1}), \tag{5.1}$$

$$b_t = \left(\frac{q}{t}\right)\left(\frac{l_t}{l_{t-1}}\right) + \left(\frac{t-q}{t}\right) b_{t-1}, \quad (5.2)$$

$$\hat{y}_{t+h|t} = l_t + b_t^h. \quad (5.3)$$

To sum up, ATA can be given in 5 forms, namely the additive damped form $ATA_{add}(p, q, \emptyset)$ (equations (1.1)-(1.3)), multiplicative damped form $ATA_{mult}(p, q, \emptyset)$ (equations (2.1)-(2.3)), simple form $ATA(p, 0, \emptyset)$ (equation (3.1)), additive trend form $ATA_{add}(p, q, 1)$ (equations (4.1)-(4.3)) and finally multiplicative trend form $ATA_{mult}(p, q, 1)$ (equations (5.1)-(5.3)).

The parameter values that minimized the in-sample one step ahead sMAPE are used as optimum values and optimization is carried out for all the parameters simultaneously.

Obtaining the point forecasts:

1. *The data sets are tested for stationarity using the Augmented Dickey-Fuller test using the function ndiffs from the forecast package in R.*

2. *The data sets are tested for seasonality using the SeasonalityTest function from the 4Thetamethod.R code downloaded from the competitions GitHub repository for $\alpha = 0.20$ and seasonality periods 4 for quarterly, 12 for monthly, 7 for daily. The data sets that could be classified as seasonal by this test are de-seasonalized by the classical multiplicative decomposition method. The hourly and weekly data are treated slightly different as they have not been put through the seasonality test and all have been de-seasonalized using the classical multiplicative decomposition method for periods 168 and 52 respectively. Note that if the length of the series was not sufficient to calculate the seasonality indexes, they were treated as non-seasonal.*

3. *Arbitrary models were used for different data types.*

   Yearly: $ATA_{add}(p, 1, \emptyset)$ was used to obtain forecasts where $\emptyset \in \{0.10, 0.15, \cdots, 1\}$.

   All other data types: The forecasts from two models were simply averaged to obtain point forecasts for each horizon. The first forecasts are obtained from the simple model $ATA(p, 0, \emptyset)$. The second forecasts are obtained from either the $ATA_{add}(p, 1, 1)$ or the $ATA_{mult}(p, 1, 1)$ whichever has a smaller in-sample sMAPE. Then the first and the second forecasts are averaged to obtain the final point forecasts.

4. *The forecasts are re-seasonalized using the seasonality indexes from the classical multiplicative decomposition to obtain final point forecasts. If any negative forecasts are obtained they are set equal to zero.*

Obtaining the prediction intervals:

For forecasting horizon h the prediction interval is obtained by:
$\hat{y}_{n+h|n} \pm C_h$,
where $C_h = \sqrt{h} Z_{\alpha/2} S_e$, $Z_{\alpha/2}$ *is the Normal deviate corresponding to* $(1-\alpha)\%$ *confidence interval and* S_e *is the standard deviation of the one step ahead errors of model fitting. If any lower bounds are found to be negative, they are set equal to zero.*

R-package:

You can download an R package that we developed for our method using the code below:

devtools::install_github("alsabtay/ATAforecasting")

Alternatively, you can visit https://atamethod.wordpress.com/software/ to download the package. The package has a help file that can guide you to fit the necessary versions of ATA to calculate the forecasts as explained above.