# APPLICATION OF TRAMO-SEATS AUTOMATIC PROCEDURE FOR FORECASTING SPORADIC AND IRREGULAR DEMAND PATTERNS WITH SEASONALITY

**F. Lolli[(a)], R. Gamberini[(a)], A. Regattieri[(b)], B. Rimini[(a)], A. Grassi[(a)], P. Belluti[(a)]**

[(a)]Department of Engineering Sciences and Methods, University of Modena and Reggio Emilia,
Via Amendola, 2, Padiglione Morselli, Reggio Emilia, Italy
[(b)] Department of Industrial and Mechanical Plants, University of Bologna,
Viale Risorgimento, 2, Bologna, Italy

[(a)]flolli@unimore.it; rita.gamberini@unimore.it; bianca.rimini@unimore.it; andrea.grassi@unimore.it; 39698@studenti.unimore.it, [(b)]alberto.regattieri@unibo.it

**ABSTRACT**
Managing sporadic and irregular demand patterns represents a relevant issue in several industrial contexts. Two main aspects have to be underlined due to their prominence: the former is the problem of forecasting future demand profiles, and the latter choosing and determining the best re-order policy to be applied, in accordance with information gained during the forecasting step. In this paper the former issue is discussed, by focusing on the management of items with sporadic and irregular demand patterns that also present a seasonality component. TRAMO-SEATS is a versatile procedure that allows quick identification of the best SARIMA forecasting model from an available set. Results obtained by its implementation are compared with those obtained by the Croston (1972) and Syntetos-Boylan (2005) methods, which represent two modified versions of simple exponential smoothing, introduced in literature for forecasting mean demand size per period specifically in case of irregular and sporadic demand profiles. In particular, two items are analysed, with the aim of demonstrating that when the strict hypothesis required by Croston's and Syntetos-Boylan's approaches fails, alternative forecasting methods could be required. TRAMO-SEATS represents a promising and user-friendly option.

Keywords: irregular and sporadic demand patterns, TRAMO-SEATS software, Croston's and Syntetos-Boylan's forecasting methods.

## 1. INTRODUCTION
Managing irregular and sporadic demand patterns is a relevant issue in several real industrial contexts, in terms of both their supply management and the application of different forecasting methods. In this paper, only the latter issue is handled, by comparing alternative forecasting methods which are briefly explained below here. Specifically, items affected by seasonality are studied.

The paper is organized as described in the following. A synthesis of the main contributions from literature is reported in section 2. The forecasting methods applied and compared in this paper are briefly explained in section 3, while section 4 describes some experimental data. Specifically, given a brief introduction of the steps followed (in section 4.1), data preliminary analysis is discussed in section 4.2. Section 4.3 reports the results of the experimentation and finally some interesting results and guidance for practitioners are given in section 5.

## 2. LITERATURE
Since the issue of demand forecasting is extremely wide-ranging, this section aims to summarise main contributions from literature on the forecasting methods applied only in the case of irregular and sporadic demand patterns. In particular, it focuses on the application in this context of Croston (1972) and Syntetos-Boylan's (2005) approaches, along with a tool named TRAMO-SEATS (Gómez and Maravall 1996) which uses a SARIMA-based automatic procedure both for identifying models that fit better with the time series and for forecasting future demand values.

Croston (1972) published a pioneering work concerning the forecast of irregular and sporadic demand. Starting from single exponential smoothing, Croston observes that it attains inappropriate results when applied to intermittent demand patterns, that is to say when demand does not occur in frequent time periods. Computing both the expected size of non-null demand occurrences and the expected interval between such occurrences is the insight provided by Croston in order to achieve the estimator of mean demand per period. In particular, Croston considers customers' order series with demand occurrences generated by a Bernoulli process and with demand sizes (when non-null) following a normal distribution. Then, Croston applies a single exponential smoothing separately to non-null demand sizes and inter-demand intervals. Finally he combines them to obtain such an estimator.

Several modifications and experimental analysis of Croston' approach have been successively proposed in literature (Rao 1973; Johnston and Boylan 1996;

Syntetos and Boylan 2005). In particular, Syntetos and Boylan firstly (2001) explain the detection of a mistake in Croston's mathematical derivation of the expected estimate of demand per time period and then (2005) they introduce a factor equal to (1- $\alpha/2$) applied to Croston's original estimator of mean demand, with $\alpha$ equal to the smoothing parameter used for updating the inter-demand intervals, in order to obtain a theoretically unbiased estimator. The derivation of the new estimator is based on Croston's assumptions of stationary, identically, independently distributed series of demand sizes and demand intervals, geometrically distributed inter-demand intervals and normally distributed demand sizes.

The SARIMA (Seasonal Autoregressive Integrated Moving Average) model, which is discussed in section 3, represents a appreciable robust approach due to its applicability to a wide variety of operative conditions. For a more detailed discussion on the application of SARIMA models see Jarrett (1991) and Bowerman and O'Connel (1993). However, it seems to require the close attention of time series analysts and considerable computing resources (Maravall 2006). Thus, although the SARIMA model has been neglected for same years due to its complexity, in particular in real-world applications, several statistics software have been developed in order to make it automatically applicable. Contextually, alternative methodologies for seasonal adjustment of time series are introduced (see for example Burman 1980, Hilmer and Tiao 1982). In particular, the TRAMO-SEATS (TS) tool allows the automatic application of the SARIMA model without requiring considerable computing resources, thus improving its implementation in real-world environments.

In synthesis, on one hand several authors investigate the possibility of improving the applicability of SARIMA-based methods by reducing their computational efforts in a wide range of operative conditions, while on the other hand *ad hoc* forecasting methods for irregular and sporadic demand patterns have yet to be improved. Nevertheless their comparison in the case of irregular and sporadic demand patterns, especially with trend and/or seasonal components, still remains a field that needs to be widely investigated.

Hence, comparing the forecasting performances obtained both by automatic software based on the SARIMA model (TS) and by forecasting methods *ad hoc* for irregular and sporadic demand patterns is the aim of this paper, without resorting to the initial research of trend and/or seasonal components of the time series. In fact, the advantage of the automatic software TS is that it can be used without any specific knowledge of time series analysis, and thus it represents a useful operative tool in a practical perspective.

## 3. METHODS APPLIED

In this section, a brief analytic explanation of the forecasting methods applied below is proposed.

For a more detailed discussion about them, see Makridakis *et al*. (1997).

### 3.1. Croston's method - CR

Croston (1972) proposes a method that takes account of both demand size and inter-arrival time between demand, which are assumed to have constant means and variances for modelling purposes, and to be mutually independent. Demand is assumed to occur as a Bernoulli process. Croston's method indicates the following forecasting steps: single exponential method evaluation, only when demand occurs, both for the smoothed demand size at the end of the review time period $t$ ($Z_t$), and the smoothed interval between non-null demands ($P_t$), using the same smoothing constant value. Thus, the equations are as follows:

$$Z_t = \alpha Y_t + (1-\alpha) Z_{t-1} \qquad (1)$$

$$P_t = \alpha G_t + (1-\alpha) P_{t-1} \qquad (2)$$

where $G_t$ is the actual value of the time between consecutive transactions at the time $t$, and $\alpha$ is the smoothing parameter.

If no demand occurs, then the smoothed estimates remain exactly unchanged. Otherwise, if demand occurs the estimates are updated. If demand occurs in every time period, Croston's estimator is identical to SES (Single Exponential Smoothing) in each time period.

The forecast of demand per period ($\bar{F}_{t+1}$) at the end of time period $t$ is given by:

$$\bar{F}_{t+1} = \frac{Z_t}{P_t} \qquad (3)$$

### 3.2. Syntetos-Boylan's Approximation - SBA

An error in Croston's mathematical derivation of expected demand size is reported by Syntetos and Boylan (2001), who propose a revision to approximately correct Croston's demand estimates: the SBA method.

Several variations have been applied to Croston's method since its introduction in 1972, but SBA is considered to be the most effective by several authors.

The forecast of demand per period ($\bar{F}_{t+1}$) at the end of time period $t$ is given by:

$$\bar{F}_{t+1} = \frac{\left(1-\frac{\alpha}{2}\right)Z_t}{P_t} \quad (4)$$

Note that equation 4 is similar to equation 3, except for the presence of a corrective factor ( $1-\frac{\alpha}{2}$ ) depending on the smoothing parameter $\alpha$ . In fact, both $Z_t$ and $P_t$ have the same meaning as those in equation 3.

### 3.3. Seasonal Auto-Regressive Integrated Moving Average – SARIMA

This is a group of methods which consist of two parts: an autoregressive (AR) part and a moving average (MA) part. In order to define the SARIMA model, which is the more generic forecasting method belonging to this group, each of its constituent components is introduced.

An autoregressive forecasting model of the order of *p*, AR(*p*), has the form:

$$F_t = \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + \ldots + \rho_p Y_{t-p} + \varepsilon_t \quad (5)$$

where $\varepsilon_t$ is a residual term that represents random events that are not explained by the model, while $\rho_i$ is a coefficient related to time period *i*.

A moving average forecasting model uses lagged values of the forecast error to improve the current forecast. A first-order moving average term uses the most recent forecast error, a second-order term uses the forecast error from the two most recent periods, and so on. A moving average forecasting model of the order of *q*, MA(*q*), has the form:

$$F_t - \varepsilon_t + \vartheta_1 \varepsilon_{t-1} + \vartheta_2 \varepsilon_{t-2} + \cdots + \vartheta_q \varepsilon_{t-q} \quad (6)$$

where $\varepsilon_i$ and $\vartheta_i$ are respectively the residual term and the coefficient related to time period *i*.

When the time series is stationary (the average and variance do not change over time), then an ARMA(*p,q*) can be applied in the following form:

$$F_t = \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + \cdots + \rho_p Y_{t-p} + \varepsilon_t + \vartheta_1 \varepsilon_{t-1} \quad (7)$$

AR and MA are combined: *p* is the degree of AR and *q* is the degree of MA.

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. It is applied in some cases where data shows evidence of non-stationarity, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to remove the non-stationarity. When the “integrated” part of the model is removed, then the process is the same as ARMA. The model is referred to as an ARIMA(*p,d,q*) model where *p*, *d* and *q* are integers greater than or equal to zero and refer respectively to the order of the autoregressive, integrated, and moving average parts of the model. When one of the terms is zero, it is usual to drop AR, I or MA. For example, an I(1) model is ARIMA(0,1,0), and a MA(1) model is ARIMA(0,0,1).

SARIMA(*p,d,q*) x $(P,D,Q)_s$ is used in case of seasonality of the order *s*. The procedure is the same as ARIMA but in this case there are three other degrees: *P*, *D* and *Q*; they have the same meaning as *p*, *d*, *q* but they are only applied to the seasonal data in the periods *t*, *t-s*, *t-2s*,..., where *s* is the seasonality length.

In synthesis, the seven parameters (*p*, *q*, *d, P*, *D, Q, s*) uniquely define each SARIMA model that suitably fits the original time series. Box and Jenkins (1976) formalise the complete procedure to be applied to identify the model, avoiding over-fitting occurrences (the need to test too many parameters). Many techniques and methods have been suggested to add mathematical rigour to the search process, including Akaike’s criterion (1974) or Schwarz’s criterion (1978), which penalise models based on their number of parameters. In any case, since nowadays statistical commercial software allows the user to test several SARIMA models very quickly, the identification of the model can only be based on its forecasting performances (Gamberini *et al.* 2010).

In this paper, the TS software is used to both identify the proper SARIMA model and to generate forecasts. It is composed of two modules, TRAMO (Time Series Regression with ARIMA Noise, Missing Observations and Outliers) and SEATS (Signal Extraction in ARIMA Time Series) respectively. The former, after eliminating deterministic effects from the time series (e.g. fixed holidays such as Easter, anomalous values and so on), automatically identifies and estimates the ARIMA model, applying the same tools listed in the Box-Jenkins procedure (e.g. sample autocorrelation ACF and partial autocorrelation PACF, Akaike’s criterion, Schwarz’s criterion, maximum likelihood estimation, and so on). The latter, using the results obtained by TRAMO, achieves the decomposition of the time series in its non-observable components, such as cycle-trend, seasonal component and irregular component. In synthesis, the aggregation of TRAMO and SEATS allows the identification of the final SARIMA model and thus it represents a full software package for the analysis, decomposition and forecasting of the time series. For a more detailed discussion about such a tool, the reader can refer to Pollock (2002) and Maravall (2006).

## 4. CASE STUDY

Below, data analysed in a case study is described. Specifically, with the presentation of the steps followed during experimentation, the data collected is commented on, introducing the results obtained.

### 4.1 Framework of the experimentation

The steps followed in the experimentation are presented below:

- Selection of two time series, characterised by irregularity and sporadicity, along with seasonality. In particular, analysed items do not satisfy the hypothesis required by CR and SBA approaches.
- For each time series, application of CR and SBA methods and utilisation of TS software, with the aim of achieving forecasts for the different number of months ahead (1, 3, 6, 12).
- Comparison of the aforementioned forecasting methods in terms of different accuracy measures.

Several accuracy measures are presented in literature to compare the performances of forecasting methods. For a more detailed discussion about them, see Makridakis (1993).

Define $T$ as the number of forecasted time periods, $F_t$ as the forecasted demand size in time period $t$, A as the mean demand size occurring in the forecasted time periods and $D_t$ as the real demand size occurring in time period $t$, for $t = 1,..., T$. In accordance with guidelines reported in Regattieri *et al.* (2005), the following accuracy measures are adopted (equations 8, 9, 10):

- *MAD/A:* Represents the Mean Absolute Deviation (*MAD*) divided by the average demand size. By describing the incidence of the mean absolute forecasting error in the mean existing demand, this index allows the performance of forecasting approaches on time series to be evaluated with very different mean values, as introduced by Regattieri *et al.* (2005).

$$MAD / A = \sum_{t=1}^{T} \frac{|F_t - D_t|}{T} \Big/ A \tag{8}$$

- *$MSE/A^2$:* Represents the arithmetic Mean of the sum of the Squares of the forecasting Errors (*MSE*), divided by the squared average demand size. Low values of $MSE/A^2$ address the adoption of forecasting approaches with a high incidence of low errors between true values and estimated ones. Otherwise, high $MSE/A^2$ indicates that high errors sometimes occur. Specifically, the ratio with $A^2$ is proposed again in order to compare values obtained in series characterised by consistent differences in the mean demand size.

$$MSE / A^2 = \sum_{t=1}^{T} \frac{(F_t - D_t)^2}{T} \Big/ A^2 \tag{9}$$

### 4.2 Preliminary analysis of data

Two monthly time series, i.e. Item1 and Item2, are collected from a real manufacturing environment, with a length of one hundred and twenty time periods (time series length of ten years).

Item 1 and Item 2 are characterised by irregularity and sporadicity. The former concerns the variability of the demand sizes, while the latter is related to the presence of frequent time periods in which demand does not occur. Therefore, two coefficients are computed (*CV* and *ADI*) in accordance with the definitions reported in Willemain (1994). Specifically, *CV* represents the coefficient of variation of non-null demands, while *ADI* represents the average number of time periods between two successive non-null demands. Hence *CV* and *ADI* establish the mark respectively of demand sizes' irregularity and of the sporadicity of demand pattern. Alternatively, in accordance with definitions reported in Syntetos (2001), $CV^2$ can be computed, that is the squared version of *CV*. Table 1 reports the values of *ADI* and *CV* along with the average demand size for both time series.

Furthermore, following the guideline introduced by Croston (1972), the demand is split into its two subcomponents, i.e. the demand sizes and the intervals between non-null demands. Then, the best distribution functions (*ddp*) that fit each of the series of subcomponents obtained are attained by using the software AutoFit®. It indicates a list of the best fitting distribution functions, presenting them in descending order of preference. In table 1 both *ddp* functions (i.e. *ddp* describing the series of non-null demand size and *ddp* describing the series of intervals between non-null demand sizes) are reported for each time series.

Finally, selected series present trend and seasonal components in the demand patterns, as shown in the bottom row of table 1.

Table 1: Statistical analysis of selected time series

| | Item1 | Item2 |
|---|---|---|
| Average demand size | 3.43 | 2.83 |
| *ADI* | 1.35 | 1.50 |
| *CV* | 0.74 | 0.68 |
| *ddp* demand sizes | Inv. Gauss. | Cauchy |
| *ddp* intervals | Chi-squared | Chi-squared |
| Further characteristics | Trend and seasonality | Trend and seasonality |

### 4.3 Experimental results

Even if, as often occurs in industrial practice, selected items do not satisfy the hypothesis of CR and SBA forecasting approaches, they are selected for predicting future values of the demand patterns, in accordance with preceding experimentations reported by the authors themselves. Subsequently, the TS tool is adopted, in order to compare the results obtained.

The application of CR and SBA methods is preceded by the choice of smoothing parameters. In an intermittent demand context, low smoothing constant values are recommended in literature. Smoothing constant values in the range of 0.05 - 0.2 are viewed as realistic

(Croston 1972; Willemain *et al*. 1994; Johnston and Boylan 1996). In this case study, as well as in Syntetos and Boylan (2005), four values are simulated: 0.05, 0.10, 0.15, and 0.20. Finally, the choice of smoothing parameters is based exclusively on the forecasting performance.

Table 2 and table 3 respectively report the forecasting errors obtained by the three aforementioned methods for the two selected items. For each accuracy measure evaluated and for each number of time periods, the lower forecasting error is underlined. Note that forecasting performance tends to improve for each forecasting method when the number of time periods ahead is increased. Figure 1 depicts the histograms of the forecasting errors for Item1 and Item2, in terms of both MAD/A and MSE/$A^2$

Despite how little data has been tested, some guidelines can be extrapolated. Firstly, the SBA method outperforms the CR method six times out of eight, and thus it appears to be an improvement in respect of Croston's method, even if the theoretical assumptions required for their implementation are violated. Moreover, whilst SBA reaches good performance in item 1 forecasting, TS achieves satisfactory results by analysing item 2, where consistent improvements are registered, especially when the number of forecasting time periods tends to increase (i.e. 6 months and 12 months).

Hence, further research is addressed in order to publish consistent experimentation comparing SBA and promising TS in real life series, representing a wide range of real-life occurrences.

## 5. CONCLUSIONS

This paper focuses on the comparison of alternative forecasting methods in case of irregular and sporadic demand patterns with seasonal components. Despite the great amount of contributions from literature on forecasting methods, this field needs further investigation due to its high criticality.

In particular, three alternative forecasting methods are compared, i.e. Croston's and Syntetos-Boylan' approaches and an automatic procedure called TRAMO-SEATS that is based on the best SARIMA model identification and application for forecasting aims. While the former methods represent *ad hoc* forecasting methods for irregular and sporadic demand patterns, the latter is a statistics software based on SARIMA modelling, which is robust and a forecasting method for general purposes.

Gamberini *et al*. (2010) obtain significant results in applying the SARIMA method in case of irregular and sporadic demand patterns with trend and/or seasonal components. In this paper a further investigation about the effectiveness of such a method is proposed. In particular, TRAMO-SEATS represents an automatic procedure for both identifying and applying the best SARIMA method found, and therefore it can be usefully applied in real industrial environments with satisfactory results.

Further research focuses on the implementation of wide experimentation for exploring TS potentials in the field of forecasting irregular and sporadic demand patterns with seasonality. Specifically, TS appears promising, since in the proposed case study it achieves results comparable with those attained by SBA, while coupling low implementation time, given its automatic behaviour and capability of jointly testing a wide range of SARIMA.

## REFERENCES

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

Bowerman, B. L., O'Connell, R. T., 1993. Forecasting and Time Series: An Applied Approach. *Duxbury Press*, Belmont, CA.

Box, G.E.P., Jenkins, G.M., 1976. *Time series analysis: Forecasting and control.* Rev. ed. Holden-Day, San Francisco.

Burman, J. P., 1980. Seasonal adjustment by signal extraction. *Journal of the Royal Statistical Society*, 143, 321–337.

Croston, J. D., 1972. Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, 23, 289–303.

Gamberini, R., Lolli, F., Rimini, B., Sgarbossa, F., 2010. Forecasting of sporadic demand patterns with seasonality and trend components: An empirical comparison between Holt-Winters and (S)Arima methods. *Mathematical Problems in Engineering*, doi:10.1155/2010/579010.

Gómez, V., Maravall, A., 1996. Programs TRAMO and SEATS. Instructions for the User (with some updates). *Working Paper 9628*, Servicio de Estudios, Banco de España.

Hillmer, S. C, Tiao, G. C., 1982. An ARIMA-model-based approach to seasonal adjustment. *Journal of the American Statistical Association*, 77, 63–70.

Johnston, F. R., Boylan, J. E., 1996. Forecasting for items with intermittent demand. *Journal of the Operational Research Society*, 47, 113–121.

Jarrett, J., 1991. *Business Forecasting Methods*. Blackwell, Oxford.

Makridakis, S. G., Wheelright, S. C., Hyndman, R. J., 1997. *Forecasting: Methods and applications*. John Wiley & Sons, Inc.

Makridakis, S. G., 1993. Accuracy measures: Theoretical and practical concerns. *International Journal of Forecasting*, 9, 527–529.

Maravall, A., 2006. An application of Tramo-Seats automatic procedure; direct versus indirect adjustment. *Computational Statistics & Data Analysis*, 50, 2167–2190.

Pollock, D. S. G., 2002. A review of TSW: the windows version of the Tramo-Seats program. *Journal of applied econometrics*, 17, 291–299.

Rao, A., 1973. A comment on Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, 24, 639–640.

Regattieri, A., Gamberi, M., Gamberini, R., Manzini, R., 2005. Managing lumpy demand for aircraft spare parts. *Journal of Air Transport Management*, 11, 426–431.

Schwarz, G. E., 1978. Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464.

Syntetos, A. A., Boylan, J. E., 2001. On the bias of intermittent demand estimates. *International Journal of Production Economics*, 71, 457–466.

Syntetos, A.A., Boylan, J.E., 2005. The accuracy of intermittent demand estimates. *International Journal of Forecasting*, 21, 303–314.

Willemain, T. R., Smart, C. N., Schwarz, H. F., 2004. A new approach to forecasting intermittent demand for service parts inventories. *International Journal of Forecasting*, 20, 375–387.

Table 2: Results achieved in terms of MAD/A

| | 1 month | | | 3 months | | | 6 months | | | 12 months | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CR | SBA | TS | CR | SBA | TS | CR | SBA | TS | CR | SBA | TS |
| Item1 | 0.90 | 0.88 | 1.03 | 0.65 | 0.72 | 0.89 | 0.72 | 0.72 | 0.89 | 0.17 | 0.05 | 0.07 |
| Item2 | 0.65 | 0.64 | 0.65 | 0.36 | 0.34 | 0.35 | 0.35 | 0.27 | 0.15 | 0.35 | 0.27 | 0.05 |

Table 3: Results achieved in terms of $MSE/A^2$

| | 1 month | | | 3 months | | | 6 months | | | 12 months | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CR | SBA | TS | CR | SBA | TS | CR | SBA | TS | CR | SBA | TS |
| Item1 | 1.03 | 1.01 | 1.68 | 0.65 | 0.63 | 0.88 | 0.58 | 0.57 | 0.80 | 0.03 | 0.003 | 0.005 |
| Item2 | 0.62 | 0.56 | 0.57 | 0.21 | 0.19 | 0.15 | 0.13 | 0.11 | 0.025 | 0.13 | 0.07 | 0.003 |

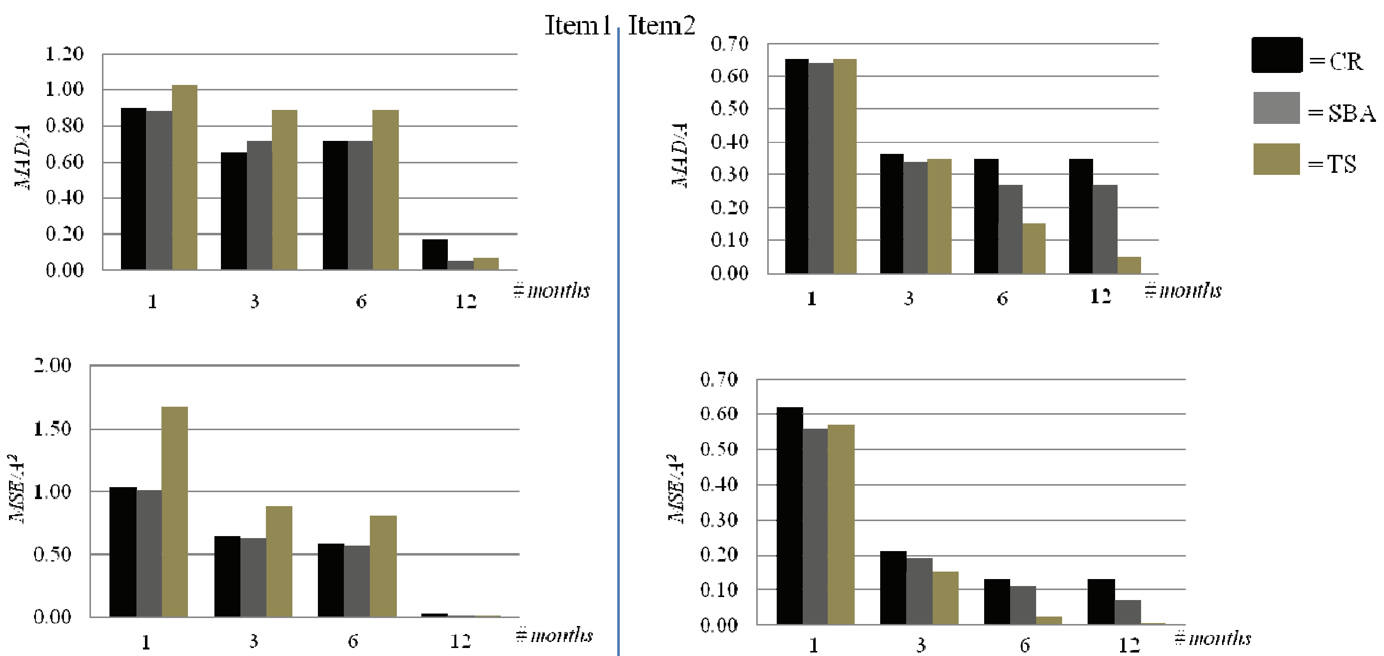


Figure 1: Forecasting errors in terms of MAD/A and $MSE/A^2$