
Biasedness of Forecasts Errors for Intermittent Demand Data

Submitted 11/09/20, 1st revision 15/10/20, 2nd revision 03/11/20, accepted 23/11/20

Mariusz Doszyna¹

Abstract:

Purpose: Intermittent demand is defined as infrequent or sporadic. Many forecasting errors are inappropriate for intermittent data. In some periods, there could be no demand, so division by zero must be avoided. Usually, forecasts are computed for many products; therefore, errors should be scale-independent (or relative). Many ex-post forecast errors, such as MASE (Mean Absolute Scaled Error) or MAE (Mean Absolute Error), indicate as best very low forecasts, sometimes even zero forecasts. Therefore, many researchers think that measures taking into account stock and consumer service levels should be used instead of conventional forecasts. It might suggest that typical forecast errors are useless for intermittent data. In this article, the contradictory hypothesis is verified. It is stated that only unbiased forecast errors should be used if the conclusions are to be correct.

Design/Methodology/Approach: Definition of unbiased forecast error is proposed and verified for popular forecast errors, such as ME (Mean Error), MSE (Mean Square Error), MAE, or MASE. The theoretical properties of these errors are considered concerning their biasedness. Forecasts are made based on Croston's and TSB methods, but also average and median were used as forecasting methods to emphasize conclusions.

Findings: In the empirical example, forecast errors are computed for intermittent demand times series to verify theoretical conclusions. The general conclusion is that only unbiased forecast errors provide proper indications according to forecast accuracy. This finding is true in general, not only for intermittent demand.

Practical Implications: Presented considerations might be useful for enterprises dealing with intermittent demand forecasting such as distribution centers, warehouse centers, and so on.

Originality/value: To the author's knowledge, forecast error bias was not analyzed before in the literature. A new forecast error is proposed, which was named RMSSE (Root Mean Square Scaled Error).

Keywords: Unbiasedness of forecasts errors, intermittent demand forecasting, RMSSE (Root Mean Square Scaled Error), Croston's method, TSB method.

JEL classification: C53, E27, L81.

Paper Type: Research study.

Acknowledgment: Project financed under the program of the Minister of Science and Higher Education in Poland under the name "Regional Initiative of Excellence" in the years 2019–2022, project number 001/RID/2018/19, financing amount 10,684,000.00 PLN.

¹Prof. Economics/Department of Econometrics and Statistics, Institute of Economics and Finance, University of Szczecin, Poland, mariusz.doszyn@usz.edu.pl

1. Introduction

Intermittent demand is particular and is usually defined as infrequent or sporadic. That kind of time series often consists of only a few demands and many zeros. Therefore, intermittent demand data require specific forecasting methods and special forecasting accuracy measures. Many popular forecast errors, such as MAPE (Mean Absolute Percentage Error), could not be applied because of zero division. For intermittent time-series, demand is often zero. Thus MAPE (and other percentage errors) are undefined. It is common knowledge that the best intermittent demand forecasting method indicates the forecast error used. To be more precise, such error measures like MASE or MAE favor methods yielding lower forecasts, sometimes even zero forecasts, which is hard to accept, especially stock management and consumer service levels.

This article states that unbiased forecast errors should be used to compare the accuracy of forecasting methods. Generally, biased forecasts errors favor underestimated or overestimated forecasts. Biasedness of the following forecast errors will be verified: ME, MSE, MAE, MASE, RMSSE. The last error – RMSSE (Root Mean Square Scaled Error) is an author's proposal. All the above-mentioned errors will be verified on a theoretical and empirical basis. In the empirical example, intermittent demand times series will be analyzed.

The article is organized in the following manner. In the second section literature review is presented. The methodological part definition of biasedness is presented, and the biasedness of analyzed errors is discussed about the proposed definition. Also, forecasting methods are shortly described. In the third section, forecasts for ten intermittent demand time series are computed, and forecast errors are estimated. Forecast errors are then evaluated concerning biasedness. In conclusion, unbiased forecast errors are pointed out, and future research directions are indicated.

2. Literature Review

A comprehensive review of forecast errors is presented in (Hyndman and Koehler, 2006). In this paper, forecasting error measures are divided into scale-dependent measures, percentage errors, relative errors, relative measures, and scaled errors. The most popular scale-dependent measures include Mean Error (ME), Mean Square Error (MSE), or Mean Absolute Error (MAE). Sometimes different variants of the above measures are proposed, where, instead of means, medians are calculated. That kind of error is robust concerning outliers. Scale-dependent errors are useless if forecasts errors for many products have to be analyzed. Each error has a different unit (or scale), so it is impossible to compare them.

Scale-independent is the percentage of errors. The most popular is the Mean Absolute Percentage Error (MAPE). Sometimes also symmetric Mean Absolute Percentage Error (sMAPE) is applied. However, the symmetry of sMAPE was questioned (Goodwin and Lawton, 1999; Koehler, 2001). These errors are inappropriate for

intermittent demand because of division by zero. In many periods intermittent demand is zero. Therefore percentage errors are just not defined.

Measures based on relative errors are also scale-independent. A relative error for a single period is a ratio of errors for analyzed and benchmark methods (Hyndman and Koehler, 2006). This class of errors Means Relative Absolute Error (MRAE), and Geometric Mean Relative Absolute Error (GMRAE) is recommended. These measures may compare different forecasting methods, but they are always tied with the benchmark method.

Sometimes also relative measures, which are quotients of given error measures (but not single errors) for two methods, are applied (Hyndman and Koehler, 2006; Syntetos and Boylan, 2005; Syntetos, 2001). They could be used for scale-dependent or percentage errors. It is emphasized that relative error measures based on geometric mean are robust to outliers (Syntetos, 2001). Those kinds of measures require at least two forecasts for the same series to compute a mean or a median. In the case of a single forecast, these measures become a relative error. In comparing forecasting methods also non-parametric alternatives are used, such as Percentage Better (PB) or Percentage Best (PBt) forecasts, where fractions of better (or best) forecasts are computed for a given method. However, these measures do not take error sizes into account, which could be misleading.

The above errors are scaled about out-of-sample values (values for an ex-post forecast horizon). In (Hyndman and Koehler, 2006), a new measure is proposed, Mean Absolute Scaled Error (MASE), scaled on an in-sample MAE from naïve forecasts. MASE is recommended for intermittent data (Hyndman, 2006). It is scaled, and it is easily applicable for intermittent demand series. However, two series with the same forecast errors should be noticed, but different in-sample values will have a different MASE, which might be confusing. Another widely known scaled measure is the MAE/Mean Ratio, a quotient of the MAE/ME, where all values come from an ex-post forecast horizon (out-of-sample values). The problem is that for intermittent data, ME is often close to zero, which makes the distribution of MAE/ME highly skewed.

As mentioned, error measures like MASE or MAE could indicate as best methods yielding low forecasts, sometimes even zero forecasts (Teunter and Duncan, 2009). An attempt to solve this problem was proposed in (Prestwich *et al.*, 2014), where mean-based measures are presented. In mean-based measures, forecasts are compared with point empirical values but with an in-sample mean (if there is stationarity). Many mean-based errors could be defined. The disadvantage of that kind of error is that the forecast horizon's actual values do not matter because in-sample means are used instead.

Interesting inventory-based measures like Cumulated Forecast Error (CFE), Number of Shortages (NOS), and Periods In Stock (PIS) are presented in (Wallström and Segerstedt, 2010). Generally, these errors simulate what would happen to fictitious

stock if its level would depend only on the given forecasting method. There are also many other proposals about intermittent demand forecast accuracy. In (Snyder, Ord, and Beaumont, 2012) and (Kolassa, 2016), it is suggested that predictive distributions should be evaluated instead of point forecasts. Many inventory-based measures are also used in (Syntetos, Babai, and Gardner, 2015), (Teunter and Duncan, 2009), (Engelmeyer, 2016).

The most popular intermittent demand forecasting method is Croston's method (Croston, 1972; Syntetos, 2001). It is based on exponential smoothing applied separately to demand size and demand intervals. Croston's method is biased and has some other drawbacks; therefore, some modifications were proposed, such as SBA (Syntetos–Boylan Approximation) (Syntetos and Boylan, 2005). Are SBA and Croston's methods, there are demand intervals that could be updated only in periods with non-zero sales. It often leads to overestimated forecasts if there are obsolesces. In the TSB method, sales probability that could always be updated is used instead of demand intervals (Teunter, Syntetos, and Babai, 2011). There are also other proposals dealing with obsolescence, such as hyperbolic exponential smoothing (Prestwich, Tarim, Rossi, and Hnich, 2014a). In intermittent demand, forecasting also SES (Simple Exponential Smoothing) or MA (Moving Average) are used, often as a benchmark (Syntetos, 2001). Simpler methods sometimes even give better results (Doszyń, 2019).

3. Methods

Before deciding which forecasting method to choose, it is necessary to specify a criterion for using the best one. Most often, the selection of the forecasting methods is based on one (or more) prediction errors. However, not all of them lead to identical conclusions.

Single forecast error could be expressed as:

$$e_t = x_t - \hat{x}_t \tag{1}$$

x_t - forecasted variable,

\hat{x}_t - forecast for the period t ,

$t = 1, 2, \dots, h$ - forecasted periods (forecast horizon).

In the whole article is assumed that in-sample periods are $t = 1, 2, \dots, n$ and forecasted periods are $t = 1, 2, \dots, h$, where n is a number of in-sample periods and h is a forecast horizon (number of forecasts). Also it is assumed that time series are stationary, but conclusions might be easily generalized to cases, when variables are functions of time.

Most forecast errors are based on mean error (\bar{e}_t), mean of squared errors ($\overline{e_t^2}$) or mean of absolute errors ($\overline{|e_t|}$). Forecasts are unbiased if $\bar{e}_t = 0$. In that case mean of actual values is equal to forecasts mean. Hence forecasts mean is on the same level as mean of the forecasted variable.

The question is if all forecasts errors are equally sensitive to forecasts biasedness? It can be assumed that not every error is the same from the point of view of biasedness. Before the forecasting method is chosen, therefore, it is necessary to assess the biasedness of forecasts errors, as these errors are common criterion for the choice of the forecasting method.

A forecast error that favors biased forecasts, i.e. overestimated or underestimated forecasts, will be classified as biased error. Below, definition of biasedness of forecasts error is proposed.

Definition

Forecasts error b is unbiased if it reaches the optimal (usually minimal) value for unbiased forecasts

Let's assume that we have actual x_1, x_2, \dots, x_h and forecasted values $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_h$ in analyzed periods $t = 1, 2, \dots, h$ and prediction error b , based on these values.

Error b is unbiased if it takes an optimal value b_{opt} (usually minimal) if forecasts are unbiased, what is true when $\bar{e}_t = 0$. In other words, forecast error b is unbiased if

$$b = b_{opt} \wedge \bar{e}_t = 0 \quad (2)$$

where

b - considered forecast error,

b_{opt} - the optimal (usually minimum) value of error b ,

\bar{e}_t - mean value of forecasts error.

Based on the above definition, it can be concluded that forecasts error can be considered unbiased if it reaches its optimal value for unbiased forecasts. Unbiased forecasts error increases as the forecasts biasedness increases. On the other hand, if a forecast error is biased, it can reach its minimum value for predictions that differ from the expected value of the variable being analyzed.

In the next step, selected popular forecasts errors will be analyzed from the point of view of their biasedness. Presented errors are often the basis of other forecasts errors, so the conclusions regarding them may also refer to other errors.

From the point of view of forecasts biasedness the primary error is ME (Mean Error):

$$ME = 1/h \sum_{t=1}^h e_t = \bar{e}_t \quad (3)$$

ME is an unbiased forecasts error. This error reaches an optimal value of zero when the average of predictions is equal to the average of the forecasted variable

$$ME = 0 \Leftrightarrow \bar{e}_t = 0 \quad (4)$$

As the forecasts biasedness increases, the (absolute) value of ME increases. ME can be negative, so in its case the optimal value (equal to zero) is not the same as the minimum value. For example, for the Mean Square Error (MSE) or the Mean Absolute Error (MAE), the optimal value is the minimum error value. To sum up, according to the proposed definition, ME error is an unbiased forecasts error.

MSE is probably the most popular forecasts error. It is the mean of squared errors

$$MSE = 1/h \sum_{t=1}^h e_t^2 = \overline{e_t^2} \quad (5)$$

MSE is an unbiased forecast error. According to the proposed definition of forecast error biasedness it could be noticed that

$$MSE = \min \Leftrightarrow \bar{e}_t = 0 \quad (6)$$

MSE could be decomposed by subtracting and adding the mean of actual values

$$MSE = E(x_t - \hat{x}_t)^2 = E(x_t - \bar{x}_t + \bar{x}_t - \hat{x}_t)^2 \quad (7)$$

Because $E[(x_t - \bar{x}_t)(\bar{x}_t - \hat{x}_t)] = 0$ then

$$MSE = E(x_t - \bar{x}_t)^2 + E(\bar{x}_t - \hat{x}_t)^2 = D^2(x_t) + (\bar{x}_t - \hat{x}_t)^2 \quad (8)$$

It can therefore be concluded that MSE is equal to the sum of the variance of predicted variable in the forecasts horizon $D^2(x_t)$ and square of forecasts biasedness $(\bar{x}_t - \hat{x}_t)^2$. Therefore MSE reaches its minimum when forecasts are unbiased. In that case MSE error is reduced to the variance of predicted variable. So the MSE is an unbiased forecasts error. If forecasts are biased, the variance $D^2(x_t)$ does not change and the MSE error increases due to biasedness by the factor $(\bar{x}_t - \hat{x}_t)^2$. Hence, MSE takes into account biasedness of forecasts.

It could be also shown by taking the first derivative (and checking the sign of the second) that $MSE = E(x_t - \hat{x}_t)^2$ is minimal when $\hat{x}_t = \bar{x}_t$, so forecast are equal to the mean of the forecasted variable. This confirms the unbiasedness of MSE.

In case of other errors based on MSE, for example Root Mean Square Error (RMSE), the conclusions are analogous, so RMSE is also an unbiased forecasts error. It also addresses other errors based on the mean of squared errors.

Previous forecasts errors are characterized by the lack of biasedness. This is not the case with the next error, i.e. the Mean Absolute Error (MAE). This error is the mean of absolute errors

$$MAE = 1/h \sum_{t=1}^h |e_t| = \overline{|e_t|} \quad (9)$$

MAE reaches a minimum if the forecasts are at the median level of the forecasted variable. For asymmetric distributions median is different than mean and in these cases MAE is a biased forecast error. MAE does not reach the minimum value for forecasts at the mean level, but only for forecasts at the median level, therefore, according to the proposed definition, it should be considered a biased forecast error.

To sum up, it could be noticed that for asymmetric distributions, when $M_e(x_t) \neq \bar{x}_t$

$$MAE = \min \Leftrightarrow \bar{e}_t \neq 0 \quad (10)$$

Similar conclusions relate to other errors based on absolute deviations.

Presented errors (ME, MSE, MAE) are scale-dependent, which might be problematic if forecasts for many products are to be at once evaluated. Therefore, scale-independent errors, such as MASE are often recommended:

$$MASE = \frac{1/h \sum_{t=1}^h |e_t|}{1/(n-1) \sum_{t=2}^n |x_t - x_{t-1}|} \quad (11)$$

where h is a number of forecasts and n is the number of in-sample periods.

MASE could be treated as a MAE in the forecasts horizon divided by in-sample MAE for naïve forecasts. If $MASE < 1$, then a verified method is better than naïve (in-sample) forecasts. MASE can not be computed only if all in-sample values are equal.

As mentioned, MAE is biased if median is different than mean and it also applies to MASE. For asymmetric distributions MASE should be treated as a biased forecasts error. MASE is biased because it is based on absolute forecasts errors.

Therefore, new measure is proposed that is similar to MASE, but is based on squared forecasts errors. It is called RMSSE (Root Mean Square Scaled Error):

$$RMSSE = \sqrt{\frac{1/h \sum_{t=1}^h e_t^2}{1/(n-1) \sum_{t=2}^n (x_t - x_{t-1})^2}} \quad (12)$$

The logic behind RMSSE is similar to that in MASE. Mean squared errors in the forecast horizon are divided by (in-sample) mean squared errors for naïve forecasts. If $RMSSE < 1$ then forecasts outperform naïve forecasts. It is a scaled error, always possible to calculate for intermittent data if not all in-sample values are equal. Moreover, RMSSE is unbiased, what is emphasized in the presented article.

Forecasts will be calculated, beyond others, by means of Croston's and TSB methods.

In Croston's method the demand size and demand intervals are updated only in periods with non-zero sale. If $x_t > 0$, then:

$$\hat{x}_t^+ = \hat{x}_{t-1}^+ + \alpha(x_t^+ - \hat{x}_{t-1}^+) \quad (13)$$

$$\hat{t}_t = \hat{t}_{t-1} + \beta(q_t - \hat{t}_{t-1}) \quad (14)$$

where

x_t^+ - demand size (non-zero sale),

\hat{x}_t^+ - smoothed demand size,

\hat{t}_t - smoothed demand interval,

q_t - number of periods since the last non-zero sale,

$\alpha, \beta \in \langle 0,1 \rangle$ - smoothing factors.

If $x_t = 0$, then $\hat{x}_t^+ = \hat{x}_{t-1}^+$ and $\hat{t}_t = \hat{t}_{t-1}$. Smoothed values are a relation of these two counterparts: $\hat{x}_t = \hat{x}_t^+ / \hat{t}_t$, hence the smoothed demand size is divided by the smoothed demand interval.

In the TSB method not demand intervals but sales probability is used.

If $x_t > 0$, then:

$$\hat{x}_t^+ = \hat{x}_{t-1}^+ + \alpha(x_t^+ - \hat{x}_{t-1}^+) \quad (15)$$

$$\hat{p}_t = \hat{p}_{t-1} + \beta(1 - \hat{p}_{t-1}) \quad (16)$$

where

\hat{p}_t - smoothed sales probability.

If $x_t = 0$, then:

$$\hat{x}_t^+ = \hat{x}_{t-1}^+ \quad (17)$$

$$\hat{p}_t = \hat{p}_{t-1} + \beta(0 - \hat{p}_{t-1}) \quad (18)$$

Smoothed demand is a product of adjusted demand size and sales probability: $\hat{x}_t = \hat{x}_t^+ \hat{p}_t$. In TSB method sales probability is updated in each period, what is better on account of obsolesces.

4. Empirical Results

In the research, ex-post forecasts were calculated for ten intermittent demand time series. These are weekly time series, where the first 205 weeks consist of in-sample values and the last 5 weeks are out-of-sample values, for which ex-post forecast errors were estimated. Analyzed data come from a company selling mostly tools and work clothes. Demand is identified with sales because, in the considered company, the consumer service level is almost one. Basic information about the considered time series are presented in the table below.

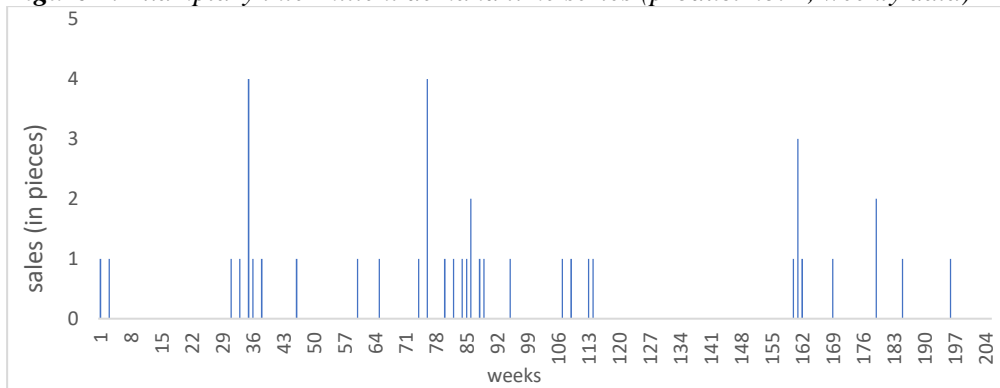
Table 1. Basic information about ten analyzed intermittent demand time series

Products	Number of observations	Sales frequency	Mean	Median	Max
1	198	0.35	1.00	0	14
2	205	0.15	0.20	0	4
3	205	0.27	0.60	0	7
4	205	0.18	0.20	0	2
5	190	0.12	0.20	0	7
6	15	0.20	0.20	0	1
7	205	0.16	0.20	0	4
8	165	0.19	0.20	0	2
9	110	0.41	0.80	0	10
10	30	0.20	1.80	0	17
Min	15	0.12	0.20	0	1
Max	205	0.40	1.80	0	17

Source: Own elaborations.

Sales were analyzed for the 205 weeks, but some products were introduced later. Hence the number of observations is in the range 15 - 205. Sales frequency is understood as a share of weeks with non-zero sales. It is visible that intermittence is present. Sales frequency was between 0.12 - 0.40 with a mean equal to 0.22, so on average, there was only one week for five with positive (higher than zero) sale. Mean sale for analyzed products was between 0.20 - 1.80; hence there were mostly slow-moving items. Each product's sales frequency was below 0.50, so the median for each item was equal to zero. In the case of intermittent demand, there are often outliers. Therefore maximum sales were also checked, for some product sales were indeed high. The highest weekly sale was equal to 17 pieces (product no. 10). The sales time series for the exemplary product (product no. 2) is presented in the graph below.

Figure 1. Exemplary intermittent demand time series (product no. 2, weekly data)



Source: Own elaborations based on the data from the analyzed company.

Four forecasting methods were applied: average, median (zero forecasts), Croston's (CR) method and TSB method.

Croston's and TSB methods were described in the methodological part. In each of these methods there are two smoothing factors, α and β . They were set at the level $\alpha = \beta = 0.1$. In case of Croston's and TSB methods first values were always taken as a starting ones.

The simple average was used as a forecasting method because, in the article problem of forecast errors, biasedness is considered about biasedness of forecasts. To remind, forecast error is defined as biased if it favors underestimated or overestimated forecasts. Forecasts calculated on the average level are unbiased. It is true for stationary times series, and intermittent times series usually have that property. Also, these are considered in the presented research.

Forecasts were also estimated on the median level. Because sales frequency for all ten products was lower than 0.50, the median was always equal to zero. Therefore forecasts on the median level are called zero forecasts. These forecasts are highly

underestimated, and they could be treated as an extreme case. If considered forecast error points zero forecasts as best, it will prove that this error is biased because it favors highly underestimated forecasts.

For these four forecasting methods (average, zero forecasts, Croston's, and TSB methods), five forecast errors were calculated: ME, MSE, MAE, MASE, and RMSSE. Their errors were described in the methodological part. The first three errors (ME, MSE, MAE) are scale-dependent, so they could not be directly compared (e.g., averaged). Therefore also scale-independent errors are considered (MASE, RMSSE). The results for all these errors are presented in the tables below.

Table 2. Scale-dependent errors for considered forecasting methods

No.	ME				MSE				MAE			
	Average	Zero forecasts	CR	TSB	Average	Zero forecasts	CR	TSB	Average	Zero forecasts	CR	TSB
1	0	1.00	-0.14	0.02	4.00	5.00	4.02	4.00	1.60	1.00	1.69	1.59
2	0	0.20	0.03	0.13	0.16	0.20	0.16	0.18	0.32	0.20	0.30	0.24
3	0	0.60	0.12	0.28	1.44	1.80	1.45	1.52	0.96	0.60	0.89	0.79
4	0	0.20	-0.02	-0.26	0.16	0.20	0.16	0.23	0.32	0.20	0.33	0.47
5	0	0.20	0.01	0.11	0.16	0.20	0.16	0.17	0.32	0.20	0.31	0.25
6	0	0.20	0.09	-0.15	0.16	0.20	0.17	0.18	0.32	0.20	0.26	0.41
7	0	0.20	-0.02	-0.04	0.16	0.20	0.16	0.16	0.32	0.20	0.33	0.34
8	0	0.20	-0.07	0.01	0.16	0.20	0.16	0.16	0.32	0.20	0.36	0.31
9	0	0.80	0.09	0.33	0.96	1.60	0.97	1.07	0.96	0.80	0.94	0.89
10	0	1.80	0.33	-0.04	5.76	9.00	5.87	5.76	2.16	1.80	2.09	2.17
Min	0.00	0.20	-0.14	-0.26	0.16	0.20	0.16	0.16	0.32	0.20	0.26	0.24
Max	0.00	1.80	0.33	0.33	5.76	9.00	5.87	5.76	2.16	1.80	2.09	2.17

Source: Own elaborations.

ME informs about forecast biasedness. The two extreme methods are average and zero forecasts. Averages are unbiased because each considered time series in-sample average was always equal to the out-of-sample average. This equality would sometimes not be true, but this equality is very probable for stationary time series, which often describes intermittent demand data. Hence in the presented example, an average could be treated as a completely unbiased forecasting method. For this method, ME is always equal to zero.

On the other extreme, there are zero forecasts, which were obtained as median sales values. This method is obviously highly underestimated and not acceptable due to the consumer service level. ME for zero forecasts was between 0.20 - 0.80; hence these forecasts are worst about biasedness. The remaining two methods (Croston's and TSB) were only slightly biased in both directions. Sometimes they were underestimated and sometimes - overestimated.

According to the verified hypothesis, ME and MSE are unbiased forecasts, errors, and MAE is a biased one. ME is unbiased "by definition"; it is always equal to zero for unbiased forecasts, which was already discussed. The absolute value of ME grows with biasedness of forecasts. In the methodological part, it was stated that MSE is unbiased. It is now visible because MSE is the lowest for average method and highest - for zero forecasts. For the average method, MSE is between 0.16 - 5.76, while for zero forecasts, it is between 0.20 - 9.00. The range for TSB is the same as for the average method. For Croston's method, MSE is between 0.16 - 5.87. Hence the results are very similar. Therefore MSE is the lowest for unbiased forecasts and highest - for (biased) zero forecasts.

Conclusions are quite different for MAE, which, according to theoretical properties, is a biased forecast error. MAE is the lowest for zero forecasts, for which it is between 0.20 - 1.80. MAE favors highly underestimated (zero) forecasts. Zero forecasts are on the median level, and MAE reaches a minimal value exactly for the median. According to MAE, the average method, which is unbiased, is much worse. For this method, MAE is in the interval of 0.32 - 2.16. Regarding MAE, TSB and Croston's methods are worse than zero forecasts, which is an unacceptable forecasting result. This example shows that mean and median are different; forecast errors based on absolute deviations favor biased forecasts. In the case of intermittent demand, distributions are highly positively skewed. Therefore errors like MAE indicate as best underestimated (even zero) forecasts. Therefore, in such cases, forecast errors should be based on squared (not absolute) errors.

Another desired property of forecasts errors for intermittent data is scale-independence. Two that kind of measures is presented in the table and graphs below.

MASE and RMSSE are useful when errors have to be compared for many products, which is often the case for intermittent demand forecasting systems.

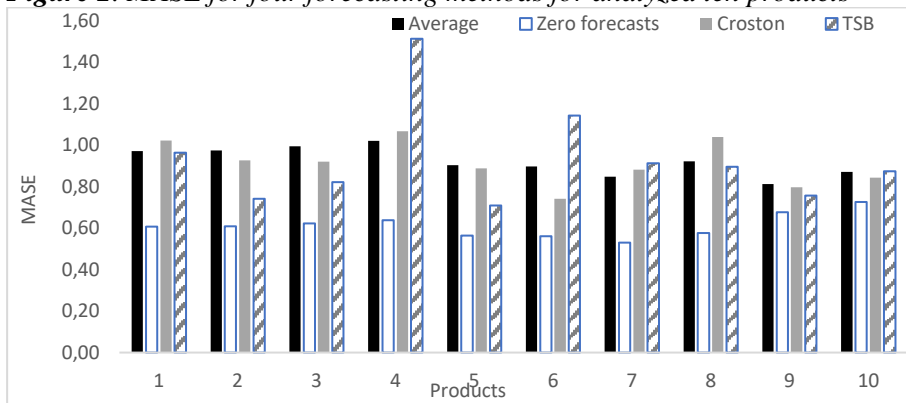
MASE, similarly as MAE, is based on absolute errors, therefore it is also biased. Also conclusions are similar as in case of MAE. Mean MASE is lowest for zero forecasts and much higher for average method. For TSB and Croston's methods mean MASE is almost the same as for average method. Hence MASE is best for zero forecasts, which is not surprising if biasedness of this error is taken into account.

Table 3. Scale-independent errors for considered forecasting methods

Products	MASE				RMSSE			
	Average	Zero forecasts	CR	TSB	Average	Zero forecasts	CR	TSB
1	0.97	0.61	1.02	0.96	0.68	0.76	0.68	0.68
2	0.97	0.61	0.93	0.74	0.52	0.59	0.52	0.55
3	0.99	0.62	0.92	0.82	0.65	0.73	0.66	0.67
4	1.02	0.64	1.07	1.51	0.65	0.72	0.65	0.77
5	0.90	0.56	0.89	0.71	0.43	0.48	0.43	0.45
6	0.90	0.56	0.74	1.14	0.67	0.75	0.69	0.71
7	0.85	0.53	0.88	0.91	0.52	0.59	0.52	0.53
8	0.92	0.58	1.04	0.90	0.65	0.72	0.65	0.65
9	0.81	0.68	0.80	0.76	0.50	0.64	0.50	0.52
10	0.87	0.73	0.84	0.87	0.48	0.60	0.49	0.48
Mean	0.92	0.61	0.91	0.93	0.58	0.66	0.58	0.60

Source: Own elaborations.

Figure 2. MASE for four forecasting methods for analyzed ten products

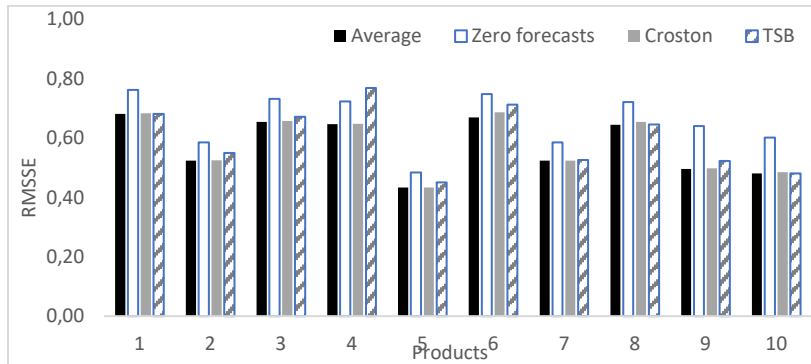


Source: Own elaborations.

RMSSE is an author’s proposal. The logic behind this measure is similar to that of MASE, but RMSSE is based on squared errors, so it is unbiased forecasts error. RMSSE is lowest for average method. It is also low for Croston’s and TSB methods and highest - for zero forecasts. That kind of conclusions are more reasonable with

regard to intermittent demand forecasting and especially - due to consumer service level.

Figure 3. RMSSE for four forecasting methods for analyzed ten products



Source: Own elaborations.

5. Summary and Concluding Comments

In the enterprise, it is often necessary to choose the best forecasting method to predict products' sales. Before that kind of decision, proper forecasts error should be settled. When the forecasted variable's distribution is symmetric, so the mean is equal to the median, all forecasts errors might be applied. They will provide consisted of conclusions. However, most economic variables are asymmetric, most often positively skewed. Intermittent demand is almost always highly positively skewed. In such cases, the median is lower than the mean, and errors based on absolute deviations favor biased (underestimated) forecasts. Measures based on absolute errors reach minimum value for median (not mean). In the case of intermittent demand, it may conclude that zero forecasts are the best. Underestimated forecasts are unacceptable due to the consumer service level. If sales frequency is below 0.50, then the median is always zero. Hence zero forecasts are the best about, e.g., MAE or MASE. That kind of error was classified in this article as biased forecasts errors. To avoid that kind of problem, measures based on squared errors ought to be applied. They are unbiased forecast errors because they reach the minimum for forecasts on the mean level.

To sum up, for asymmetric distributions for which the median is different from mean unbiased forecasts, errors should be used. In the case of intermittent demand data, usually forecasts for many products are calculated. Therefore scale-independence is also important. In the article, a new error fulfilling these requirements was proposed. It was named RMSSE (Root Mean Square Scaled Error).

In the future, research biasedness and other properties of different forecast errors will be verified, also in the context of intermittent demand data.

References:

- Croston, J.D. 1972. Forecasting and stock control for intermittent demands. *Operational Research Quarterly* (1970–1977), 23(3), 289-303.
- Doszyń, M. 2019. Intermittent demand forecasting in the Enterprise: Empirical verification. *Journal of Forecasting*, 38, 459-469.
- Engelmeyer, T. 2016. *Managing Intermittent Demand*. Wiesbaden, Springer.
- Goodwin, P., Lawton, R. 1999. On the asymmetry of the symmetric MAPE. *International Journal of Forecasting*, 15, 405-408.
- Hyndman, R.J. 2006. Another look at forecast accuracy metrics for intermittent demand. *Foresight: the International Journal of Applied Forecasting*, 4(4), 43-46.
- Hyndman, R.J., Koehler A.B. 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22, 679-688.
- Koehler, A.B. 2001. The asymmetry of the sAPE measure and other comments on the M3-competition. *International Journal Forecasting*, 17, 537-584.
- Kolassa, S. 2016. Evaluating Predictive Count Data Distributions in Retail Sales Forecasting. *International Journal of Forecasting*, 32(3), 788-803.
- Prestwich, S. D., Tarim, S. A., Rossi, R., Hnich, B. 2014a. Forecasting intermittent demand by hyperbolic–exponential smoothing. *International Journal of Forecasting*, 30, 928–933.
- Prestwich, S., Rossi, R., Tarim, S.A., Hnich, B. 2014. Mean–Based Error Measures for Intermittent Demand Forecasting. *International Journal of Production Research*, 52(22), 6782-6791.
- Snyder, R.D., Ord, J.K., Beaumont, A. 2012. Forecasting the intermittent demand for slow – moving inventories: A modelling approach. *International Journal of Forecasting*, 28, 485-496.
- Syntetos, A.A. 2001. Forecasting of intermittent demand. PhD Thesis, Buckinghamshire Chilterns University College, Brunel University, UK.
- Syntetos, A.A., Babai, M.Z., Gardner E.S. 2015. Forecasting intermittent inventory demands: simple parametric methods vs. bootstrapping. *Journal of Business Research*, 68, 1746-1752.
- Syntetos, A.A., Boylan, J.E. 2005. The accuracy of intermittent demand estimates. *International Journal of Forecasting*, 21, 303-314.
- Teunter, R.H, Syntetos, A.A., Babai, M. Z. 2011. Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research*, 214, 606-615.
- Teunter, R.H., Duncan, L. 2009. Forecasting Intermittent Demand: a Comparative Study. *Journal of the Operational Research Society*, 60, 321-329.
- Wallström, P., Segerstedt, A. 2010. Evaluation of forecasting error measurements and techniques for intermittent demand. *International Journal of Production Economics*, 128, 625-636.