Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

SIMULATION BASED FORECAST DATA GENERATION AND EVALUATION OF FORECAST ERROR MEASURES

Sarah Zeiml
Klaus Altendorfer

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

Thomas Felberbauer
Jamilya Nurgazina

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

ABSTRACT

Production planning is usually performed based on customer orders or demand forecasts. The demand forecasts in production systems can either be generated by manufacturing companies themselves, i.e. forecast prediction, or they can be provided by customers. For both alternatives, forecast prediction, as well as the customer-provided forecasts, the quality of those forecasts is critical for success. In this paper, a simulation model to generate forecast data that mimic different forecast behaviors is presented. In detail, an independent forecast distribution and a forecast evolution model are investigated to discuss the value of customer-provided forecasts in comparison to the simple moving average forecast prediction method. Main findings of the paper are that Root-Mean-Square-Error and Mean-Absolute-Percentage-Error describe the forecast error well if no systematic effects are present and Mean-Percentage-Error provides a good measure for systematic effects. Furthermore, systematic effects like overbooking are significantly reducing the value of customer-provided forecast information.

1 INTRODUCTION

Production planning is usually performed based on customer orders or demand forecasts. Such demand forecasts in production systems can be generated by manufacturing companies themselves applying qualitative methods, quantitative methods, or a combination of both (Hopp and Spearman 1999). An alternative to forecast prediction is that customers of the manufacturing company provide demand forecasts in a rolling horizon and the production system produces according to this information. For both possibilities, forecast prediction as well as the customer-provided forecasts, a key factor for the production planning performance is the quality of those forecasts (Forslund and Jonsson 2007). However, looking at the literature there are a lot of different forecast error measures available and it is not clear which ones should be applied from a practical perspective. Furthermore, customer-provided forecasts can in addition to pure stochastic errors also have a systematic behavior which might lead to different representations in different forecast error measures (Shcherbakov et al. 2013; Chai and Draxler 2014). Another relevant aspect when looking at forecasts is the timing of the forecast, i.e. how far in advance a forecast is stated, and the frequency of the forecast updates (Forslund and Jonsson 2007). For a better understanding of the basic relationships between the aspects mentioned above, the aim of the conducted simulation study is the modelling and simulation of different forecasting processes, i.e. forecast prediction and customer-provided forecasts, and the generation of forecast data for which different forecast error measures are tested and discussed. In case of customer-provided forecasts, a simple method for data generation is the application of independent forecast disturbances, i.e. the forecasts are stochastic and their updates are independent of the past forecast values. This simple method simplifies the data generation, i.e. a desired forecast behavior can easily be generated, however, it ignores the interdependencies of forecast values identified in the real world.

978-1-7281-3283-9/19/$31.00 ©2019 IEEE

Therefore, Heath and Jackson (1994) provided the martingale model of forecast evolution which characterizes a behavior where past forecast values are updated and the forecast quality improves with respect to decreasing time before delivery (Heath and Jackson 1994). Both methods are applied in this study to generate forecast data based on simulation. In case of forecast prediction, a common forecasting method is the simple moving average which is tested in this paper as basis for identification of value of customer-provided forecast information.

The following research questions are addressed in this paper:

- RQ1: How do different forecast error measures describe the forecast quality of customer-provided forecast data which is either generated with independent forecast disturbance or with forecast evolution methods?
- RQ2: What is the effect of systematic forecast disturbances in customer-provided forecasts on the studied forecast error measures and which ones are best suited to identify systematic forecast behaviors?
- RQ3: What systematic and unsystematic forecast errors in customer-provided forecasts lead to lower forecast quality than the forecast prediction method moving average, i.e. what are customer-provided forecast characteristics that lead to no information advantage?

The first research question provides an overall insight and understanding of the different forecast error measures and their application. Besides its application, also the performance of the error measures is discussed. In the second research question the practical problem is addressed that the customer forecasts may be biased and, therefore, the existing forecast error measures are reviewed for their ability to identify biased behaviors. The third research question provides insights into the problem of applying customer-provided forecasts or replacing them by forecast prediction methods. To answer these research questions, a simulation model providing forecast data either based on forecast prediction, i.e. moving average, or based on customer-provided forecasts, i.e. independent forecast disturbance and forecast evolution, is developed in this paper. Managerial insights are generated with a numerical study on different scenarios for customer-provided forecast behavior including different systematic overbooking effects. Note that these scenarios are motivated by practical observations from different companies, however, these are very stylized settings.

The paper is structured, as follows. In Section 2 we provide a review of the relevant literature and in Section 3 the forecast processes as well as the simulation model are introduced. Additionally, several forecast error measures are presented and discussed in Section 3. It is followed by the numerical study in Section 4, which contains the results of the simulation and discusses those on the basis of the characterized research questions. After the numerical study, the conclusions and further research are provided.

2 LITERATURE REVIEW

As motivated in the introduction, there are two different sources of forecast information which can influence the planning decisions: the forecasts which are received directly from the customer, i.e. the customer-provided forecasts, and the demands which are calculated on basis of historical demands realizations, i.e. forecast predictions (Zotteri and Verganti 2001). The decision making process in production systems depends strongly on future demands, as they are often used as an input for capacity planning (Hopp and Spearman 1999; Forslund and Jonsson 2007). Production planning nowadays alleviate sharing reliable customer and demand information as those are key for manufacturers to cope with demand volatility (Durango-Cohen and Yano 2006). For forecast predictions, the available information in the present is used for the prediction of future production planning. The used forecasting methods for predicting the future demands can be differentiated in two categories. Hopp and Spearman (1999) characterize the qualitative forecasting method and the quantitative forecasting method. The qualitative forecasting method is relying on the expertise of people which estimate the future demands. Note that qualitative forecasts are not focused in this paper. The quantitative forecasting method is a process where the forecasts are calculated on basis of historical data, using a numerical measure or a mathematical model. The simplest quantitative forecast method, that is commonly used in practice, is the moving average, where past values of a predefined horizon are used to compute a prediction of the future demand (Hopp and Spearman 1999). However, Hopp and Spearman (1999) state that there is no guarantee for perfect prediction of the future demand. According to Fildes and Kingsman (2011), the best method for forecast prediction doesn't exist, as the method is dependent on the generation process of the forecasts. A common practice for making forecasts in organizations include generic methods, for example, the moving average (Fildes and Kingsman 2011).

Forslund and Jonsson (2007), for example, emphasize that the quality of forecast information has a strong impact on how the forecasts can be used. For customer-provided forecasts, they identify that collaboration is a key factor for the information quality gained by the members of the supply chain (Forslund and Jonsson 2007). Note that scenarios with different levels of forecast quality are generated in the simulation study conducted in this paper. Altendorfer et al. (2016) have studied the effect of forecast errors in unbiased customer-provided forecasts when MRP (material requirements planning) is conducted. They identified that forecast errors have a significant impact on the capacity needed and also the overall costs. Enns (2002) is examining the effects of demand uncertainty and forecast bias in a system with customer-provided forecasts and conclude that both are influencing customer delivery service levels. He also characterizes timing uncertainty and demand uncertainty which both have an impact on the planning process (Enns 2002). The specific aspect of timing of information is studied by Güllü (1996) with a forecast evolution process for customer-provided forecasts. He compares a production system without forecast information and with forecast information where the forecasts are updated two periods ahead. He shows a significant improvement in production system performance for updated forecasts. However, the forecast evolution is restricted to two periods in the future (Güllü 1996). A model to describe the increasing quality of (customer-provided) forecasts with the increased knowledge for decreasing time to the final delivery period is the martingale model of forecast evolution of (Heath and Jackson 1994). This model is applied in the current paper to generate the customer-provided forecast data in the forecast evolution scenarios. An application of this forecast evolution model is presented in Felberbauer and Altendorfer (2014) where the dynamics and uncertainties of different forecast behaviors were investigated for a simple production system structure. However detailed analysis of the forecast quality in customer-provided forecasts in comparison to forecast predictions are not addressed there (Felberbauer and Altendorfer 2014).

The literature reviewed above, shows that many authors are dealing with the topic of forecast errors and demand uncertainty but do not differentiate between customer provided forecasts and forecast predictions or just focus on one of them. Furthermore, only few authors distinguish between systematic and non-systematic forecast error effects. Therefore, a research gap is the comparison of forecast quality for customer provided forecasts and forecast predictions with respect to systematic and non-systematic forecast error effects.

3 MODEL DESCRIPTION

The sections above describe how forecasts, which are used in a production system, can be generated. Also the information quality of forecasts can differ and the errors in forecast data have an impact on the planning process and the costs in an organization. This section is describing the forecast processes used in the simulation model. For forecast prediction, the moving average method is implemented. For modelling customer-provided forecasts, independent forecast disturbance and forecast evolution are introduced. Furthermore, different commonly used forecast error measures for the evaluation of the generated forecasts are presented.

3.1 Forecast generation processes

Before the different forecast generation processes are introduced, the used variables are defined. For the calculation of the forecasts, the forecast amount or demand predictions, the final order amounts and the long-term forecast are used. The due date period is defined by the index i. The variable for the forecast amounts is $x_{i,j}$, where the index j represents the periods before delivery, i.e. $x_{45,3}$ is the forecasted amount for period 45 stated 3 periods in advance. The final order amount is represented by $x_{i,0}$, where the zero indicates that zero periods before delivery are left, for the period i. The long-term forecast is characterized as $\widetilde{x_\iota}$. Note that in our numerical setting $\widetilde{x_\iota}$ is constant for all periods and all materials, i.e. the long term forecast provided by the customer is constant.

The forecast process consists of a starting point, where the long-term forecast $\widetilde{x_\iota}$ is generated. When the forecast horizon is reached, the forecast values $x_{i,j}$ are altered (forecasts are calculated) each period until the final due date at which the process ends.

Forecast prediction – Moving Average

The forecast prediction method used in this paper is the moving average. It provides the base line for the numerical study. The moving average is a common quantitative forecast method and according to Hopp and Spearman (1999) it is the simplest forecasting method, as it is just the average of the actual observed demands (Hopp and Spearman 1999). The forecast for period $i+1$, therefore equals $x_{i+\mathrm{j},j} = \sum_{j=i-n+1}^{i} x_{j,0}/n \quad j\epsilon\{1,\dots,J\}$. Whereby n is the number of periods applied for averaging and J is the number of forecasted periods, i.e. the forecast is equal for all future periods until the next forecast update.

Customer-provided forecasts – Independent forecast disturbance

The independent forecast disturbance is the first approach to mimic customer-provided forecasts. This method uses a long-term forecast value $\widetilde{x_\iota}$ and an error value ($\varepsilon_{\mu_j\,\sigma_j}$). The forecast amount for period i is then $x_{i,j} = \widetilde{x_\iota} + \varepsilon_{\mu_j\,\sigma_j}$ which means that it equals the long term forecast but has a stochastic error value. This error value $\varepsilon_{\mu_j\,\sigma_j} \sim N(\mu_{j,}\sigma_j)$ is normally distributed whereby $\sigma_j = \alpha_j * \widetilde{x_\iota}$ and $\mu_j = fb_j * \widetilde{x_\iota}$ are applied with fb_j and α_j defining the distribution itself that is scaled with $\widetilde{x_\iota}$. Note that fb_j and α_j can be directly compared between different materials or scenarios because they are unscaled. Furthermore, α_j describes the unsystematic noise and fb_j identifies the systematic behavior of booking too much or too less. This forecast generation method implies that forecasts are generated each period independently of the forecasts one period earlier.

Customer-provided forecasts – Forecast Evolution

The second approach to mimic customer-provided forecasts is forecast evolution. It differs from the independent forecast behavior as the error value calculated is added to the forecast value from the previous period, i.e. $x_{i,j} = x_{i,j+1} + \varepsilon_{\mu_j\,\sigma_j}$. The calculation of the error value is characterized as stated above. This method, which is introduced in Güllü (1996), implies that forecasts are not independent and evolve over time, as with each period new information is added and the forecast is updated (Güllü 1996). The method is based on the martingale model of forecast evolution published by (Heath and Jackson 1994).

3.2 Simulation model

One main contribution of this paper is the development of a simulation model that is able to generate forecast data according to the above stated processes with systematic and unsystematic errors. In addition to the different forecast behaviors, the model is also calculating various forecast error measures for the evaluation of the generated forecasts. The different scenarios implemented in the simulation model are predefined by an input table and various parameters. The data generating process is modelled in the simulation software AnyLogic and for the parametrization a table-function according to Table 1 is used. This means that different scenarios can be generated concerning their systematic and unsystematic forecast error with respect to periods before delivery. The following further parameters are needed to specify the scenarios and the simulation model. The forecast horizon which indicates how many periods in advance of the due date the forecast is altered with the error value is set to 10 periods, i.e. the forecast is constant for periods further in the future with value $\widetilde{x_\iota}$ (see also Figure 1). Within the forecast horizon the forecast amount is changed periodically, a typical forecast evolution behavior is visualized in the Figure 1. The window size for moving average *n* is defined as 24 periods.

Table 1: Input table with the fb-basis 0.025 and alpha 0.1.

Product	Periods before delivery	Forecast bias	alpha
1	10	μ_{10}	α_{10}
1	…	…	…
1	0	μ_0	α_0

As the error value is calculated with the normal distribution, the values can be positive as well as negative. Therefore, the issue could occur, that the calculated forecast amounts are negative values. With the assumption that the forecasts can never be a negative value, the negative forecast amounts are saved and set off to the next period. This leads to the following slightly changed forecast values for the forecast evolution method $x_{i,j} = (\hat{x}_{i,j+1} + \varepsilon_{\mu_j \sigma_j}) + \min(0, x_{i-1,j})$ with $\hat{x}_{i,j} = \max(0, x_{i,j})$. Note that the independent forecast disturbance works similar. In Figure 1, two different customer-provided forecast behaviors are visualized for the duration of 20 periods. The line with the square markers indicates a systematic forecast error, which is added to the forecasted amount and the line with the diamond markers is representing the forecast evolution behavior. For the systematic behavior (square markers), the long-term forecast is set to 800 in this example and for the forecast evolution (diamond markers), the long-term forecast is set to 880 (note that this value is chosen for visualization reasons only). The forecast horizon is set to 10 periods for forecast evolution, i.e. forecasts stay at their long-term values from 20 to 11 periods before delivery and the first change happens 10 periods before delivery. The first change with the systematic forecast error behavior happens in period eight (i.e. forecasts stay at their long-term values until 9 periods before delivery), where the forecast is increased to 820. After period six, the forecast is slowly decreasing and it returns to the base value 800 in period three. This systematic forecast behavior, which is used as one scenario in the simulation study, represents a systematic overbooking of the customer. Note that in this example for the systematic forecast error, there is no unsystematic noise added. However, most scenarios tested in the numerical example combine systematic and unsystematic errors.

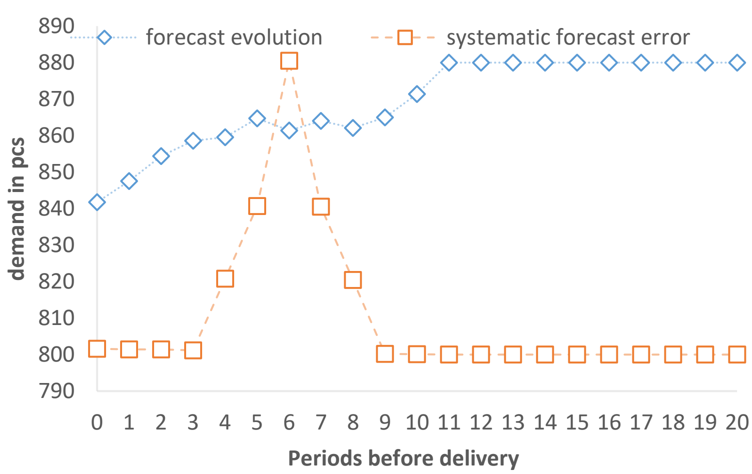

Figure 1: Example for simulated forecast behavior with systematic error and forecast evolution.

The forecast evolution behavior shows one forecast stream that is decreasing with every progressing period to the final due date. There are a few small fluctuations, e.g. five periods before delivery the forecast slightly increases, however the values are decreasing steadily to 840. Note that this is only one possible realization of a forecast evolution.

3.3 Forecast Error Measures

To evaluate the forecast quality, the following forecast error measures are used in the numerical model.

Mean Squared Error (MSE)

The mean squared error, which is suggested by Hopp and Spearman (1999), is calculated as $MSE_j = 1/m\sum_{i=1}^{m}\left(x_{i,j} - x_{i,0}\right)^2$. It indicates the deviation between the actual demand and the forecast j periods before delivery whereby m is the number of values in the forecast history. The mean squared error can only be positive and the value should be as small as possible (Hopp and Spearman 1999). A mean squared error value of zero would indicate perfect accuracy. Note that higher deviations have higher weights.

Root Mean Squared Error (RMSE)

The root mean squared error, which is suggested by Shcherbakov et al. (2013), is calculated with the square root of the MSE, i.e. $RMSE_j = \sqrt{MSE_j}$. The RMSE is also useful for showing the error distribution. According to Chai and Draxler (2014) it is a statistical standard measure, which is often used. The value should be as close to zero as possible (Chai and Draxler 2014).

Mean Absolute Percentage Error (MAPE)

The original equation of the mean absolute percentage error is $MAPE_j^* = 1/m\sum_{i=1}^{m}\left|x_{i,j} - x_{i,0}\right|/x_{i,0}$, but in our setting it can occur that the final order amount $x_{i,0}$ is zero; therefore, the division would result in an error or infinite value. That is why the calculation was altered to $MAPE_j = \sum_{i=1}^{n}\left|x_{i,j} - x_{i,0}\right|/\sum_{i=1}^{n}x_{i,0}$, where the sum of the absolute error is divided by the sum of final order amounts. It is one of the more popular error measures, as it is percentage based and the value should be as close to zero as possible, see (Shcherbakov et al. 2013). The interpretation of the MAPE is easier than the other error measures which are introduced because the value is percentage based. A value of zero is indicating that the model does not have a mean error, which is implying that the value should be as close to zero as possible.

Mean Percentage Error (MPE)

The mean percentage Error is calculated similar to the MAPE, however no absolute values are taken, i.e. $MPE_j = \sum_{i=1}^{m}\left(x_{i,j} - x_{i,0}\right)/\sum_{i=1}^{m}x_{i,0}$. The value of the mean percentage error should best be as close to zero as possible and is the metric for the percentage of the error (Shcherbakov et al. 2013). The mean percentage error of zero indicates that there is no mean error of the model. However, in comparison to the MAPE, the MPE is able to have negative values, which give an indication if the error value was higher than the final order or if it was lower. A negative effect concerning the MPE is that positive and negative deviations can diminish each other.

Mean Absolute Deviation (MAD)

The MAD is according to Hopp and Spearman (1999) a very common forecast error measure, which can only result in positive values. It is calculated as $MAD_j = 1/m\ \sum_{i=1}^{m}\left|x_{i,j} - x_{i,0}\right|$. The objective should be to keep the value of the MAD as small as possible. As it is the absolute difference, the MAD cannot tell if the forecast value was higher than the final order or lower (Hopp and Spearman 1999).

Mean Deviation (MD)

The Mean Deviation is the measure of the mean difference between the forecasted amount and the final order amount, i.e. $MD_j = 1/m \sum_{i=1}^{m} x_{i,j} - x_{i,0}$. As the MD is not taking the absolute values, they can be negative, which indicates that the forecasted amount was smaller than the final order amount. Therefore, the value should be as close to zero as possible (Shcherbakov et al. 2013).

Coefficient Variation (CV)

Since the RSME seems to be a promising error measure, however it is not normalized, i.e. different products cannot directly be compared, a modification of the RMSE is introduced here. The Coefficient of Variation (CV) is calculated as $CV_j = (RMSE_j * m)/\sum_{i=1}^{m} x_{i,0}$. The purpose of the CV is that the values remain comparable for the numerical study. The CV like the RMSE cannot be applied to show a systematic direction of forecast deviations. The value should be close to zero.

Error measure selection for numerical example

A detailed analysis of the error measures introduced above shows that the error measures presented are all relatively similar and also offer a similar interpretation. Therefore, not all of these error measures are presented in the following numerical study to answer the research questions. However, they can be clustered into three sets of error measures. The first set of error measures provides a forecast quality information similar to the variance from statistics, these error measures are MSE, RMSE and CV. Since MSE and RMSE are not normalized, the CV was introduced and is used in the numerical study. The second set of error measures provides information on the absolute deviation between forecast values and final order amount, i.e. larger deviations do not have a higher weight as in the variance. These measures are MAD and MAPE. Since MAPE is a percentage value which is better suited for comparing different materials, it is applied in the numerical study. The third set of error measures, consisting of MD and MPE, shows an average deviation between forecast values and final order amounts. Even though positive and negative deviations are crossing out, these measures show potential to identify systematic forecast errors. Since the MPE is a percentage value, this error measure is used for the numerical study. Note that in the simulation model all of these error measures are calculated but only the above mentioned are presented.

4 NUMERICAL STUDY

In this numerical simulation study a set of different scenarios is tested to answer the research questions, where the noise parameter α_j and the forecast bias fb_j are varied for the customer-provided forecasts. Note that the moving average forecast prediction needs no further parametrization for its implementation and is applied as baseline for the comparison of all scenarios with customer-provided forecasts. The following scenarios for the customer-provided forecasts are tested:

- Forecast evolution behavior without systematic influences (FEV1), i.e. $fb_j = 0$ and α_j is set constant for all j values α_j=a in this scenario and $a \epsilon \{0.025, 0.05, 0.1, 0.15, 0.2\}$ are tested. Based on the forecast updates, this scenario leads to an increase in forecast quality when the due date approaches.
- Forecast evolution behavior with systematic influences (FEV2). Here fb_j is parameterized according to Table 2 and α_j=a is again applied with $a \epsilon \{0.025, 0.05, 0.1, 0.15, 0.2\}$. Note that a typical overbooking behavior further in the future, followed by a decrease of the amounts which were overbooked when the final due date approaches is modelled there (see also Figure 1). To parameterize the strength of this overbooking a base value b is applied with $b \epsilon \{0.025, 0.05, 0.1, 0.15, 0.2\}$.
- Independent forecast distribution without systematic influences (IFD1), i.e. $fb_j = 0$ and α_j is set according to $\alpha_j = j * a$ in this scenario and $a \epsilon \{0.01, 0.02, 0.04, 0.06, 0.08\}$ are tested. Based on

the independence between past forecasts and the updated values, also this scenario leads to an increase in forecast quality when the due date approaches.

- Independent forecast distribution with systematic influences (IFD2). Here a too high long term forecast is modelled which is kept until the last period before the final order, i.e. fb_0=b and $fb_j = 0$ for all j>0 with $b \epsilon \{0.025,\ 0.05, 0.1, 0.15,\ 0.2\}$. The unsystematic error $\alpha_j = j * a$ in this scenario and $a\ \epsilon\ \{0.01, 0.02, 0.04, 0.06, 0.08\}$. This implies that customers are overbooking by the same amount for all forecasted values, but their final order is then lower.

Note that all parameter combinations for the respective scenarios have been simulated, however, only a selected set of results is presented in this paper due to space limitations.

Table 2: forecast bias fb_j with respect to periods before delivery (PBD).

PBD	0	1	2	3	4	5	6	7	8	9	10
	b * 0	b * 0	b * 0	b * -1	b * -1	b * -2	b * 2	b * 1	b * 1	b * 0	b * 0

Some general implementation details for the simulation study are: the long-term forecast value $\widetilde{x_i}$ for the model is 800 pieces. The runtime of the developed simulation model is 2080 time units, one time unit represents one week and the simulation run was replicated 50 times. Over all of those replications, the average value of the defined forecast measures is reported.

4.1 Forecast Error Measures Discussion (RQ1)

For the first research question, the forecast errors are compared for the scenarios FEV1 and IFD1. This means there are no systematic influences in this setting and forecast quality increases when the final due date approaches. Note that only two values $a\ \epsilon\ \{0.025,\ 0.2\}$ for FEV1 and $a\ \epsilon\ \{0.01, 0.08\}$ for IFD1 are shown in Figure 2, however, the other results also confirm the stated findings.

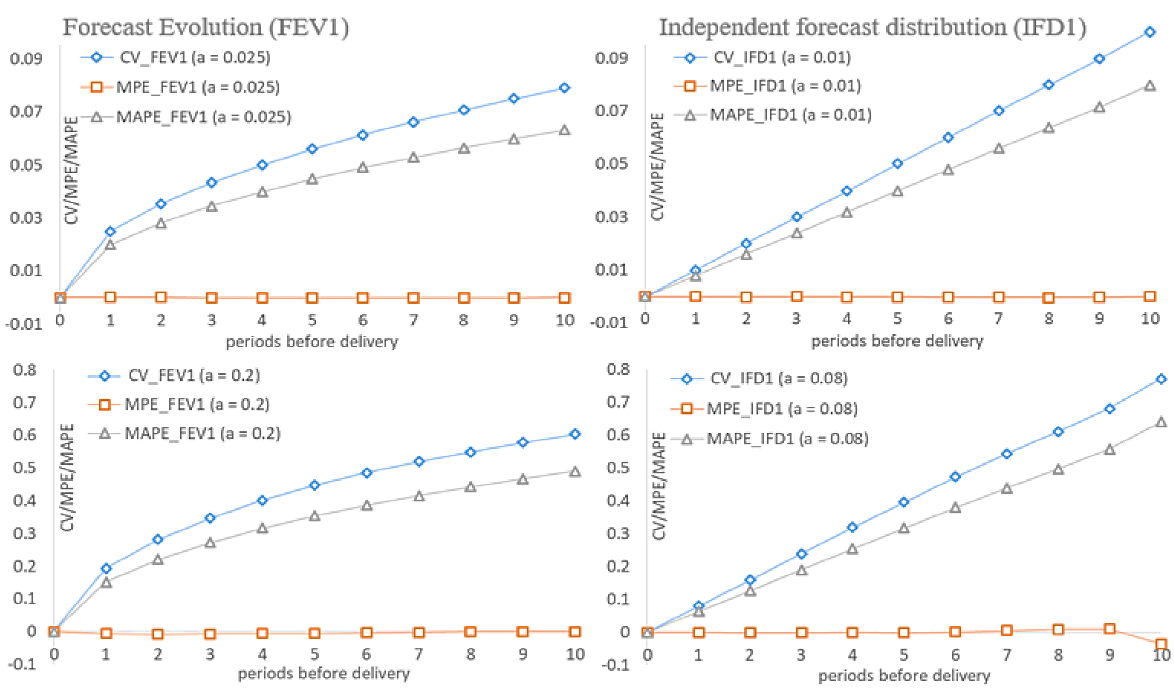

Figure 2: RQ1 forecast error comparison for unsystematic errors.

The first research question is to examine how different forecast error measures describe the forecast quality of customer-provided forecasts. The results in Figure 2 show that CV and MAPE are both appropriate error measures to identify the forecast quality. However, the MPE shows no forecast error and is, therefore, not an appropriate measure to identify unsystematic forecast errors. A detailed comparison between forecast evolution modelling (FEV1) and independent forecast disturbance (IFD1) shows that updating forecast values, as done in the FEV, leads to a smoother forecast error behavior than in IFD. This shows that the FEV is able to mimic the real world forecast updating process well and that forecast errors in this scenario are non-linear. IFD provides for the CV exactly the error values as intended by the calculation scheme for α_j which is a positive aspect related to modelling and data generation, however, it ignores the interrelation between forecasts that we assume to occur in practical applications. Managerial insights are therefore that CV and MAPE can both be applied to identify the unsystematic forecast errors and that FEV should be applied when simulation is used to generate data in a similar way as customer-provided forecasts. However, from a scientific aspect also IFD might be useful to parameterize forecast error exactly as intended for the respective study.

4.2 Error Measures for Systematic Forecast Disturbances (RQ2)

The second research question investigates how a systematic forecast behavior influences forecast error measures and which ones are best suited to identify these systematic behaviors. For this study, the scenarios FEV2 and IFD2 are applied. Note that only two values $a \in \{0.025,\ 0.2\}$ for FEV2 and $a \in \{0.01, 0.08\}$ for IFD2 are shown in Figure 3. Furthermore, only one value for the systematic forecast influence b=0.1 for FEV2 and b=0.1 for IFD2 are selected. However, the other results which are available but not shown also confirm the stated findings.

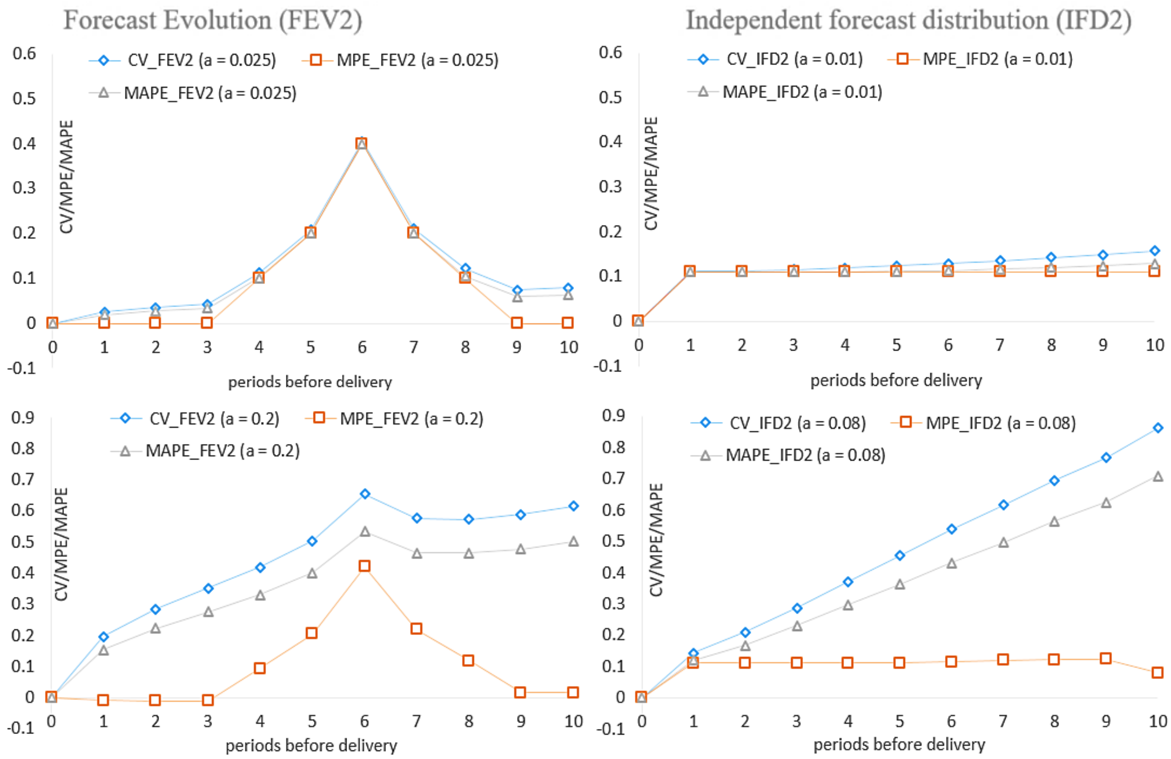

Figure 3: RQ2 forecast error comparison for systematic errors.

In general, the results from Figure 3 confirm the findings of RQ1 in the sense that MAPE and CV are showing a similar behavior. Concerning the description of systematic forecast influences, however, an additional finding in this subsection is that MPE is able to show the two different systematic behaviors of customer-provided forecasts very well. For the FEV2 setting, where customers have an overbooking peak six periods in advance of the final due date, this behavior is for both levels of unsystematic errors very well described by the MPE. Also for the IFD2 scenario, the MPE is for both levels of unsystematic errors able to exactly identify the systematic behavior in customer-provided forecast. This behavior is that until the final order the forecasts are too high, i.e. capacities at the supplier are reserved with that behavior, and the final order is then lower. The reason for this good performance of MPE to report systematic errors is that unsystematic errors are crossing out in the MPE calculation, see also Figure 2, therefore only systematic behaviors remain to be calculated. A managerial insight here is that MPE is an appropriate forecast error measure for systematic behaviors in customer-provided forecasts and CV or MAPE could both be applied for unsystematic behaviors. Looking at the results for CV and MAPE in detail shows that these measures do not only report that unsystematic forecast error but also the systematic behavior is included there. For a better understanding of customer-provided forecasts, this leads to a further research gap to develop further forecast error measures that provide a better distinction of systematic and unsystematic forecast errors. For practical application in manufacturing companies this is especially important since for systematic forecast errors there correction strategies can be developed and for unsystematic forecast errors, either safety stock or overcapacity might be needed. Since MAPE and CV showed a similar performance, in the following subsection only the CV is applied.

4.3 Information Advantage in Customer-Provided Forecasts (RQ3)

The last research question focusses on characteristics of forecasts that decrease the information quality of customer-provided forecasts in a way that it is better to do forecast prediction based on moving average forecast instead of using the customer-provided forecast data. In detail, the whole set of parameters for a and b in the FEV1 and FEV2 is simulated and the respective CV is evaluated. Within each single simulation run, in addition to the customer-provided forecasts, also the moving average forecast is calculated and the respective CV value is evaluated. Whenever the CV of moving average is lower than the one calculated with customer-provided forecasts, there is no information advantage in using these customer-provided forecasts and the production system performance can be improved by applying the moving average forecasts. The simulation results on FEV1 show that for all unsystematic error levels tested, i.e. $a \epsilon \{0.025,\ 0.05, 0.1, 0.15,\ 0.2\}$, the forecast error CV is lower for the customer-provided forecasts in comparison to the forecast prediction method moving average. Note that the respective results are omitted here because of space limitations. This is an interesting managerial insight since it implies that whenever there is no forecast bias and forecasts are continuously updated by the customer, this provides an advantage for the manufacturing system. Note that the cases for IFD1 and IFD2 cannot be used for comparison since the moving average forecast leads in the current simulation model implementation to a perfect forecast since the final order amount has no uncertainty based on $\alpha_0 = 0$. A correction of this implementation is left for further research, however, we conjecture that FEV modelling is much more relevant for practical application.

For the scenario FEV2, Figure 4 shows the results in the following format. The x-axis represents the periods before delivery, and the y-axis are the a-values which describe the level of unsystematic uncertainty. Three different levels of systematic forecast uncertainty are shown. The line visualizes the border for a-values (unsystematic error) above which the forecast prediction moving average leads to lower forecast error values CV than the customer-provided forecasts. For example, in the scenario FEV2 (b=0.05), i.e. a very low unsystematic error, six periods before delivery the moving average forecast provided a better result for a-values greater/equal 0.05. Looking at the results for FEV2 in detail shows that exactly in the periods where the systematic error reaches its peak, i.e. around six periods before delivery, the moving average forecast provides better results even for very low a-values. A managerial insight is therefore, that systematic errors significantly reduce the value of customer-provided forecast information and foster the application of simple forecast prediction methods. Note that also other forecast prediction methods could be applied here. An interesting opportunity provided by the developed data generation model is that specifically the forecast evolution with respect to periods before delivery can be studied. The developed simulation model provides the opportunity to study a set of different customer-provided forecast behaviors and is therefore a fruitful basis for further research on this topic.

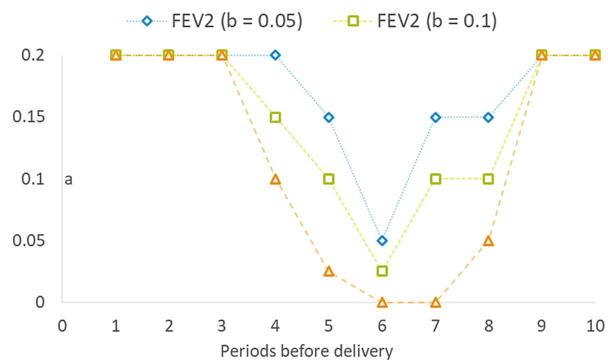

Figure 4: Moving average comparison with forecast evolution (FEV2).

5 CONCLUSION

In this paper a simulation model to generate customer-provided forecast data is developed and applied to evaluate the applicability of common forecast error measures and the value of customer information. In detail, two different methods for customer-provided forecast generation are implemented, i.e. forecast evolution behavior and independent forecast disturbance. Furthermore, the simple moving average forecast prediction method is applied to evaluate the value of customer-provided forecasts compared to the simple moving average approach.

A numerical study is conducted to answer the three research questions and the following managerial insights have been identified. The forecast error measures CV (based on RMSE) and MAPE have been identified as appropriate measure for evaluating the forecast quality if no systematic forecast behavior occurs. The MPE is not appropriate to identify unsystematic forecast error since positive and negative deviations diminish each other. However, for systematic forecast behavior, i.e. specific overbooking behaviors are tested, the MPE is a good predictor for the systematic forecasting effects. In this case CV (based on RMSE) and MAPE report a mixture of systematic and unsystematic forecast effects. Therefore, in future research a development of distinct forecast error measures for systematic and unsystematic effects is needed. With respect to the value of customer-provided forecast information, the results show that systematic forecast errors lead to a significant forecast error increase. The simple moving average forecast prediction method outperforms the respective customer-provided forecasts.

In general, the developed simulation model provides the opportunity to further investigate the influence of forecasts on production system performance. For future research one interesting research direction is development of analytical models for the identification and correction of systematic errors which occur in the forecast and evaluation of its performance. Limitations of the current study are that only few forecast behavior scenarios have been tested and only some common forecast error measures haven been applied. However, all of these limitations can be addressed in further research.

ACKNOWLEDGEMENTS

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

REFERENCES

Altendorfer, K., T. Felberbauer, and H. Jodlbauer. 2016. "Effects of Forecast Errors on Optimal Utilisation in Aggregate Production Planning with Stochastic Customer Demand". *International Journal of Production Research* 54(12):3718–3735.

Chai, T. and R. R. Draxler. 2014. "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? – Arguments Against Avoiding RMSE in the Literature". *Geoscientific Model Development* 7(3):1247–1250.

Durango-Cohen, E. J. and C. A. Yano. 2006. "Supplier Commitment and Production Decisions Under a Forecast-Commitment Contract." *Management Science* 52(1):54–67.

Enns, S. T. 2002. "MRP Performance Effects due to Forecast Bias and Demand Uncertainty". *European Journal of Operational Research* 138(1):87–102.

Felberbauer, T. and K. Altendorfer. 2014. "Comparing the Performance of Two Different Customer Order Behaviors within the Hierarchical Production Planning." In *Proceedings of the 2014 Winter Simulation Conference,* edited by A. Tolk, S.Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley and J. A. Miller, 2227-2238. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc..

Fildes, R. and B. Kingsman. 2011. "Incorporating Demand Uncertainty and Forecast Error in Supply Chain Planning Models". *Journal of the Operational Research Society* 62(3):483–500.

Forslund, H. and P. Jonsson. 2007. "The Impact of Forecast Information Quality on Supply Chain Performance". *International Journal of Operations & Production Management* 27(1):90–107.

Güllü, R, 1996. "On the Value of Information in Dynamic Production/Inventory Problems Under Forecast Evolution". *Naval Research Logistics* 43 (2): 289–303.

Heath, D. C.and P. Jackson. 1994. "Modeling the Evolution of Demand Forecasts with Application of Safety Stock Analysis in Production/Distribution Systems". *Institute of Industrial Engineers Transactions* 26 (3):17–30.

Hopp, W. J. and M. L. Spearman. 1999. Factory physics. 2nd. edn. McGraw-Hill / Irwin: Boston.

Shcherbakov, M. V., A. Brebels, N. Shcherbakova, A. P. Tyukov, T. A. Janovsky, and V. A. Kamaev. 2013. "A Survey of Forecast Error Measures". *World Applied Sciences Journal (Information Technologies in Modern Industry, Education & Society)* 24(24):171-176.

Zotteri, G. and R. Verganti. 2001. "Multi-level approaches to Demand Management in Complex Environments: An Analytical Model". *International Journal Production Economics* 71(1-3):221-233.

AUTHOR BIOGRAPHIES

SARAH ZEIML is a research associate at the Department of Operations Management at the University of Applied Sciences Upper Austria. She received her bachelor's degree in Business Informatics from the Johannes Kepler University in Linz. Besides discrete event simulation modeling and analysis, her research interest also include information uncertainty and forecasting. Her email address is sarah.zeiml@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 PhD 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 academic director of the sutdy programme 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 PhD degree developing solution methods for stochastic project management. His email address is thomas.felberbauer@fhstp.ac.at.

JAMILYA NURGAZINA is a research associate at the Department Media and Digital Technologies at the St. Pölten University of Applied Sciences (Austria). She received her Master's degree in Industrial Management from Heriot-Watt University. Her research interests include information uncertainty forecasting, information visualization and supply chain management. Her e-mail address is jamilya.nurgazina@fhstp.ac.at.