A study of different Croston-like forecasting methods

[1]Anders Segerstedt, [1, 2]Erik Levén

[1] Industrial Logistics, Luleå University of Technology, Luleå, Sweden,
[2] Sveaskog, Kalix, Sweden

Abstract

An often-adopted technique for short-term forecasting is the single exponential smoothing (*SES*); it is available in most computer systems for material- and production control. With help from it, it is possible to update continually reorder points for an efficient inventory control. However, the ability for *SES* to forecast an item when the forecasting time periods often have zero demand is questioned; i.e. slow moving items or when demand is intermittent. Croston presented an idea and method to separate ordinary exponential smoothing in to two parts; the time between demand (withdrawals), and demand size. The forecasts then update only when there is a demand. Since then modifications of Croston's idea have been suggested and also an idea to measure the fraction of periods with zero demand compared to non-zero demand, i.e. techniques to handle intermittent demand. From the beginning four different suggestions to treat intermittent demand are tested. The tests showed that some complementary modifications were interesting to investigate. The different techniques are compared with Mean Squared Error (*MSE*), Cumulated Forecast Error (*CFE*) and with a new bias measure "Periods in Stock" (*PIS*). Our tests show that Croston's original idea may be to prefer; some techniques overestimate and others underestimate demand in certain circumstances; and one technique is not to prefer at all.

Keywords: Forecasting, Inventory control, Intermittent demand, Croston, Exponential smoothing

1. Introduction

An often-adopted technique for short-term forecasting is the single exponential smoothing (*SES*) from Brown (1959, 1962). The limited computational effort and low requirement of memory storage appropriate for yesterday's restricted computer capacity was an advantage compared to e.g. moving averages. The ability for *SES* to forecast an item when the forecasting time periods often have zero demand has been questioned; i.e. slow moving items or when demand is intermittent. Croston (1972) presented a method that separates the forecasts in to two parts; in time between withdrawals or demand and demand size. The forecasts update only when there is a demand. E.g. Willemain et al (1994) verify the usefulness of Croston's method

This paper studies, discuss and compares different forecasting techniques connected to the original idea of Croston. Syntetos and Boylan (2005) recommended an adjustment of the Croston method due to a systematic error notified by Syntetos and Boylan (2001). Levén and Segerstedt (2004) suggested a modification of the Croston method where a demand rate is directly calculated when a demand has happened. The main idea behind the modification in Levén and Segerstedt (2004) is that time between demand and demand size is not independent. However, this modification has shown poor results. Therefore, Wallström and Segerstedt (2010) suggest another modification, a "forward coverage" instead of a "backward coverage", a modification that is also tested here. Teunter et al (2011) suggest another method, a combination of updates every time period, for estimating the probability of a demand occasion, and every time a demand occurs for estimating the demand size.

All these techniques are tested and compared with the same real demand data. The data covers 18 months, but the specific with this data is that every item's withdrawal or demand is specified with the date (YYMMDD) and amount. This was just the type of data we were looking for; other data (from Sweden) were soon accumulated to month values, and thereafter accumulated to annual values.

Silver et al (1998) say to evaluate the performance of forecasts "no single measure is universally best". Nevertheless, evaluations of forecasting performances are sometimes done using only one measure of forecasting errors. Most common measures are Mean Absolute Deviation (*MAD*) or Mean Squared Error (*MSE*). Here we compare the different techniques with Mean Squared Error (*MSE*), Cumulated Forecast Error (*CFE*) and with the new bias measure "Periods in Stock" (*PIS*).

For a broader literature review about previous studies with similarities to this investigation we refer to Wallström (2009) and Wallström and Segerstedt (2010).

The paper has the following disposition; first in this section a short introduction. Section 2 describes the tested different forecasting techniques. In section 3, the different measures we use in the forthcoming analysis are presented. Section 4 presents the main and summary results of the study. Finally, in section 5 some conclusions from the tests are presented and discussed. Some related research and to 'forecast or not to forecast?' are also discussed.

Notations

X_t	Demand in period t
$\hat{X}_t$	Demand forecast in period t
α	Smoothing parameter, value 0-1
β	Smoothing parameter, value 0-1
t_n	Time period for the current or latest demand, $t_n = t$
t_{n-1}	Time period for the previous demand
T_t	$= t_n - t_{n-1}$, time interval between the latest and previous demand in period t
$\hat{T}_t$	Forecasted inter-demand intervals in period t
$\hat{d}_t$	Forecast of the demand rate in period t
N	Number of demand occasions, $N \le T$
$\dot{X}_n$	$= X_{t_n}$, demand in demand occasion n
$\dot{X}_{n-1}$	$= X_{t_{n-1}}$, demand in demand occasion n-1
$\hat{p}_t$	Estimated probability of demand for period t
T	Number of time periods, $T \ge N$, in our study T = 536

2. Forecasting methods

2.1 Single Exponential Smoothing (SES)

It is a technique applied in different fields, such as forecasting (cf. Brown (1959)), but it is a weighted average of previous outcome also applicable to process regulation (cf. Montgomery (2005)).

$$\hat{X}_{t+1} = \hat{X}_t + \alpha (X_t - \hat{X}_t) . \tag{1}$$

How often should the forecast be renewed, and should every individual item be renewed? With a high resolution of the forecast intervals, a short time period until the next calculation of a new forecast, the probability for periods with zero demand increases. If several zero demand periods will happen, the forecast will decrease and eventually approach zero. This scenario will happen when the items are slow-moving and with intermittent demand. Wallström (2009) and Wallström and Segerstedt (2010) show that the method with low smoothing parameter can still in some circumstances manage and "compete" with the methods especially designed for intermittent demand.

2.2 Croston

Croston (1972) presented a solution for slow-moving items. He suggests that the forecast shall be divided in two parts; one for the demand size and one for the inter-demand interval. The forecast recalculates only when there is a demand.

$$\text{If } X_t = 0\text{, then } \hat{X}_{t+1} = \hat{X}_t,\ \hat{T}_{t+1} = \hat{T}_t, \tag{2}$$

$$\text{If } X_t \neq 0\text{, then } \hat{X}_{t+1} = \hat{X}_t + \alpha (X_t - \hat{X}_t),\ \hat{T}_{t+1} = \hat{T}_t + \beta (T_t - \hat{T}_t), \tag{3}$$

where $T_t = t_n - t_{n-1}$.

The two exponential smoothing forecasts are then combined to estimate the mean demand per period length:

$$\hat{d}_{t+1} = \frac{\hat{X}_{t+1}}{\hat{T}_{t+1}} \tag{4}$$

2.3 Croston according to Syntetos Boylan (CrSyBo)

Syntetos and Boylan (2001) claimed that the procedure by Croston only show modest benefits in comparison to exponential smoothing in practical situations. The cause is explained due to the assumed distributions for time between withdrawal and amount of withdrawal, where Croston according to Syntetos and Boylan made an original mistake, which they correct with a bias correcting function the bias correction is based on a Taylor series expansion (Eaves and Kingsman, 2004).

Syntetos and Boylan (2005) suggested a modification to the *Croston* method. The modification can be described as a bias correcting function. In eq. (5) a bias corrector is added to the original *Croston*. The forecast updates are the same as for the original *Croston*:

$$\hat{d}_{t+1} = \left(1 - \frac{\beta}{2}\right)\frac{\hat{X}_{t+1}}{\hat{T}_{t+1}}. \quad (5)$$

Syntetos and Boylan modification of Croston's method is hereafter called *CrSyBo*.

2.4 Modified Croston (ModCr)

Levén and Segerstedt (2004) presented another modification of Croston's idea. Every time there is a demand, a new experienced demand rate is calculated. The update occurs when there is a demand, but maximum is once per time unit, day. If there are several demands during a time unit, the demands are added together. The demand rate is the quotient between the demand and the inter-demand interval:

$$If\ X_t = 0\text{, then } \hat{d}_{t+1} = \hat{d}_t \quad (6)$$

$$If\ X_t \neq 0\text{, then } \hat{d}_{t+1} = \hat{d}_t + \alpha\left(\frac{\dot{X}_n}{t_n - t_{n-1}} - \hat{d}_t\right). \quad (7)$$

(Where $\dot{X}_n = X_t$) A withdrawal every time period (day) transforms eq. (7) to eq. (1).

Levén and Segerstedt (2004) write: "The way we modify Croston's method avoids the bias Syntetos and Boylan have found (although we were not aware of their paper from the beginning)." We meant naturally the bias Syntetos and Boylan found in the derivation of *Croston*; not that *ModCr* never should show bias. Unfortunately in literature is found that Levén and Segerstedt claimed that *ModCr* avoids bias. What is bias and what is not bias; is mostly not entirely clear and requires proper analysis and discussion; it will be further treated in this paper. However, *ModCr*, except Levén and Segerstedt (2004), has shown unsatisfactory results with overestimation of demand (Boylan and Syntetos, 2007; Teunter and Sani, 2009; Wallström and Segerstedt, 2010).

The article, Levén and Segerstedt (2004), has received attentions; but most of the audience may not have noticed the idea and aim behind *ModCr*: time between withdrawal and amount of withdrawal are not independent; only one smoothing parameter to maintain; solve the problem of which forecasting interval to use, days; create a routine to update reorder points continuously. - The idea also wanted to collect and calculate an experienced demand rate or a takt time associated with the technique of Cover-Time Planning/Takt planning for materials requirement planning (Segerstedt, 1995 and 2017).

2.4.1 Forward Modified Croston (FModCr)

Wallström and Segerstedt (2010) in their tests discovered a difference, between the mean of the different *ModCr*'s demand rates and the mean demand rate for the whole time horizon. This may indicate that *ModCr* is wrongly designed; if the time between demands and demanded quantity not are independent then eq. (8) may model reality better than eq. (7):

$$\hat{d}_{t+1} = \hat{d}_t + \alpha\left(\frac{\dot{X}_{n-1}}{t_n - t_{n-1}} - \hat{d}_t\right) \quad (8)$$

The difference is that in eq. (8) the previous withdrawal is assumed to cover demand up to now and the new withdrawal covers future demand. If time between demands and demanded quantity are independent, then a construction like equation (7) or (8) is of less important; but if they are dependent, the construction is crucial. Eq. (7) assumes a "backward coverage", the new demand, or withdrawal covers a demand that has already been experienced, but eq. (8) assumes that the current withdrawal, or demand, will cover demand until the next withdrawal, a "forward coverage". Here is tested both assumptions.

2.5 Teunter, Syntetos, Babai (TeunterSB)

Teunter et al (2011) present a new idea to forecast intermittent demand. Every time period a probability for demand is updated; and a forecast for expected demanded quantity is updated only when there is a demand:

$$\text{If } X_t = 0\text{, then } \hat{X}_{t+1} = \hat{X}_t,\ \hat{p}_{t+1} = \hat{p}_t + \beta(0 - \hat{p}_t), \tag{9}$$

$$\text{If } X_t \neq 0\text{, then } \hat{X}_{t+1} = \hat{X}_t + \alpha(X_t - \hat{X}_t),\ \hat{p}_{t+1} = \hat{p}_t + \beta(1 - \hat{p}_t). \tag{10}$$

$$\hat{d}_{t+1} = \hat{X}_{t+1} \cdot \hat{p}_{t+1} \tag{11}$$

Silver et al (1998) discuss that a smoothing constant between 0.1-0.3 is mostly suitable for *SES* when forecasts are done monthly. If the forecast intervals are shorter, days, it is plausible that the smoothing constant is lower.

3. Forecast accuracy

3.1 Mean Square Error (MSE)

In this section, we present and discuss the different measures we use in the forthcoming analyses. Common measures for forecasting errors and its variability are *MSE* and Mean Absolute Deviation (*MAD*). Silver et al (1998) recommend the use of *MSE*, because *MSE* is related to standard variation of forecast errors. However, *MSE* is more sensitive to outliers and errors smaller than one due to the squared function. Which mean that in an evaluation of different forecasting methods *MSE* and *MAD* sometimes present a different result:

$$MSE = \frac{1}{T}\sum_{t=1}^{T}\left(X_t - \hat{X}_t\right)^2, \tag{12}$$

$$MAD = \frac{1}{T}\sum_{t=1}^{T}\left|X_t - \hat{X}_t\right|. \tag{13}$$

3.2 Cumulated Forecast Error (CFE)

MSE and *MAD* do not measure or reveal a systematic overestimate or underestimate of demand; if there is a systematic bias (error). A common measurement of bias is Cumulated Forecast Error (*CFE*). *CFE* is the cumulated sum of all forecast errors and CFE_t is the cumulated forecast error from period 1 to period *t*, and CFE_T the cumulated forecast error from period 1 to period *T*, i.e. the cumulated forecast error during the whole investigated time interval. If the forecast is unbiased, the *CFE* values should be close to zero. However, if CFE_T is zero it might also be due to “luck”; an earlier bias below zero might be covered with more recent errors above zero. To diminish this phenomena Wallström and Segerstedt (2010) suggest two additional *CFE* periods to be measured, namely where the maximum and minimum values occurs. $CFE_{\max}$ is equal to the greatest “shortage” during the forecast and $CFE_{\min}$ is equal to the greatest “surplus”. The reason for this interpretation is the definition of the forecast error. The forecast is subtracted from the actual demand and therefore a systematic overestimate of demand results in negative *CFE*-errors:

$$CFE_t = \sum_{i=1}^{t} (X_i - \hat{X}_i) = X_t - \hat{X}_t + CFE_{t-1}, \quad t = 1, 2, \ldots, T\,, \tag{14}$$

$$CFE_{\max} = \max_{t \in \{1, 2, \ldots, T\}} \left(CFE_t \right), \tag{15}$$

$$CFE_{\min} = \min_{t \in \{1, 2, \ldots, T\}} \left(CFE_t \right). \tag{16}$$

The quotient between CFE_t and MAD_t often serves as a tracking signal; it assumes the quotient has a normal distribution. To be able to trace whether a forecast is biased or not in a running situation, a tracking signal must be used.

3.3 Periods in Stock

“Periods in Stock” (*PIS*) measures the total number of periods the forecasted units of items have spent in fictitious stock or the number of fictitious stock-out periods. A period is equal to the length of the used time period. In this case, we assume a period is measured in days. Wallström and Segerstedt (2010) exemplify how the *PIS* work; assume there is a forecast of one unit per period during a total time period of tree time periods. The forecast is one unit for every period. If the demand is zero during all three periods, *PIS* in period 3 is equal to + 6. The item from day one has spent three periods in stock, the item from the second period have spent two periods in stock and the last item has spent one period in stock. In the beginning of the first period the one item is delivered to a fictitious stock; if there is been no demand during the first day, the result is plus one *PIS*. If then there is a demand of one in period 2 and a demand of one in period 3, the result is still plus one *PIS* in period 3. A positive number is a sign that the forecasting method tends to overestimate the demand. A negative number is a sign of underestimation of the demand because it shows *periods in/of shortages*. Therefore, for *PIS* the error subtraction is reversed, forecast minus demand, compared to CFE_t . PIS_t is the integration over time of CFE_t with a reversed sign due to the reversed forecast error; it measures not only the difference between forecast and outcome but also how long it takes to correct forecasting mistakes:

$$PIS_t = \sum_{i=1}^{t} (\hat{X}_i - X_i)(t - i + 1) = PIS_{t-1} + \sum_{i=1}^{t} (\hat{X}_i - X_i)\,, \tag{17}$$

$$PIS_t = PIS_{t-1} - CFE_t = -\sum_{i=1}^{t} CFE_i \ , \tag{18}$$

$$PIS_T = -\sum_{t=1}^{T} CFE_t \ . \tag{19}$$

In the following text when we discuss *CFE* and *PIS* and present figures with results, we mean the value in the end of the studied time interval CFE_T and PIS_T respectively.

4. Test results

The data comes from, for us, an anonymous company; the different data covers 18 months, 536 days to be exact. The information we have about the data is that it comes from a company in a European country and that our information provider later was told not to send out the data and keep it anonymous. We have here studied only 10 articles or items. The motive is to discover, and show, the variance and the difficulties that can be experienced in real demand data. The demand for some of the studied articles is shown in forthcoming figures (histograms) to present an overlook how the demand behave and looks like. All calculations and analysis has been done with Excel. In the shown figures, Excel has calculated the trend for the data, but no trend estimates, or corrections are used in the analyses of the different forecasting methods. None of the articles shows any clear trend in their demands.

The important feature of these data is that they contain daily consumption for 18 months. Usually a period of 18 month only contains less than 18 measures, because daily demand is just accumulated for the month and saved and stored as a month measure. Our aim was just to test forecasting techniques with daily updates. We have 536 periods with demand or no-demand. If we had had only monthly values, we would need a period of 44 years to get as many measuring points as we have in our data! (This also raises a question about the reliability and validity of previous published studies concerning forecasting error measurements and techniques for intermittent demand.)

Syntetos, Boylan and Croston (2005) presented an approach to categorise demand patterns in smooth, erratic, intermittent and lumpy; that are based on threshold values for the squared coefficient of variation (CV^2) for the demand and the average inter-demand interval (*p*). Kostenko and Hyndman (2006) modified the threshold values to 0.5 and 4/3. Most studied items are of the "worst" category lumpy ($CV^2 > 0.5, p > 4/3$) the rest are intermittent ($CV^2 \le 0.5, p > 4/3$) and none is erratic ($CV^2 > 0.5, p \le 4/3$) or smooth ($CV^2 \le 0.5, p \le 4/3$). According to Syntetos et al. *Croston* should be used within the smooth category and *CrSyBo* in every other category. The categorisation of the 10 items is presented in Fig. 1.

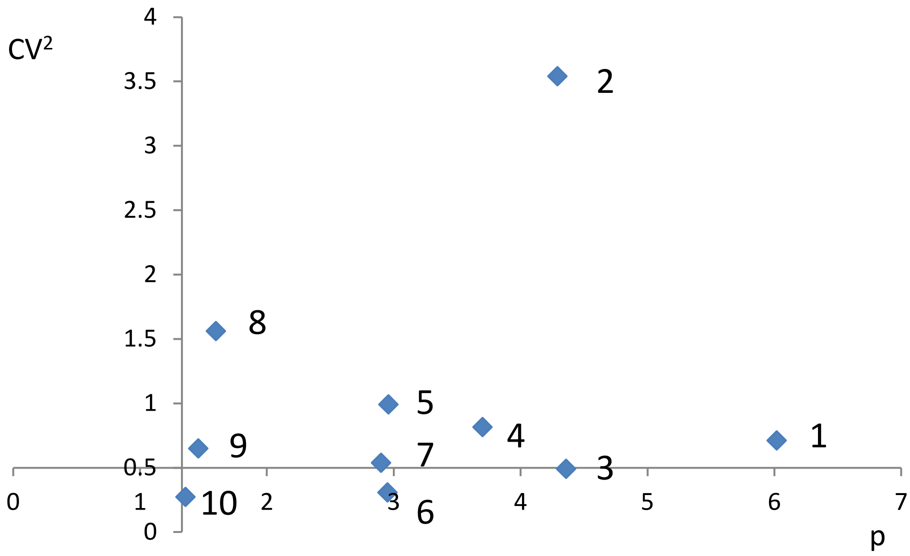

Figure 1. Categorisation of the 10 investigated items

Every studied item is presented with a heading containing: item number (1-10), Average, density, p =1/density and CV^2. - Average is $\sum_{t=1}^{536} X_t / 536$. Density is $N / 536$. Additionally is presented between which values the demand of the article varies ($X_t \in \{\min X_t, \max X_t\}$)

The start value for the forecasts in all investigations is the mean value ($\sum_{t=1}^{536} X_t / 536$) or derivations from the mean value. The results are dependent on the start values; but the main differences between the different forecasting methods presented below would probably not be overturned by different start values. Three different smoothing constants, 0.05, 0.15 and 0.30 are tested. A smoothing constant less than 0.05 means that more than 95% of the previous value is used for a forecast; and as we start with the mean value a smaller smoothing constant do not seem meaningful. Silver et al (1998) mean a larger value greater than 0.3 should raise the question of the validity of the underlying level in such a case a trend model is more appropriate. The trends Excel found for our studied articles are modest. A too large smoothing constant also seems unrealistic then maybe the previous outcome also is the best forecast.

The mean value, average, of the forecasts $\sum_{t=1}^{536} \hat{d}_t / 536$ for all the different forecasting methods with different parameters (0.05, 0.15, 0.30) are presented for all items together with Mean Square Error (*MSE*), Cumulated Forecast Error (*CFE*) and Periods in Stock (*PIS*). From a huge file of different items, the items were drawn manually and randomly; but the items were selected, and omitted, so a difference in “density” would appear.

The aim from the beginning was to test just *ModCr*, *F ModCr*, *CrSyBo* and *TeunterSB*; but the results trigged that also *Croston* must be included and new variants of *ModCr*.

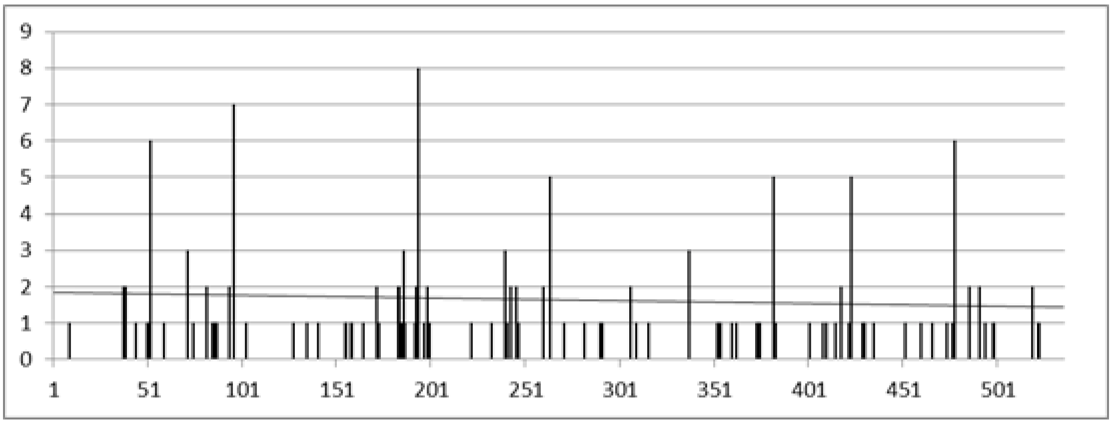

Figure 2. Demand pattern *item 1*

Item 1: Average 0.278, density 0.17, p =6.02, CV^2= 0.712, $X_t \in \{1,8\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**0.729**	0.97	-241.1	56326
	0.15	**0.803**	1.12	-281.0	74063
	0.30	**0.816**	1.27	-288.1	78687
F ModCr	0.05	**0.655**	0.90	-201.7	44283
	0.15	**0.708**	0.98	-230.2	58233
	0.30	**0.728**	1.08	-240.6	62379
Croston	0.05	**0.296**	0.72	-9.9	1676
	0.15	**0.326**	0.75	-25.5	6847
	0.30	**0.370**	0.80	-49.3	14004
CrSyBo	0.05	**0.289**	0.72	-5.9	602
	0.15	**0.301**	0.74	-12.4	3238
	0.30	**0.314**	0.77	-19.6	5713
TeunterSB	0.05	**0.259**	0.72	9.8	-5826
	0.15	**0.283**	0.73	-2.7	-657
	0.30	**0.291**	0.75	-7.3	1518

Table 1. Results *Item 1*

Item 1's demand pattern is shown in figure 2. The withdrawals, demands, are mostly 1 unit, but they alternate between 1 and 8. To exclude 6 or 7 outliers does not seem reasonable; these extreme values can be the result from a special customer ordering much more than the others. The histogram line is too wide compared to the time scale, it covers more than a time period, but the figure presents a picture of the demand pattern. The time unit used here is days, for *item 1* during this 536 days there is a withdrawal 89 days (or times); from this we calculate a ***density*** 89/536 = 0.17. (On average 17% of the time units has a withdrawal.) The mean value ($\sum_{t=1}^{536} X_t / 536$) for *item 1* is 0.278. These measures we show for every studied item; but no other item has a lower density than *item 1*.

A quick conclusion from table 1 and 2 is that both backward and forward *ModCr* overestimate demand; but forward seems better. *TeunterSB* needs two smoothing constants; the period constant used is 0.10 in all situations, in this analysis and also in the investigation of other items. Some experiments have been done with another value, β , but the most important smoothing constant seems to be α . An optimisation of β may create a better outcome.

Item 2: Average 1.205, density 0.23, p = 4.29, CV^2 = 3.54, $X_t \in \{1, 49\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**1.844**	21.07	-341.0	84594
	0.15	**2.093**	22.25	-474.2	117022
	0.30	**2.245**	24.17	-556.3	138192
FModCr	0.05	**1.608**	20.62	-214.8	41323
	0.15	**1.719**	20.94	-274.3	51765
	0.30	**1.728**	21.42	-279.	57456
Croston	0.05	**1.195**	20.17	6.1	9959
	0.15	**1.280**	19.74	-39.51	14539
	0.30	**1.374**	19.23	-90.28	24421
CrSyBo	0.05	**1.165**	20.54	22.1	5951
	0.15	**1.184**	20.83	11.9	2171
	0.30	**1.168**	21.27	20.2	-1797
TeunterSB	0.05	**1.181**	20.84	11.5	10657
	0.15	**1.220**	21.16	-9.3	10656
	0.30	**1.254**	21.75	-27.3	13561

Table 2. Results *Item 2*

Item 2's withdrawals, demands, are often 1 unit, but they alternate between 1 and 49. To exclude 49 as an outlier does not seem reasonable, even this extreme value can be the result from a special customer ordering much more than the others. Likewise, it is found for other items, so no extreme vales are excluded in the analyses. *Item 2* presents, like *item 1*, Forward *ModCr* overestimates demand less than backward *ModCr*. Like for *item 1 TeunterSB* shows a tendency to underestimate demand.

Item 3 behave in a different way than what will be shown by the other investigated items. (Forward) *FModCr* overestimates demand more than backward *ModCr*. *CrSyBo* and *TeunterSB* underestimate demand despite that the density is relatively low, cf. table 3. Wallström and Segerstedt (2010) discovered and claim that *CrSyBo* shows a tendency to underestimate demand when the demand has a higher density, i.e. when the demand is not so intermittent; but here the demand is rather intermittent, density is low. The demand does not show any exceptional trend or something else that can explicitly explain the divergent outcome, see figure 3.

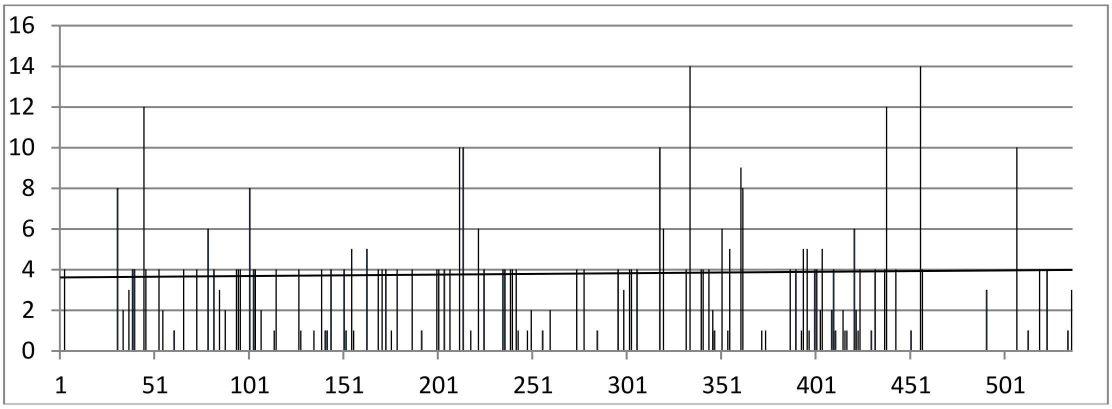

Figure 3. Demand pattern *item 3*

Item 3: Average 0.879, density 0.23, *p*=4.36, CV^2= 0.490, $X_t = \{1, 14\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**1.554**	4.91	-361.1	67744
	0.15	**1.698**	5.36	-438.8	87881
	0.30	**1.752**	5.82	-467.7	96887
FModCr	0.05	**1.674**	5.09	-425.4	89039
	0.15	**1.864**	5.87	-527.4	112869
	0.30	**2.008**	7.21	-605.4	123453
Croston	0.05	**0.957**	4.29	-31.3	2557
	0.15	**0.983**	4.39	-56.1	8837
	0.30	**1.045**	4.55	-89.3	18030
CrSyBo	0.05	**0.913**	4.28	-18.98	-712
	0.15	**0.909**	4.36	-16.6	-1440
	0.30	**0.888**	4.45	-5.3	-3904
TeunterSB	0.05	**0.869**	4.39	5.1	-523
	0.15	**0.870**	4.40	4.9	-1676
	0.30	**0.873**	4.43	2.8	-2693

Table 3. Results *Item 3*

Item 4: Average 1.026, density 0.27, *p*=3.70, CV^2=0.815, $X_t = \{1,18\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**1.856**	6.93	-443.2	88204
	0.15	**2.053**	7.96	-549.3	106690
	0.30	**2.150**	9.27	-601.4	114893
FModCr	0.05	**1.469**	6.18	-236.0	56483
	0.15	**1.599**	6.60	-305.6	64852
	0.30	**1.668**	7.14	-343.0	68884
Croston	0.05	**1.015**	5.97	6.4	6846
	0.15	**1.113**	6.17	-46.1	11253
	0.30	**1.248**	6.73	-118.6	22187
CrSyBo	0.05	**0.989**	5.97	20.0	3782
	0.15	**1.029**	6.13	-1.4	1731
	0.30	**1.061**	6.49	-18.3	1502
TeunterSB	0.05	**0.998**	5.99	15.4	2666
	0.15	**1.032**	6.12	-2.7	2368
	0.30	**1.066**	6.35	-21.0	3963

Table 4. Results *Item 4*

Both forward and *backward ModCr* overestimate demand, compared to ordinary *Croston*, *CrSyBo* and even *TeunterSB*. The difference between *ModCr* and the other techniques is that the individual withdrawal and the time since the last withdrawal influence more and create a high variation. Therefore, another experiment was also performed, to calculate both forward and backward *ModCr* in a different way to "average" the measures. Backward *ModCr* is modified according to equation (20) and *FModCr* according to equation (21); both just an average of the two last events:

$$\hat{d}_{t+1} = \hat{d}_t + \alpha\left(\frac{\dot{X}_n + \dot{X}_{n-1}}{t_n - t_{n-2}} - \hat{d}_t\right) \quad (20)$$

$$\hat{d}_{t+1} = \hat{d}_t + \alpha\left(\frac{\dot{X}_{n-1} + \dot{X}_{n-2}}{t_n - t_{n-2}} - \hat{d}_t\right) \tag{21}$$

We call it an average modification *avModCr* respective *avFModCr*. Table 5 presents the results of testing *item 5* also with these modifications and with *SES* and *Croston*. The overestimation decreases for both a backward and a forward assumption. A forward assumption is better but the backward assumption improves more in this case. Figure 4 shows the demand pattern of *item 5*. *AvModCr* and *avFModCR* are tested also for other items then *item5* and *item6*, and it diminishes the overestimation shown by *ModCr* and *FModCr*.

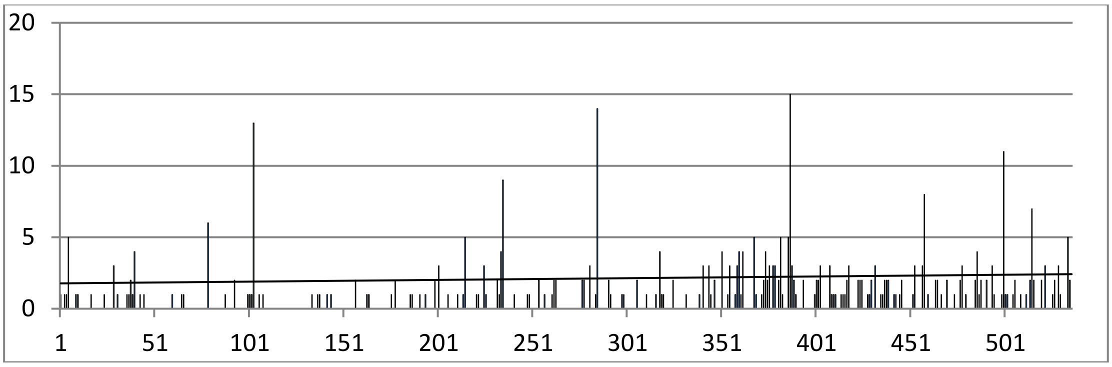

Figure 4. Demand pattern *item 5*

Item 5: Average 0.731, density 0.34, *p*=2.96, CV^2=0.991, $X_t = \{1,15\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**1.335**	3.01	-322.4	85242
	0.15	**1.448**	3.45	-383.1	105685
	0.30	**1.478**	3.99	-399.6	112146
avModCr	0.05	**1.063**	2.72	-177.4	55799
	0.15	**1.118**	2.92	-206.5	63947
	0.30	**1.139**	3.18	-218.3	67011
FModCr	0.05	**1.106**	2.74	-199.6	49062
	0.15	**1.183**	2.97	-240.8	58489
	0.30	**1.216**	3.21	-257.9	62534
avFModCr	0.05	**1.003**	2.69	-145.2	43257
	0.15	**1.052**	2.89	-171.4	47355
	0.30	**1.090**	3.17	-191.8	52436
Croston	0.05	**0.710**	2.57	11.7	4913
	0.15	**0.777**	2.63	-23.9	9867
	0.30	**0.855**	2.81	-65.6	21108
CrSyBo	0.05	**0.693**	2.57	21.2	2755
	0.15	**0.719**	2.62	7.3	3022
	0.30	**0.727**	2.73	3.0	5733
TeunterSB	0.05	**0.724**	2.55	4.2	1528
	0.15	**0.735**	2.60	-1.3	2213
	0.30	**0.742**	2.66	-5,2	3490
SES	0.05	**0.716**	2.58	8.4	156
	0.15	**0.726**	2.70	3.4	18
	0.30	**0.729**	2.90	1.9	4

Table 5. Results *Item 5*

Item 6: Average 0.819, density 0.34, *p*=2.95, CV^2= 0.308, $X_t = \{1,7\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**1.378**	2.18	-299.8	68988
	0.15	**1.349**	2.17	-284.4	76760
	0.30	**1.325**	2.24	-271.8	776830
FModCr	0.05	**1.394**	2.19	-308.3	70504
	0.15	**1.349**	2.15	-284.2	76888
	0.30	**1.315**	2.21	-266.7	76510
Croston	0.05	**0.905**	1.86	-46.7	4970
	0.15	**0.889**	1.87	-38.0	8839
	0.30	**0.912**	1.93	-50.2	14976
CrSyBo	0.05	**0.869**	1.86	-27.4	-2135
	0.15	**0.815**	1.85	1.9	-4228
	0.30	**0.772**	1.89	24.9	-8690
TeunterSB	0.05	**0.840**	1.89	-12.0	83
	0.15	**0.827**	1.89	-5.0	-165
	0.30	**0.820**	1.91	-1.2	-775
SES	0.05	**0.839**	1.86	-11.1	-211
	0.15	**0.823**	1.94	-2.6	-15
	0.30	**0.820**	2.12	-0.6	-1
avModCr	0.05	**1.133**	1.94	-168.4	36978
	0.15	**1.070**	1.91	-134.8	38397
	0.30	**1.034**	1.94	-116.1	37062
avFModCr	0.05	**1.144**	1.96	-174.2	36961
	0.15	**1.083**	1.93	-142.2	39122
	0.30	**1.048**	1.97	-123.3	38176

Table 6. Results *Item 6*

Item 7: Average 0.784, density 0.35, *p*=2.90, CV^2= 0.538, $X_t = \{1,11\}$

	alfa	Average	MSE	CFE	PIS
Mod Cr	0.05	**1.333**	2.44	-294.1	69516
	0.15	**1.355**	2.57	-306.2	76835
	0.30	**1.362**	2.81	-310.0	80215
FModCr	0.05	**1.275**	2.37	-263.1	64562
	0.15	**1.285**	2.46	-268.37	70962
	0.30	**1.302**	2.66	-277.6	77442
Croston	0.05	**0.821**	2.10	-20.4	877
	0.15	**0.828**	2.13	-24.0	5037
	0.30	**0.877**	2.25	-50.0	13642
CrSyBo	0.05	**0.801**	2.09	-9.4	-2451
	0.15	**0.766**	2.12	9.3	-5258
	0.30	**0.745**	2.19	20.5	-8239
TeunterSB	0.05	**0.785**	2.14	-0.9	154
	0.15	**0.778**	2.15	3.0	-855
	0.30	**0.779**	2.19	2.7	-992

Table 7. Results *Item 7*

Item 8: Average 2.146, density 0.63 p=1.60, CV^2 = 1.56, $X_t \in \{1, 41\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**2.821**	15.01	-360.8	94493
	0.15	**2.911**	16.34	-408.2	107605
	0.30	**3.020**	18.96	-466.2	120500
FModCr	0.05	**2.617**	14.75	-252.1	67379
	0.15	**2.663**	15.48	-276.4	75623
	0.30	**2.733**	17.07	-313.9	85557
Croston	0.05	**2.151**	14.37	-2.9	3172
	0.15	**2.232**	15.15	-46.0	13887
	0.30	**2.383**	17.02	-127.1	33546
CrSyBo	0.05	**2.097**	14.36	25.9	-5137
	0.15	**2.065**	15.02	43.7	-11843
	0.30	**2.026**	16.29	64.5	-20862
TeunterSB	0.05	**2.135**	14.39	0.2.0	1103
	0.15	**2.155**	14.94	-4.2	2467
	0.30	**2.196**	16.03	-20.7	5990
SES	0.05	**2.134**	14.46	6.4	127
	0.15	**2.141**	15.38	2.8	17
	0.30	**2.145**	16.94	0.5	2

Table 8. Results *Item 8*

Item 9: Average 2.104, Density 0.68 p=1.46, CV^2 =0.648, $X_t \in \{1, 25\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**2.610**	6.57	-270.2	69427
	0.15	**2.627**	6.81	-278.8	74584
	0.30	**2.655**	7.35	-292.7	77365
FModCr	0.05	**2.585**	6.60	-256.9	69562
	0.15	**2.606**	6.91	-267.2	76225
	0.30	**2.633**	7.58	-281.2	81282
Croston	0.05	**2.123**	6.01	-9.8	4469
	0.15	**2.163**	5.70	-30.2	10142
	0.30	**2.244**	5.45	-72.6	19486
CrSyBo	0.05	**2.070**	6.29	18.7	-3062
	0.15	**2.000**	6.50	56.7	-12876
	0.30	**1.907**	6.89	107.5	-27951
TeunterSB	0.05	**2.100**	6.36	2.9	733
	0.15	**2.098**	6.49	4.7	1063
	0.30	**2.103**	6.77	2.7	817

Table 9. Results *Item 9*

Item 10: Average 4.235, Density 0.74, p=1.36, CV^2= 0.273, $X_t = \{1,19\}$

	alfa	Average	MSE	CFE	PIS
ModCr	0.05	**5.149**	14.19	-489.9	112377
	0.15	**5.160**	14.87	-495.4	122979
	0.30	**5.187**	16.15	-509.4	127822
FModCr	0.05	**5.137**	14.23	-483.5	113917
	0.15	**5.159**	15.08	-494.9	127240
	0.30	**5.209**	16.67	-521.1	137880
Croston	0.05	**4.311**	13.35	-41.5	6787
	0.15	**4.353**	14.23	-63.8	17173
	0.30	**4.481**	15.88	-131.7	35312
CrSyBo	0.05	**4.203**	13.33	16.3	-9258
	0.15	**4.027**	14.12	111.3	-31740
	0.30	**3.809**	15.36	228.6	-65236
TeunterSB	0.05	**4.264**	13.57	-16.3	-436
	0.15	**4.236**	13.72	-0.7	-510
	0.30	**4.222**	14.08	7.0	-1933
SES	0.05	**4.258**	13.32	-12.7	-243
	0.15	**4.237**	14.10	-1.2	-7
	0.30	**4.233**	15.26	1.3	3

Table 10. Results *Item 10*

5. Conclusions, discussions and possible future extensions

From the limited items tested can be concluded that eq. (8) is often better than eq. (7). In many cases the assumption of a forward coverage intuitively makes more sense; a large withdrawal from inventory will cover future consumption. But our limited study shows that in some cases it can also be the right assumption; a large withdrawal is due to an expected or already experienced large demand. In the simulation study of Levén and Segerstedt (2004) the times between the demands and the demand size is independent, so forward or backward calculation was not crucial there.

In many cases in our study *CrSyBo* underestimate demand, it verifies the study of Teunter and Sani (2009); because *PIS* is negative. *TeunterSB* also for some items show a tendency to underestimate demand. During a time interval they forecast the total amount well, *MSE* is favourable, but with a time lagging, a bit too late, because *PIS* is negative. Therefore our study contradicts the suggestion that *Croston* should be used only within the “smooth category” (Syntetos et al., 2005). If the forecast technique is used for inventory control the underestimation is a more serious problem than overestimation. An underestimation will lead to lost sales and shortages. A forecast that suggest large inventories can be stopped and controlled by “days of inventory”, a large cover-time (inventory position divided by expected demand rate). Therefore, a risk adverse user of forecasting techniques for inventory control would prefer the original *Croston* method compared to the modification *CrSyBo*.

Even if *FModCr* sometimes seems better than *ModCr* it still overestimates demand too much, and it seems not a suitable method for forecasting. The different demand patterns (fig. 2, 3 and 4) and most other demand patterns show larger withdrawals that depart from other usual smaller withdrawals. As already discussed, such extreme values can be the result from a customer ordering much more than the others. Nevertheless, these extreme values seem to “destroy” *FModCr* and *ModCr* and a levelling should/must be done of the different withdrawals, as shown by *avFModCr* and *avFModCr*. A moving average of two withdrawals is probably not enough, tree would probably be better; but then we are back to the original *Croston*. We can conclude *ModCr* is not a good idea.

Tiacci and Saetta (2009) meant that most studies so far did not consider that the demand forecasting method which provides data to an inventory control system also interact with the control system. First attempts to address this issue are in works of Sani and Kingsman (1997) and Eaves and Kingsman (2004), according to Tiacci and Saetta (2009). Eaves and Kingsman (2004) evaluate forecasting methods with help of resulting inventory holding costs with specified service levels. How possible shortage is treated is not very clear. In their study *CrSyBo* is preferred, a possible explanation can be bias of underestimation found by Teunter and Sani (2009), Wallström and Segerstedt (2010) and in this paper.

Tiacci and Saetta (2009) made a simulation study and showed that traditional measures of forecast errors cannot be taken as singlehanded indicators for the choice among different demand forecasting methods; this is also the conclusion of Wallström (2009), Wallström and Segerstedt (2010). Ferbar Tratar (2010) showed and argued about the value of introducing also a stock control policy when choosing forecasting parameters. Syntetos, Nikolopoulos and Boylan (2010) state that when a forecasting method is used as an input to an inventory system it should always be evaluated with respect to its consequences for stock control through accuracy implications measure in addition to its performance on the standard forecast accuracy measures. Levén and Segerstedt (2004) is to the best of our knowledge the first paper to evaluate forecasting methods with number of shortages and average inventory, even if it is on a small scale. Levén and Segerstedt (2004) compared *ModCr* and *SES* by the average inventory and shortage volume created in a simulation study. The test in Levén and Segerstedt (2004) has a $CV^2 = \left(\sqrt{17.99}/4.93\right)^2 = 0.740$ and *p*=5; which explains the difficulty for especially *SES*. *PIS* is a bias measure supposed to simulate and imitate an inventory control situation; further studies will show how close it is to real inventory control measures.

From a practical view, the important thing is to be able to sale. Without something to sale with a satisfactory delivery time, there will be no income. This makes inventories necessary. A lack of inventory prevents income. The risk with a large inventory is that it cannot be sold, it will become obsolete. Therefore, “days of inventory” has become a key figure. Another risk with large inventories is lack of liquidity, no money to pay suppliers and employees. For a practician “inventory holding cost” is rather abstract.

Reading forecasting literature it is easy to get the opinion that everything, every item in an inventory should be forecasted. Nevertheless, it may not be necessary. Already in the 1970s, it was usual to signal items that should be deleted from the inventory and instead bought separately when a demand occurred. A “delete-rule” was defined as follows:

$W =$ Number of withdrawals or demands per year; $D =$ total demand per year; $Q =$ used order quantity for replenishment. An item was signalled for delete, and manual assessment, if $W < \max\left\{2, 2 \cdot D/Q\right\}$ (Segerstedt, 1976; but the idea came from a CEO directive from the company ASEA (later ABB).)

When it comes to pure spare parts, a ‘delete’-rule, like above, does not work; then special judgements must be taken. E.g. for a special roll-bearing 5 units must be kept in inventory in case the machine fails, because it is necessary to change all 5 and to prevent a long production downfall. Even a well-functioned forecasting method for intermittent demand will mostly fail to solve the original problem.

Experience shows that the inventories of many companies are filled with obsolete goods, which should be scrapped instead of taking up valuable storage space. Many items with intermittent demand are to be purchased separately if required and not stocked. - However, it will meet resistance to change tactics from stocking to purchase when needed. To write off, depreciate, stocked items will reduce the inventory value; this in turn will reduce the profit of the year. Top management can get nervous; it will perhaps reduce their bonus. A solution can be if the item has no withdrawal during one year to depreciate its value with 10 %; if the item has no withdrawal even the next year depreciate with additional 20 %; if the item has no withdrawal even the following year depreciate with additional 30 % etc. until the value is zero. (The company Benzlers used a similar depreciation rule already in the 1980s.)

This paper ends up with in favour of the original *Croston* forecasting method. From a practical point of view, it is inconvenient to decide what forecasting interval to use week or month, then the idea of updating the forecast every day when there has been a demand or withdrawal is appealing. The idea from *ModCr* with one smoothing parameter can be solved in *Croston* by using the same for both quantity and time. Some items have a high frequency of withdrawals and other a lower; is this a problem? Maybe, maybe not; can they have the same smoothing constant? To get an idea of how demand looks and develops, the number of withdrawals is the important information. Therefore the same smoothing constants ($(\alpha, \beta = \alpha)$ may be used despite different frequencies of withdrawal. Left for future studies, those items with high frequency may stand a smaller smoothing constant than those with lower frequencies. Some type of ‘aging’-function, can possibly solve this.

Suppose we have two forecasted items with the same positive PIS_t; item *i* exhibits several small demand and item *j* shows larger withdrawals with a lower frequency. Which item has the most wrong forecast? Which article has the most serious forecast error? Our assumption is that it is item *i*; with many observations we, or the system, has missed to correct the forecast. A common “tracking signal” for forecast errors is:

$$\frac{CFE_t}{MAD_t}, \text{ therefore we suggest a complement } \frac{1}{PIS_t / \hat{T}_t} \text{ when } PIS_t > 0 \qquad (22)$$

Petropoulos and Kourentzes (2015) suggest a scaled “periods in stock, we instead suggest:

$$sPIS_t = \frac{PIS_t}{\hat{d}_t \cdot t} \qquad (23)$$

Eq. (22) and (23), which item has the most serious forecast error, we leave and suggest for future studies.

Many companies have many items in inventory, many several thousands, Ahlsell 80 000 (Wallström, 2006). The companies must keep track of them and decide when to replenish and how much; therefore, a simple, easy to understand and maintain, efficient forecasting method is necessary to support the inventory control. How the forecast method is used to decide to replenish now, or wait; is also a problem to solve. With *SES* it is usually solved by an assumed service level and from that an estimated necessary reorder point. A direct calculated probability of stock outs can be an alternative (Segerstedt, 1994; Levén and Segerstedt, 2004); neural networks combining all (Kourentzes, 2013) another. It is important that the inventory control adapts to demand changes.

6. References:

Boylan J. E., Syntetos A. A. (2007). The accuracy of a modified Croston procedure. *International Journal of Production Economics*, 107 (2), 511-517.

Brown, R. G. (1959). *Statistical forecasting for inventory control*, New York: McGraw-Hill.

Brown R. G. (1962). *Smoothing, forecasting and prediction of discrete time series*. Englewood Cliffs, N.J.: Prentice Hall.

Croston J. D. (1972). Forecasting and Stock Control for Intermittent Demands, *Operational Research Quarterly*, 42 (3), 289-303.

Eaves A. H. C., Kingsman B. G. (2004). Forecasting for the ordering and stock-holding of spare parts. *Journal of the Operational Research Society*, 55 (4), 431–437.

Jonsson P., Mattsson S.-A. (2008). Inventory practices and their impact on perceived planning performance. *International Journal of Production Research*, 46 (7), 1787–1812.

Jonsson P., Mattsson S-A. (2014). *'Best practice' vid lagerstyrning i svensk industri* (in Swedish: Best practise inventory control in Swedish industry). Forskningsrapport, Chalmers Tekniska Högskola och Plan.

Kourentzes N. (2013). Intermittent demand forecasts with neural networks. *International Journal of Production Economics*, 143 (1), 198–206.

Kostenko A. V., Hyndman R. J. (2006). A note on the categorization of demand patterns. *Journal of the Operational Research Society*, 57 (10), 1256-1257.

Levén E., Segerstedt A. (2004). Inventory control with a modified Croston procedure and Erlang distribution. *International Journal of Production Economics*, 90 (3), 361-367.

Petropoulos F., Kourentzes N. (2015). Forecast combinations for intermittent demand. *Journal of the Operational Research Society*, 66 (6), 914-924.

Sani B., Kingsman B. G. (1997). Selecting the best periodic inventory control and demand forecasting methods for low demand items. *Journal of the Operational Research Society*, 48 (7), 700–713.

Segerstedt A. (1976). Användarhandledning för förrådsstyrningssystemet D05, Gränges Metallverken (in Swedish: User guide Inventory Management System D05). Västerås: Gränges Data.

Segerstedt A. (1994). Inventory Control with Variation in Lead Times, Especially when Demand is Intermittent. *International Journal of Production Economics*, 35, 365-372.

Segerstedt A. (1995). Cover-Time Planning, A method for Calculation of Material Requirements. *International Journal of Production Economics*, 41, 355-368, 1995

Segerstedt A. (2017). Cover-Time Planning/Takt Planning: A technique for materials requirement and production planning. *International Journal of Production Economics*, 194, 25-31.

Silver E. A., Pyke D. F., Peterson R. (1998). *Inventory Management and Production Planning* (3rd ed.). New York: John Wiley & Sons.

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., Croston J. D. (2005) On the categorization of demand patterns. *Journal of the Operational Research Society*, 56 (5), 495–503.

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

Syntetos, A. A., Nikolopoulos K., Boylan J. E. (2010). Judging the judges through accuracy-implication metrics: The case of inventory forecasting. *International Journal of Forecasting*, 26 (1), 134-143.

Teunter R., Sani B. (2009). On the bias of Croston's forecasting method. *European Journal of Operational Research*, 194 (1), 177-183.

Teunter R., Duncan L. (2009). Forecasting intermittent demand: a comparative study. *Journal of the Operational Research Society*, 60, 321-329.

Teunter R. H., Syntetos A. A., Babai M. Z. (2011). Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research*, 214 (3), 606-615.

Tiacci L., Sani B. (2009). An approach to evaluate the impact of interaction between demand forecasting method and stock control policy on inventory system performances. *Intenational Journal of Production Economics*, 118 (1), 63-71.

Ferbar Tratar, L., 2010. Joint optimisation of demand forecasting and stock control parameters. *International Journal of Production Economics*, 127 (1), 173-179.

Wallström, P., 2006. *Prognoser vid Ahlsell* (in Swedish: Forecasts at Ahlsell). Master thesis, Luleå University of Technology, Luleå (LTU-EX--06/188--SE).

Wallström P. (2009). *Evaluation of forecasting techniques and forecast errors: with focus on intermittent demand*. Luleå University of Technology, Licentiate thesis (ISBN 978-91-86233-63-1).

Wallström P., Segerstedt A. (2010). Evaluation of forecasting error measurements and techniques for intermittent demand. *International Journal of Production Economics*, 128 (2), 625-636.

Willemain T. R., Smart C. N., Shockor P. H., DeSautels P. A. (1994). Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. International *Journal of Forecasting*, 10 (4), 529-538.