Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper, eds.

SIMULATION OF STOCHASTIC ROLLING HORIZON FORECAST BEHAVIOR WITH APPLIED OUTLIER CORRECTION TO INCREASE FORECAST ACCURACY

Wolfgang Seiringer
Klaus Altendorfer

Dep. for Production and Operations Management
University of Applied Sciences Upper Austria
Wehrgrabenstraße 1-3
Steyr, 4400, AUSTRIA

Thomas Felberbauer

Department of Media and Digital Technologies
University of Applied Sciences St. Pölten
Matthias Corvinus-Str. 15
St. Pölten, 3100, AUSTRIA

ABSTRACT

A two-stage supply chain is studied in this paper where customers provide demand forecasts to a manufacturer and update these forecasts on a rolling horizon basis. Stochastic forecast errors and a forecast bias, both related to periods before delivery, are modeled. Practical observations show that planning methods implemented in ERP (enterprise resource planning) systems often lead to instabilities in production plans that temporarily increase projected demands. From the manufacturer's point of view, this behavior is observed as an outlier in the demand forecast values. Therefore, two simple outlier correction methods are developed and a simulation study is conducted to evaluate their performance concerning forecast accuracy. In detail, the magnitude of each demand forecast is evaluated and if a certain threshold is reached, the forecast is corrected. The study shows that the application of the outlier correction for forecast values leads to significant forecast accuracy improvement if such planning instabilities occur.

1 INTRODUCTION

Customer provided demand forecasts are often applied for production planning in order to create production orders for the respective materials. Intuitively it is obvious that a high forecast accuracy leads to a better production system performance as it reduces uncertainty. One of the main assumptions in literature is that information quality increase when the due date approaches. However, practical observations show that based on the customers' planning system, there might be temporary fluctuations in the demand forecast values that have a high magnitude which lead to instabilities in production plans with negative impact on production system performance. In the forecast data this behavior is measurable in form of numerical outliers. For example, the customer could occasionally increase the demand forecast for a certain due date from 800 to 1200 pieces and reduce the respective forecast some periods later to 800 again. For our investigations we consider rolling horizon forecast behavior. This means we are rolling through time and forecasts are provided for several periods into the future and updated each period. Such updated demand information is usually provided electronically by the customers and such data can be imported into the ERP (enterprise resource planning) system of the supplier after a data preprocessing step. Within the ERP system, the implemented planning method uses these forecast data to compute the production schedule. For planning the production steps the demand information is important, but when forecasts are not accurate and rely on outliers they introduce nervousness into production system with negative impact on production system performance. From the perspective of the outliers are often intuitively known from past experiences. Nonetheless, the production system of the supplier must absorb it with countermeasures like higher safety stock or additional capacity. To reduce this forecast introduced uncertainty within the production system we develop two simple outlier correction methods and extend the forecast generation process presented by

978-1-6654-3311-2/21/$31.00 ©2021 IEEE

(Zeiml et al. 2020). Our extension to this forecast generation process is the modelling of temporary outliers within the demand model and the development of two outlier correction methods. The simplicity and ease of use of the developed outlier correction methods supports their integration into, e.g. MRP based, production planning frameworks. In this paper the focus lies on the investigation of the performance and applicability of the proposed outlier corrections. One specific contribution of this paper is the extension of the demand model which mimics the rolling horizon evolution of forecasts with outliers. Based on our motivation the following research questions, which are answered with a simulation study, are formulated:

- RQ1: What is the performance of the developed demand forecast outlier correction methods if each update includes an information gain but some are disturbed by outliers?
- RQ2: What is the performance of the developed demand forecast outlier correction methods if some forecast updates are biased related to specific periods before delivery, i.e. not each update includes an information gain, and some updates are disturbed by outliers?
- RQ3: What is the effect of the outlier correction methods, their parameterization and different levels of uncertainty on their performance?

In general, the paper contributes to available literature in extending the available model (see (Zeiml et al. 2020)) to create demand forecasts applying a forecast evolution method by integrating the outlier creation. From a managerial point of view, RQ1 provides a manager the information which forecast correction model works well and what are the risks of applying such a simple forecast correction. The specific focus on forecast biases in RQ2 contributes to a better understanding of how such a bias influences the performance of the respective outlier correction. Since the outlier correction methods need to be parameterized, in RQ3 a sensitivity analysis concerning outlier correction parameters and forecast uncertainty is conducted. In the sensitivity analysis different levels of forecast bias and uncertainties are tested. The remaining paper is organized as follows. In section 2, we give an introduction to the related literature concerning forecast generation and outliers introduced by the customer. Section 3 describes our forecast generation model with a simple example followed by the definition of our outlier correction methods. In section 4, the simulation model is introduced and a numerical study is conducted in section 5 to answer the research questions. Finally, we conclude and give an outlook on our further research activities related to forecast processing.

2 LITERATURE REVIEW

Among other disturbances in demand information companies have to deal with unexpected high or low demand values, commonly known as outliers. Outliers in demand information can lead to unreliable and poor forecasts. Therefore, the identification of future outlier occurrence is an essential task in time series analysis to reduce the average forecasting error (see Chen and Liu 1993; Martínez–Álvarez et al. 2011). (Martínez–Álvarez et al. 2011) predict the occurrence of outliers in time series, based on the discovery of motifs. They assume motifs as pattern sequences preceding certain data marked as anomalous if data to be predicted as motifs, such data are identified as outliers, and treated separately from the rest of regular data. They use statistical methods to evaluate the accuracy of the proposed approach regarding the forecasting of the occurrence of outliers and their corresponding forecast values. (Ho et al. 2019) focus on the prediction of an anomaly to be available to raise an alert before an outlier occurrence happens. The authors mention that most of previous publications are limited to detect an outlier after its occurrence and not before. (Taylor and Letham 2018)developed a forecasting library called Prophet in Python and R for forecasting time series data based on an additive model where non-linear trends are fitted with yearly, weekly and daily seasonality, and special or irregular events (outlier) occurs. Their implementation is based on the model of (Harvey and Peters 1990). Special events are modeled independently as another additive factor of the used prediction model. Their implementation works best with time series that have strong seasonal effects and several seasons of historical data available. Prophet is robust to missing data and shifts in trend, and typically handles outliers well. Nevertheless, these models focus on the final order amounts and deal with outliers which occur in the final order amounts not in the evolution of forecasts.

Customer provided forecasts, which are commonly transmitted via electronic data interchange between customers and suppliers and which are regularly updated until the realization of the order at the specified due date, can be modeled according to the martingale model of forecast evolution (Heath and Jackson 1994). The papers of (Norouzi and Uzsoy 2014) and (Güllü 1996) integrate the above mentioned information of forecast evolution in inventory models to improve the production system performance. These authors focus on the development of analytical models, nevertheless also discrete event simulation can be an appropriate tool to investigate the influence of different demand model extensions and its parameterization on forecast accuracy, or to test the performance of developed forecast correction models. Especially discrete event simulation is suitable for the investigation of complex production systems. As one example in the simulation study of (Zeiml et al. 2019), customer provided forecasts are investigated with a discrete event simulation model. In detail the forecast accuracy of two customer provided forecast behaviors, i.e. independent forecast distribution and forecast evolution, are compared to simple moving average method (Svetunkov and Petropoulos 2018), which is a classical time series forecasting technique. After the discussion of the performance of different forecast error measures with respect to varying forecast error parameters the question is answered when customer provided forecasts are advantageous compared to the simple moving average method based on the final orders. An additional important finding of (Zeiml et al. 2019) is that the forecast error measure MPE is appropriate to show the systematic forecast error, i.e. forecast bias, with respect to periods before delivery. Motivated by this finding, a forecast error correction model to reduce the negative effects of forecast bias with focus on the forecast evolution is developed in (Zeiml et al. 2020). For scenarios with and without forecast bias, the correction model is evaluated. The authors find, that whenever there is a significant value of the systematic error, the correction model based on MPE improves forecast accuracy. Nevertheless, when the correction model is applied and there is only an unsystematic error the correction model is counterproductive.

On the one hand, the relevant papers above show, that current literature on outliers in demand data mainly focusses on the final order amounts and most of them neglect the evolution of forecasts. This means that forecasting methods to predict such outliers in final order amounts are developed or observable factors related to the outliers which are identified. One the other hand, literature focuses on the evolution of forecasts for different demand model extensions, but neglect the identification and correction of outliers to improve forecast accuracy. However, this paper addresses that outliers occur in the forecast evolution process and not in the final order amount. The combination of outlier detection and forecast evolution and the use of simulation to investigated the performance of the developed forecast correction models is the contribution of this paper.

3 DEMAND FORECAST MODEL AND OUTLIER CORRECTION

In this section, the demand forecast generation model extension to include outliers is derived and the outlier correction methods are developed.

3.1 Demand model and forecast process

To illustrate the demand model and the forecast process observed by the supplier, test data is generated and shown in Table 1. In this table, forecast data examples for four due dates, i.e. weeks, with forecast updates each week is illustrated. The table also shows forecast uncertainty and outlier behavior. In our demand model we assume to have a constant long term forecast and periodical updates until the due date approaches (see (Zeiml et al. 2020)). In detail, the customers start to update their forecasts for $j<H$ periods before delivery (PBD) with $H = 10$ being the forecast horizon. For $j \geq H$ we assume constant forecasts $\tilde{x}_i$ without updates. Based on Table 1, $x_{31,5} = 885$ means that for due date i = 31 (i.e. week 31), $j = 5$ weeks before delivery, the customer submitted a delivery request of 885 items. The information $x_{32,0} = 686$ means, that in week $i = 32$ and PBD $j = 0$ the supplier must deliver 686 items to the customer, i.e.

$x_{i,0}$ reflects the final order amount for week i after the last update. The time when the forecast information was transmitted can be calculated by $i - j$. For example for $x_{31,5}$ the respective demand forecast information was provided in week 26. Note that the week number 26 is shown as superscript value in Table 1.

Table 1: Simple forecast data example.

	Due dates in [weeks]				
	j \ i	29	30	31	32
	10…n	800^{19*}	800^{20}	800^{21}	800^{22}
PBD [weeks]	9	640^{20}	620^{21}	715^{22}	787^{23}
	8	750^{21}	640^{22}	820^{23}	1587^{24**}
	7	780^{22}	814^{23}	710^{24}	1670^{25**}
	6	860^{23}	525^{24}	633^{25}	870^{26}
	5	793^{24}	761^{25}	885^{26}	797^{27}
	4	737^{25}	583^{26}	583^{27}	738^{28}
	3	727^{26}	645^{27}	756^{28}	725^{29}
	2	745^{27}	677^{28}	744^{29}	644^{30}
	1	719^{28}	772^{29}	750^{30}	616^{31}
	0	784^{29}	773^{30}	873^{31}	686^{32}

*denotes the week of arrived forecast information, **outlier appears PBD 8 and disappears PBD 6

In (Zeiml et al. 2020) a demand model for rolling horizon forecast updates has been introduced which is extended in this paper by an outlier generation to test the performance of the outlier correction. Specifically, only the original forecast evolution and the forecast bias from (Zeiml et al. 2020) are applied in this paper. The forecast process we rely on (Zeiml et al. 2020) is an extension of the forecast evolution process introduced by (Güllü 1996) and (Heath and Jackson 1994). In this model, each forecast update is modeled by adding a random term to the previous forecast amount. The basic version from (Zeiml et al. 2020) we use to compute our forecast values is illustrated in equation (1). Whereby $\varepsilon_{i,j}$, the random update term, is a normally distributed random variable with expected value $\beta b_j \tilde{x}_i$ and standard deviation $\alpha a_j \tilde{x}_i$. The long term forecast $\tilde{x}_i$ for the due dates i can be constant or variable. To parameterize different forecast generation scenarios, α and β are varied similar to (Zeiml et al. 2020). The variable α determines the level of random error independently of the periods before delivery j, however, the a_j values can be applied to configure an interrelation between periods before delivery j and the random updates. The variable β determines the level of forecast bias independently of the periods before delivery j, however, the b_j values can be applied to configure an interrelation between periods before delivery j and the forecast bias. The variables β and α can be directly compared between scenarios because they are unscaled. We summarize that α describes the unsystematic noise of the forecast and β identifies the forecast bias behavior of booking systematically too much or too little.

$$\begin{aligned} x_{i,j} &= x_{i,j+1} + \varepsilon_{i,j}(\tilde{x}_i, \beta, \alpha) \\ \varepsilon_{i,j}(\tilde{x}_i, \beta, \alpha) &= N(\beta b_j \tilde{x}_i, \alpha a_j \tilde{x}_i) \end{aligned} \tag{1}$$

The basic demand model according to equation (1) is extended in the current paper by a temporary outlier generation. This means that disturbances in the forecast data occur only temporarily. In Table 1, e.g., 8 periods before delivery, the demand forecast for period 32 increases by 800 pieces (from 787 to 1587) and 6 periods before delivery this increase disappears again (from 1670 to 870). Equation (2) shows how the outlier generation is modeled:

$$\begin{aligned} & x_{i,j} = x_{i,j+1} + \varepsilon_{i,j}(\tilde{x}_i, \beta, \alpha) + P_{i,j}(\gamma c_j)\lambda_{i,j}(\tilde{x}_i, \delta, e) - P_{i,j+v}(\gamma c_j)\lambda_{i,j+v}(\tilde{x}_i, \delta, e) \\ & \lambda_{i,j}(\tilde{x}_i, \delta, e) = N(\delta \tilde{x}_i, \delta e \tilde{x}_i) \\ & P_{i,j}(\gamma c_j) = 1 \text{ with Probability } \gamma c_j \text{ and } P_{i,j}(\gamma c_j) = 0 \text{ otherwise} \end{aligned} \tag{2}$$

This formulation implies that outliers are randomly generated with probability γc_j, whereby c_j represents the link to periods before delivery and γ is a scenario specific parameter leveling this probability. The amount of an outlier is normally distributed with mean $\delta \tilde{x}_i$ and standard deviation $\delta e \tilde{x}_i$, whereby δ is also scenario specific. As the outliers occur only temporarily, the parameter v indicates for how much periods an outlier occurs. In the example presented above, $\lambda_{32,8}(\tilde{x}_i, \delta, e) = 800$, $P_{32,8}(\gamma c_8) = 1$ and $v = 2$. Note that this formulation constrains the outlier generation to a fixed and predefined v value, i.e. outliers are random in their amount and occurrence probability but stay always for v weeks. This limitation is applied in the current study to simplify the scenario setup and discussion, but can easily be extended in further research.

3.2 Outlier Correction

To increase forecast accuracy, we try to detect outliers in the forecast data consisting of $x_{i,j}$ and correct them to smoothen the demand forecasts. The corrected forecast value is represented by $\hat{x}_{i,j}$. Correction is applied during the process of forecast generation for each $x_{i,j}$ for all $j < H$. Looking at the possibilities to correct the respective outliers in the forecast data stream shows that two decisions have to be taken here. Decision 1 is related to the identification of outliers. For this decision, we focus on the mean and variance of the final order amounts. Decision 2 concerns the value to which the forecast could be corrected, here either the mean order amount from the past is applied or the last forecast before the identified outlier is taken.

3.2.1 Outlier Correction Method 1 (M1)

In this method, outliers are identified based on the mean and variance of the final order amounts and the average order amount is used instead of the original forecast if an outlier is detected. Equation (3) shows the corrected forecast calculation:

$$\begin{aligned} & \hat{x}_{i,j} = \begin{cases} \mu_i(m) : x_{i,j} > F^{-1}(X, \mu_i(m), \sigma_i(m)) \\ x_{i,j} : \text{otherwise} \end{cases} \\ & \text{with} \\ & \mu_i(m) = \frac{1}{m} \sum_{k=i+1-m-j}^{i-j} x_{k,0}; \ \sigma_i(m) = \sqrt{\frac{1}{m-1} \sum_{k=i+1-m-j}^{i-j} (x_{k,0} - \mu_i(m))^2} \end{aligned} \tag{3}$$

With $F^{-1}(X,\mu,\sigma)$ being the inverse of the normal distribution at probability *X*. This means during forecast generation the $x_{i,0}$ are stored, the mean and standard deviation are computed, and the upper bounds are evaluated. This upper bound is compared to the current $x_{i,j}$ and the forecast is corrected to $\hat{x}_{i,j}$ if an outlier is presumed. Note that *m* is the number of data points applied, i.e. the representative final order amount history.

3.2.2 Outlier Correction Method 2 (M2)

In this method, outliers are identified similar to method 1, but the last forecast value for the respective due date (in M1 the average order amount is used) is used to replace the original forecast. Equation (4) shows the corrected forecast calculation:

$$\hat{x}_{i,j} = \begin{cases} \hat{x}_{i,j+1} : x_{i,j} > F^{-1}\left(X, \mu_i(m), \sigma_i(m)\right) \\ x_{i,j} : \text{otherwise} \end{cases}$$
$$\text{with} \tag{4}$$
$$\mu_i(m) = \frac{1}{m} \sum_{k=i+1-m-j}^{i-j} x_{k,0};\ \sigma_i(m) = \sqrt{\frac{1}{m-1} \sum_{k=i+1-m-j}^{i-j} \left(x_{k,0} - \mu_i(m)\right)^2}$$

Assuming that we have an unbiased forecast evolution demand model, method 2 should provide better results than method 1. However, we will also study the effects of biased forecasts on the performance. Note that $\hat{x}_{i,j} = \tilde{x}_i$ is applied for $j \geq H$.

4 SIMULATION MODEL FOR FORECAST GENERATION AND OUTLIER CORRECTION

To evaluate the performance of the developed outlier correction model, a discrete event simulation model built with AnyLogic© is used. As a discrete sequence of events (forecast updates) should be created, discrete event simulation model is identified as an appropriate solution method. In detail, only at specific periodic points in time, a change in the forecast is triggered in the simulation model. Our simulation model has implemented the forecast generation process described in Zeiml et al. (2020) and is extended with the already described outlier generation as well as the outlier correction methods M1 and M2 according to the formulation in Section 3.

Selecting the appropriate parameters for simulation is an extensive research field. We selected the simulation parameters for our numerical study based on preliminary tests with respect to demonstrate the usage and limitations of our developed outlier correction method. We decided to use a simulation runtime of 520 periods, which are 10 years, as simulation model time are weeks. Weeks are used as time periods as they are a common time frame to update forecasts. For the first 20 periods forecast history is not completely available, therefore the first 20 time periods are removed from result analysis. This results in 500 complete forecast streams for each due date and the full range of observed periods before delivery. As stochastic forecast behavior is assumed 20 replications per iteration are used.

The normalized $RSME_j$ (root-mean-squared-error) is used to measure forecast accuracy with respect to periods before delivery *j*. To enable a distinction between the uncorrected forecasts and the corrected forecasts, the notation $RMSE_j$ is applied for uncorrected forecasts and $CRMSE_j$ is applied for corrected forecasts. Equation (5) shows the specific $RMSE_j$ and $CRMSE_j$ calculation. Note that *n* is the number of forecast streams observed, i.e. number of due dates for which a full forecast history is available.

$$RMSE_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{ij} - x_{i,0}\right)^2} \Bigg/ \left(\frac{1}{n}\sum_{i=1}^{n} x_{i,0}\right);\ CRMSE_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{x}_{ij} - x_{i,0}\right)^2} \Bigg/ \left(\frac{1}{n}\sum_{i=1}^{n} x_{i,0}\right) \tag{5}$$

Additionally, a correction effectiveness indicator E_j is introduced in equation (6) to evaluate the performance gain (or loss) with respect to forecast accuracy when the forecast correction methods are applied. Whenever E_j is positive, the forecast correction leads to an increase in forecast accuracy.

$$E_j = \frac{RMSE_j - CRMSE_j}{RMSE_j} \tag{6}$$

5 NUMERICAL STUDY

Based on the previously described simulation setting and the above stated research questions, several scenarios to evaluate the performance of the developed outlier correction methods are defined. The respective results are discussed in this section.

5.1 Scenario definition

To test the performance of the outlier correction methods, the following basic scenarios A (unbiased) and B (biased) are identified. Based on these basic scenarios A and B, the parameters α, β, γ, and δ are changed to create a range of unbiased (A) and biased (B) forecast evolution scenarios. The basic scenario parameters are defined as:

- $a_j = 0.1$ for all j, i.e. randomness in forecast evolution stays constant with respect to periods before delivery
- $b_3 = b_4 = -0.1$, $b_5 = -0.2$, $b_6 = 0.2$, $b_7 = b_8 = 0.1$; i.e. a systematic overbooking occurs with a peak 6 periods before delivery (see Zeiml et al. 2020)
- $c_4 = c_7 = 0.5$, $v = 1$, $e = 0.25$; i.e. an outlier can occur 4 and 7 periods before delivery and stays for 1 period

Based on the basic scenario parameters the scenario sets A and B are defined as following:

- **Scenario set A**: *unbiased* forecast updates with outliers: $\alpha \in \{0.5,1,2\}$, $\beta = 0$, $\gamma \in \{0.5,1,2\}$, $\delta \in \{0.5,1,2\}$; i.e. 27 scenarios are tested. Basic case is $\alpha = 1$, $\beta = 0$, $\gamma = 1$, and $\delta = 1$.
- **Scenario set B**: *biased* forecast updates with outliers: $\alpha = 1$, $\beta \in \{0.5,1,2\}$, $\gamma \in \{0.5,1,2\}$, $\delta \in \{0.5,1,2\}$; i.e. 27 scenarios are tested. Basic case is $\alpha = 1$, $\beta = 1$, $\gamma = 1$, and $\delta = 1$.
- For both basic scenarios a comparison scenario without outliers is generated; i.e. basic a w/o: $\alpha = 1$, $\beta = 0$, $\gamma = 0$, and $\delta = 0$ and basic B w/o: $\alpha = 1$, $\beta = 1$, $\gamma = 0$, and $\delta = 0$.

The simulation study can be divided into two parts. In both parts the outlier correction methods are applied on the whole forecast stream. In the first part performance and correction effectiveness for the basic scenarios are tested. Therefore, the simulation is conducted for the basic scenarios Basic A and B with and without outliers (w/o), for $X = 0.9$ and $m \in \{6,24\}$. The second part which is the input for the sensitivity analysis, consists of 54 scenarios (27 for each scenario set A and B) with $X \in \{0.7,0.8,0.9,0.95,0.98,0.99\}$ and $m = 24$.

5.2 Correction method comparison for basic scenarios

In this section the results for unbiased and biased forecasts with and without outliers are presented for the basic scenarios A and B for both correction methods M1 and M2. As introduced above, the higher the $RMSE_j$ value the higher the level of forecast error is. For scenario *A Basic*, where outliers occur in PBD_7 and PBD_4, Figure 1 shows, that without correction the $RMSE_j$ is between 0.7 and 0.8 (for j=4 and j=7). The application of correction methods M1 and M2 leads in this case to a $RMSE_j$ reduction of more than 50%. Which indicates a good performance of the correction methods when an outlier occurs. Furthermore, Figure 1 shows that until the first application of the correction methods (*j*=10 to 8) the $RMSE_j$ is similar with and without outlier correction. However, looking at *j*=3 to 1, shows that the corrected forecasts have a higher $RMSE_j$ which implies that the outlier correction adds a source of uncertainty. This behavior can also be observed for scenario *B Basic* with outliers and forecast bias in Figure 3. Specifically the results in Figure 2 show, that without outliers, the correction methods start to overreact and introduce an additional error. When forecast bias and random error are present in a setting without outliers (see Figure 4), the outlier correction methods can partially diminish the negative effect of the forecast bias. That happens because the bias between PBD 8 and 4 is treated like outliers related to the final order amount. Concerning RQ1 and RQ2 the tested correction methods show a positive impact on the forecast accuracy and the information updates that are disturbed by outliers are smoothened. However, a negative note to RQ1 and RQ2 is the negative contribution to the $RMSE_j$ after the outliers disappear or if only random noise is present. Furthermore, outlier correction method M2 shows in these scenarios a consistently better performance than M1.

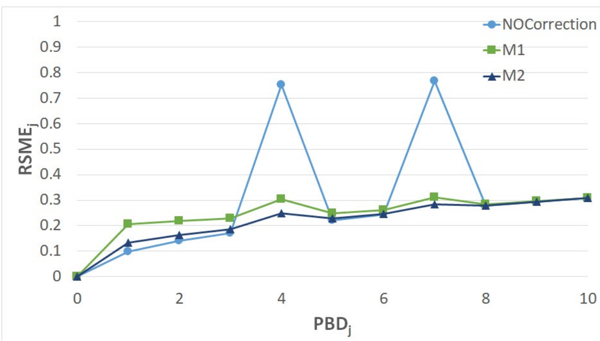

Figure 1: Scenario *A Basic* for *X*=0.9 and *m*=24.

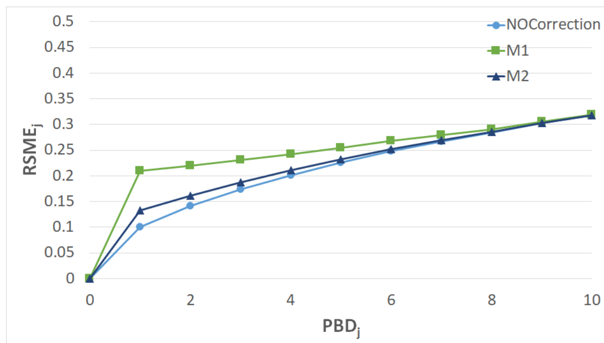

Figure 2: Scenario *A Basic (w/o)* for *X*=0.9 and *m*=24.

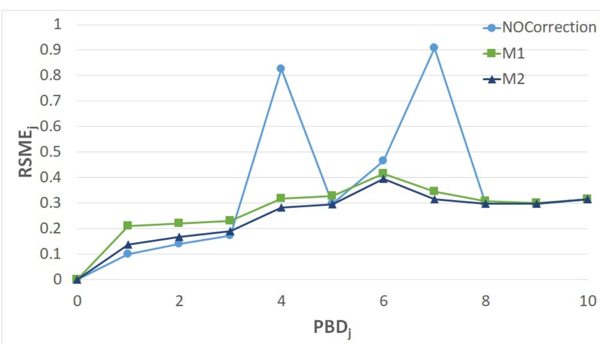

Figure 3: Scenario *B Basic* for *X*=0.9 and *m*=24.

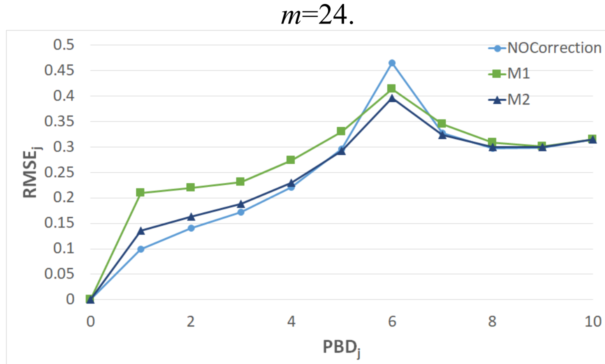

Figure 4: Scenario *B Basic (w/o)* for *X*=0.9 and *m*=24.

5.3 Performance decrease when no outliers occur

To further analyze the performance loss if no outliers occur and the correction is still applied, in this section an additional experiment with *m*=6 and *m*=24 is conducted for scenario *A Basic w/o* and scenario *B Basic w/o* for outlier correction method M2. Note that *m*=6 and *m*=24 show the influence of different lengths of forecast history, i.e. the value *m*=6 means we are using 6 instead of 24 historic order amount values to compute the outlier threshold. Figure 5 shows that without forecast bias, the correction effectiveness is negative for all $j < 10$ (m=24) and j d 10 (m=6) values and the added uncertainty increases with respect to decreasing j, i.e. more uncertainty is added when the due date approaches. Figure 6 shows that in general, the outlier correction is able to increase forecast accuracy in periods where the bias occurs, however, it adds uncertainty whenever no bias occurs. A specifically interesting finding is that a lower number of available forecast history streams leads to a higher uncertainty added. This can be observed by the fact, that in Figure 5 and Figure 6 the correction effectiveness of m=24 is always higher compared to m=6. This means an increase in m leads to an increase in the performance of the outlier correction method. Concerning RQ1 and RQ2 this shows that outlier correction has to be applied with caution and it is important to identify up front if outliers may occur in the forecast stream.

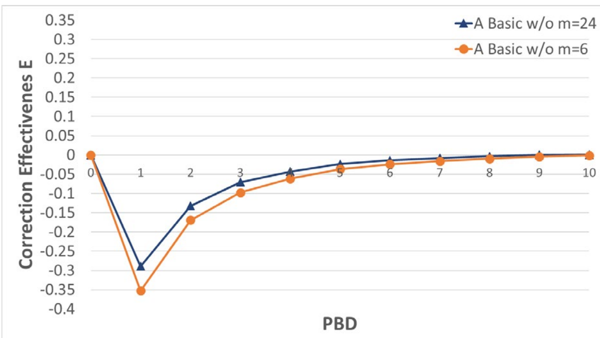

Figure 5: Performance decrease M2 X=0.9 scenario *A Basic w/o.*

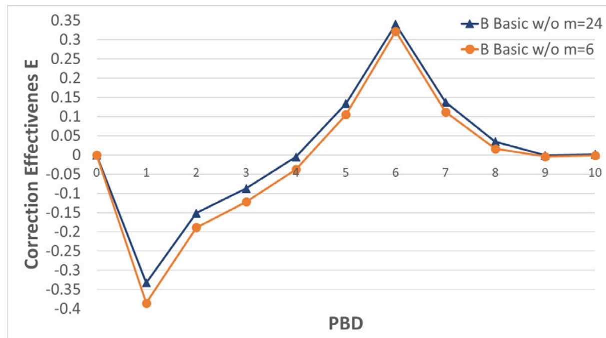

Figure 6: Performance decrease M2 X=0.9 scenario *B Basic w/o.*

5.4 Correction effectiveness related to parameter X

In this section the influence of parameter X on the correction effectiveness for the scenarios *A Basic* and *B Basic* with m=24 is observed for j=7 periods before delivery where outliers occur in our setting. The parameter X is used for the inverse of the normal distribution to compute the threshold for applying the correction method. The higher X is, the higher the threshold becomes and the probability of correction is reduced. The results in Figure 7 and Figure 8 show that the correction method M2 has always a higher correction effectiveness than M1. For the better method M2, a concave relationship between the correction effectiveness and X can be observed. This means that neither very low nor very high X values lead to the best results. For both figures it has to be considered, that only j=7 is illustrated and the correction effectiveness for the X values can change when discussing other periods before delivery j or calculating an average of correction effectiveness for all j values as conducted in the next subsection. Concerning RQ1 and RQ2 these observations show that the parameter X has an important influence on the outlier correction performance.

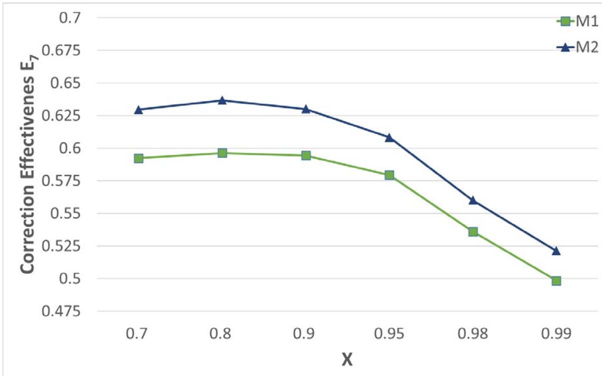

Figure 7: Correction effectiveness *A Basic* for *j*=7.

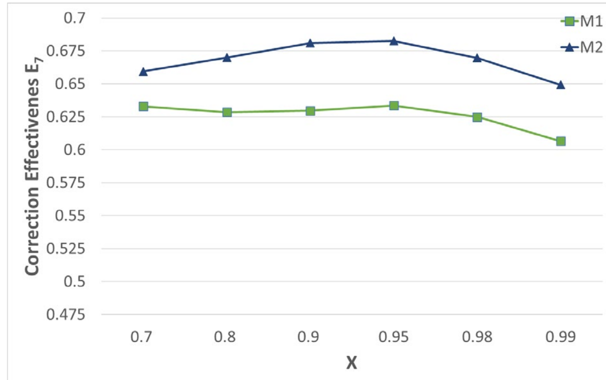

Figure 8: Correction effectiveness *B Basic* for *j*=7.

5.5 Sensitivity analysis

In this section the sensitivity of the results concerning the demand model, i.e. the forecast generation process, and the model parameter X is tested in order to answer RQ3. $E(X)$ shown in Table 2 is the average over all j values between 1 and 10. This means that the positive effects when an outlier occurs are averaged with the negative effects of additional uncertainty when no outlier occurs. Furthermore, Table 2 reports only the result for the beneficial method (out of M1 or M2; column method) with the best X value ($X \in \{0.7, 0.8, 0.9, 0.95, 0.98, 0.99\}$; column X). The results from Table 2 show that M2 performs better than M1, which can be expected related to the results from above, however, it is interesting that it outperforms M1 for all scenarios. This is interesting since for the scenarios with forecast bias, also M1, which applies the average order amount for forecast correction, might be a good method. Furthermore, the results show that high X values lead to the best results which is not intuitive from the results in section 5.4. However, this shows that it is better to have a high threshold value for the outlier identification in order to reduce the additional uncertainty induced in periods without outliers. Even though this high threshold value does sometimes neglect occurring outliers. With respect to random noise α, the results show that higher random noise leads to lower performance of the outlier correction which is intuitive since then outliers are more difficult to identify. With respect to the level of forecast bias b, the results show that a higher forecast bias leads to a better performance of the outlier correction, which is intuitive since then more biased forecast values are identified as outliers and then corrected. With respect to outlier probability γ and outlier amount δ, the results both support intuition since an increase in γ and δ leads to an increase in correction effectiveness. Overall, a broad range of correction effectiveness values between 1.74% and 28.23% depending on the parameters of the forecast generation process can be observed.

Table 2: Sensitivity Analysis

	α	γ	E(X)	method	X
β=0,δ=1	0.5	0.5	0.1485	M2	0.99
		1	0.1588	M2	0.99
		2	0.1666	M2	0.99
	1	0.5	0.0915	M2	0.98
		1	0.1031	M2	0.98
		2	0.1175	M2	0.99
	2	0.5	0.0252	M2	0.99
		1	0.0510	M2	0.99
		2	0.0846	M2	0.99

	β	γ	E(X)	method	X
α=1,δ=1	0.5	0.5	0.1214	M2	0.99
		1	0.1431	M2	0.99
		2	0.1547	M2	0.99
	1	0.5	0.1800	M2	0.99
		1	0.1912	M2	0.99
		2	0.2012	M2	0.99
	2	0.5	0.2420	M2	0.99
		1	0.2524	M2	0.99
		2	0.2607	M2	0.99

	α	δ	E(X)	method	X
β=0,γ=1	0.5	0.5	0.1089	M2	0.99
		1	0.1588	M2	0.99
		2	0.1757	M2	0.99
	1	0.5	0.0585	M2	0.99
		1	0.1031	M2	0.98
		2	0.1611	M2	0.99
	2	0.5	0.0174	M2	0.99
		1	0.0510	M2	0.99
		2	0.1054	M2	0.98

	β	δ	E(X)	method	X
α=1,γ=1	0.5	0.5	0.1020	M2	0.99
		1	0.1431	M2	0.99
		2	0.1938	M2	0.99
	1	0.5	0.1545	M2	0.99
		1	0.1912	M2	0.99
		2	0.2309	M2	0.99
	2	0.5	0.2203	M2	0.99
		1	0.2524	M2	0.99
		2	0.2823	M2	0.99

6 CONCLUSION

In this paper two selected outlier correction methods for forecast data are tested. Based on the distribution of the final order amount a threshold value is calculated for outlier identification and two different corrections methods are compared. The first correction method is simply based on historical data of the average final order amount, while the second correction method applies values of the previously shared (i.e. j+1) forecast information. To test the performance of the outlier correction methods a simulation study using AnyLogic© is performed. The results show that both outlier correction methods perform well and provide the expected smoothening of the forecasts values which consequently can reduce nervousness in the production system. However, results show that method M2 significantly performance better than M1. Furthermore, the results show that outlier correction might add an additional source of uncertainty if no outliers occur and therefore, high threshold values for outlier identification are advantageous. In future research the outlier identification could be extended to be based on the mean and variance of the forecast updates with respect to the periods before delivery instead of the final order amounts. Since simulation always treats streamlined systems, a further research activity is identified to conduct a numerical study, where the performance of the outlier correction methods with real world forecast data are tested. Another research direction is the integration of the presented demand forecast model into a simulation framework that mimics a hierarchical production planning system to identify the respective performance increase if forecasts are corrected and material requirements planning is applied.

ACKNOWLEDGMENTS

The work described in this paper was done within the Produktion der Zukunft Project (InnoFIT, #867471), funded by the Austrian Research Promotion Agency (FFG).

7 REFERENCES

Chen, C., and L.-M. Liu. 1993. "Forecasting Time Series with Outliers." *Journal of Forecasting* 12:13–35.

Güllü, R. 1996. "On the Value of Information in Dynamic Production/Inventory Problems under Forecast Evolution." *Naval Research Logistics* 43:289–303.

Harvey, A. C., and S. Peters. 1990. "Estimation Procedures for Structural Time Series Models." *Journal of Forecasting* 9:89–108.

Heath, D. C., and P. Jackson. 1994. "Modeling the Evolution of Demand Forecasts with Application of Safety Stock Analysis in Production/Distribution Systems." *IIE Transactions* 26:17–30.

Ho, S.-S., M. Schofield, B. Sun, J. Snouffer, and J. Kirschner. 2019. "A Martingale-Based Approach for Flight Behavior Anomaly Detection." In *2019 20th International Conference on Mobile Data Management,* 43–52, Los Alamitos, California: IEEE Computer Society, Conference Publishing Services.

Martínez–Álvarez, F., A. Troncoso, J. C. Riquelme, and J. S. Aguilar–Ruiz. 2011. "Discovery of Motifs to Forecast Outlier Occurrence in Time Series." *Pattern Recognition Letters* 32:1652–1665.

Norouzi, A., and R. Uzsoy. 2014. "Modeling the Evolution of Dependency between Demands, with Application to Inventory Planning." *IIE Transactions* 46:55–66.

Svetunkov, I., and F. Petropoulos. 2018. "Old Dog, New Tricks: A Modelling View of Simple Moving Averages." *International Journal of Production Research* 56:6034–6047.

Taylor, S. J., and B. Letham. 2018. "Forecasting at Scale." *The American Statistician* 72:37–45.

Zeiml, S., K. Altendorfer, T. Felberbauer, and J. Nurgazina. 2019. "Simulation Based Forecast Data Generation and Evaluation of Forecast Error Measures." In *Proceedings of the 2019 Winter Simulation Conference,* edited by N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y-J. Son, 2119–2130, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Zeiml, S., U. Seiler, K. Altendorfer, and T. Felberbauer. 2020. "Simulation Evaluation of Automated Forecast Error Correction Based on Mean Percentage Error." In *Proceedings of the 2020 Winter Simulation Conference,* edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 1572–1583, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

WOLFGANG SEIRINGER is a research Associate at the Department of Operations Management at the University of Applied Sciences Upper Austria. He received his master degree in Business Informatics from the Technical University of Vienna. Besides discrete event simulation modeling and analysis, his research interest also includes information uncertainty and forecasting. His email address is wolfgang.seiringer@fh-steyr.at.

KLAUS ALTENDORFER works as a Professor in the field of Operations Management at the University of Applied Sciences Upper Austria. He received his Ph.D. degree in logistics and operations management and has research experience in simulation of production systems, stochastic inventory models, and production planning and control. His e-mail address is klaus.altendorfer@fh-steyr.at.

THOMAS FELBERBAUER is the academic director of the study program Smart Engineering at the St. Pölten University of Applied Sciences (Austria). He works as a Professor in the field of production planning and simulation at the Department Media and Digital Technologies. He received his Ph.D. degree in developing solution methods for stochastic project management. His email address is thomas.felberbauer@fhstp.ac.at.