



Continuous modelling of machine tool failure durations for improved production scheduling

B. Denkena¹ · M.-A. Dittrich¹ · L. Keunecke² · S. Wilmsmeier¹

Received: 12 November 2019 / Accepted: 3 February 2020 / Published online: 12 February 2020
© The Author(s) 2020

Abstract

Unforeseen machine tool failures due to technical issues can cause downtimes leading to delays during production. To reduce delays, rescheduling of the production is, in most cases, necessary. However, warranting such a change requires reliable knowledge about the duration of the failure. This article presents a method to provide this knowledge by estimating the duration of a machine tool failure based on previous failure durations. Using the cross-industry standard process for data mining (CRISP-DM) and statistical methods, the embedded model for failure classification and duration is continuously improved. The method is thoroughly tested using multiple distributions, parameters and a practical use case. The results show high potential for predicting the duration of machine tool failures, which consequently could lead to improved quality of rescheduling.

Keywords Production planning · Scheduling · Manufacturing system · Failure duration

1 Introduction

Growing customer expectations and technological developments have led to an increased complexity of manufacturing systems. Companies face stochastic disturbances and cyclic demands and consequently they struggle to achieve their full manufacturing capacity [1]. Additionally, they have to deal with shorter product and technology lifecycles, increasing number of variants, customized products, short delivery times and competitive pricing [2–4]. Thus, efficient production planning and control (PPC) has become a crucial competitive advantage for many companies [1].

However, stochastic influences on the production time make exact planning difficult. For example, machine tool failures (MTF), defined by Iserman as “a permanent interruption of a system’s ability to perform a required function under specified operating conditions” [5], are very common. These are often caused by technical issues. Hence, different approaches exist in the area of maintenance (preventive or

predictive maintenance to avoid them) [6]. Nevertheless, the implementation of these approaches is not always possible. Therefore, the quality of production control depends on the ability to forecast the time to restoration (TTR) as accurately as possible and to use this information in a targeted manner when rescheduling orders.

In practice, a forecast of MTF durations is often difficult to obtain due to the poor quality of the operating data used for the forecast. These are usually recorded manually and are therefore often faulty. Hence a methodology for predicting MTF durations is required. This must ensure the quality of manually recorded failures by using suitable routines. At the same time, the quality of predicted failure durations by means of a structured and as complete as possible classification of the recorded MTF must also be guaranteed.

2 State of the art

In common methods for the modelling of MTF events, an exponential distribution [7–9] or an Erlang distribution [10, 11] is assumed for the entire failure duration of a machine tool. The Weibull distribution or the logarithmic normal distribution are also often applied [12]. To model these distributions, the mean value of the TTR (MTTR) is used. Due to the complexity of disturbances like MTF, such a description

✉ S. Wilmsmeier
wilmsmeier@ifw.uni-hannover.de

¹ Institute of Production Engineering and Machine Tools, An der Universität 2, 30823 Garbsen, Germany

² Volkswagen AG, Berliner Ring 2, 38440 Wolfsburg, Germany

is a simplification [13] because a distinction between different MTF classes is often not made. As a result the MTTR is obtained on the basis of data from many different failure events. Thus, in practice, the data is enhanced by the expert knowledge of the production planner or maintenance engineer. However, there is no systematic analysis which may lead to sub-optimal planning results.

An approach with automatic and structured failure classification and individualized forecast values for each failure class is presented by Oladokun. Variables, which affect the duration of an MTF, are identified as input for an artificial neural network. Output variable is a failure duration class which divides the duration into three possible time intervals, depending on how strongly the duration influences production planning. Even if the accuracy of the forecast model is specified with 70% [14], it only uses a fixed and small number of failure duration classes. Hence, the output cannot serve as a basis for high-quality rescheduling decisions.

Research in the area of maintenance as well as condition and health monitoring focuses on the diagnosis of faults and the prognosis of wear on tools and components [15, 16]. For example, event-related or condition-related data is used for information retrieval in condition-based maintenance [17]. The event-related data include machine failures, maintenance measures or information on the tool, the machining object or the process [17–19]. In contrast, condition-based data are measured values that reflect the condition of the unit under investigation. For this purpose, the term sensor fusion refers to various theoretical and practical approaches to combine several or different signals of machine tools to obtain more desirable results [20, 21]. In this context, this means more precise, complete, reliable or robust results than a single signal or one type of signal [22]. Neither summaries of the state of the art [20–25] nor the individual publications considered in this article provided results for the prognosis of MTF durations.

Other research areas investigate the suitability of different methods to determine an exact prognosis of accident durations [26], the occupancy time of hospital beds [27, 28] and the duration of surgeries [29]. Although these cases are similar to the problem of MTF duration prognosis discussed in this article, the solutions cannot simply be transferred because extensive data sets are needed to train the underlying models. Especially, small and medium-sized enterprises (SMEs) often do not have the resources or the incentive for a detailed and continuous data recording of their production processes [30].

3 Methodology for predicting MTF durations

The review of the state of the art shows that there is a lack of research with regards to the prognosis of MTF durations in the area of PPC. Established indicators such as unclassified

MTTR are too imprecise for the prognosis of individual MTF. Therefore, the methodology presented in Fig. 1 was developed. It was developed according to the concept of “Cross-industry standard process for data mining” (CRISP-DM). CRISP-DM defines the requirements and work steps shown in Fig. 2 for the creation of models with a high forecast quality. The four core steps are Sects. 3.1–3.4. These steps are included in the methodology for MTF prognosis. As a result, the methodology enables systematic processing of MTF events and provides the production planner with statistically validated prognosis of MTF durations.

3.1 Data comprehension

A prediction model must first be trained to forecast MTF durations. This can be done by either using historical data or information about MTF that occur after implementation. In the latter case, the use of the method must ensure that a forecast is not made until sufficient data for statistic validation is available. For data recording, tuples consisting of the MTF duration and the MTF class (e.g. tool breakage) are entered into the model via a user interface.

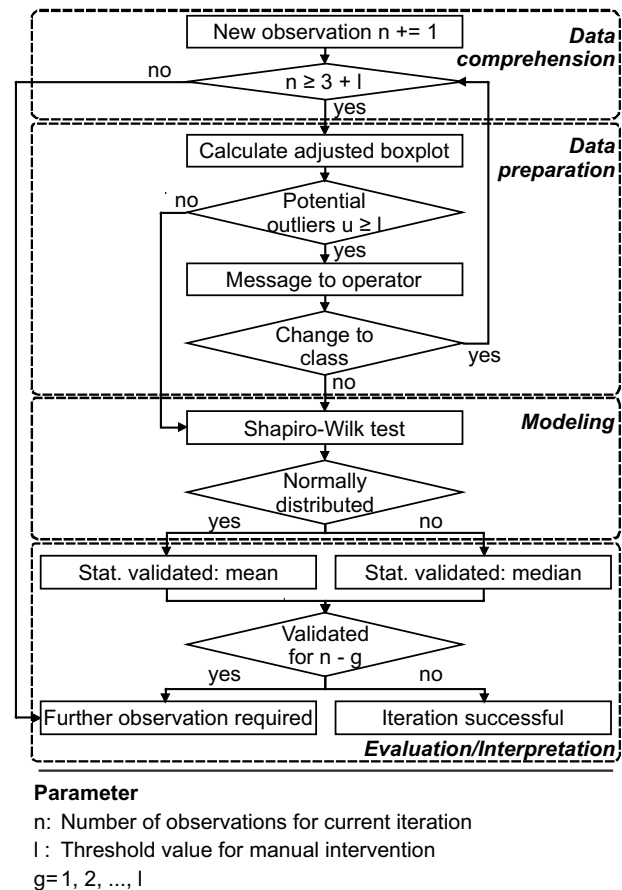


Fig. 1 Method for prognosis of MTF durations

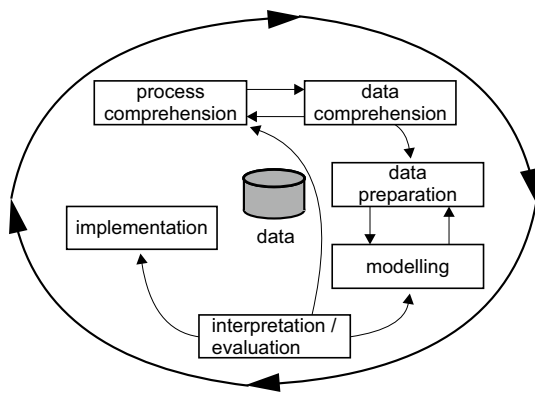


Fig. 2 Cross-industry standard process for data mining [39]

If there are at least $n = 3 + l$ observations for the current MTF class, an adjusted box plot is calculated for determining the position parameters. This enables a descriptive characterization and evaluation of the observations. The variable l represents the threshold value for manual intervention (cf. Sect. 3.2). The adjusted boxplot is used because it takes into account the skewness of a distribution when calculating the whiskers by using the Medcouple (MC) [31]. This is done to avoid the incorrect identification of many potential outliers in case of skew distributions. The MC is calculated according to Brys as follows [32]:

$$MC(X) = \text{med}_{x_i \leq x_{\text{med}} \leq x_j} h(x_i, x_j) \tag{1}$$

with

$$X = \{x_0 \geq x_1 \geq \dots \geq x_n\} \tag{2}$$

$$x_i \neq x_j \tag{3}$$

$$h(x_i, x_j) = \frac{(x_j - x_{\text{med}}) - (x_{\text{med}} - x_i)}{x_j - x_i} \tag{4}$$

X is the sorted sample with n independent observations; x_{med} [h] is the median of the sample X ; $h(x_i, x_j)$ is the medcouple matrix.

The whisker interval is calculated as follows.

For $MC \geq 0$

$$[Q_1 - 1, 5e^{-4MC}IQR; Q_3 + 1, 5e^{3MC}IQR] \tag{5}$$

For $MC \leq 0$

$$[Q_1 - 1, 5e^{-3MC}IQR; Q_3 + 1, 5e^{4MC}IQR] \tag{6}$$

IQR is the interquantile distance; Q_1 is the lower quantile; Q_3 is the upper quantile.

3.2 Data preparation

For data preparation, the observations are compared with the previously determined limit values. Observations that lie outside the whiskers represent extreme values and, thus, could have a negative influence on the prediction quality of the model. Possible causes for such extreme values are:

1. The observations represent correct MTF durations of the MTF class considered, they are real extreme values. An annotation is internally made in the model, so that these points are no longer considered when comparing with the threshold value l .
2. The observations represent the correct MTF duration for the failure class considered but a modification (e.g. maintenance measure, improvement) has taken place on the machine, which has caused a significant change in the failure distribution. The data collected up to the time of modification are excluded from further prognosis.
3. The observations are outliers which do not reflect the distribution of the failure duration as expected but are influenced by further effects (e.g. input errors). These values can be removed or corrected by the employee from the dataset.
4. The observations are classified as extreme values because they belong to a different MTF class. In this case, it is possible for the employee to transfer the observations to another or a new class.

According to the recommendations in the literature, case 2 to case 4 lead to manual interventions by the machine operator or production planner. To keep the interventions at an appropriate level, the number of potential outliers u is compared with the threshold value l . If there are enough observations for the failure class considered and at least l observations outside the whiskers at the same time, the employee is made aware of this fact. For this purpose, the adjusted box plots of the failure classes are displayed and the employee can qualitatively evaluate the extreme values. After a manual intervention in the data sets, a new check is initiated for the affected failure classes, as long as that sufficient observations are available.

3.3 Modelling

If the considered data set passes the test for extreme values, data modelling follows. For individual MTF classes, it can be assumed on the basis of the central limit value theorem that a production system consisting of overlapping stochastic input variables results in normally distributed output variables [33]. This assumption is tested during the modelling step using the Shapiro–Wilk test (SW test) [34]. Depending on the results, either the median or the mean value of the

data set under consideration is used for MTF duration prognosis. The tested data set deviates significantly from a normal distribution, if the null hypothesis of the test is rejected and the alternative hypothesis is assumed. In this case the robust duration of MTF prediction is based on the median because of its high break point compared to the mean value. This prevents individual atypical MTF durations from influencing the forecast quality and can have a positive influence on the prognosis quality [35]. In addition, the median represents the entry value with the highest probability in unimodal distributions, which is not normal or skew, and is therefore usually preferred over the mean value [36].

3.4 Evaluation and interpretation

Lastly, the statistical significance of the position parameter is examined. For normally distributed data, Kröning's method for accuracy prediction is used. Accuracy e is calculated as follows [37]:

$$\frac{e}{2} \geq \frac{s}{\sqrt{n}} \cdot t_{\left(1-\frac{\alpha}{2}; n-1\right)} \quad (7)$$

based on the confidence interval

$$\left[\bar{x} - t_{\left(1-\frac{\alpha}{2}; n-1\right)} \cdot \frac{s}{\sqrt{n}}; \bar{x} + t_{\left(1-\frac{\alpha}{2}; n-1\right)} \cdot \frac{s}{\sqrt{n}} \right] \quad (8)$$

e [h] is the forecast accuracy; n is the number of observations; α is the error probability; s [h] is the standard deviation of all MTF durations; \bar{x} [h] is the mean value of all MTF durations; $t_{\left(1-\frac{\alpha}{2}; n-1\right)}$ is the value of t distribution.

For not normally distributed data observations, j and k are identified, which correspond to $\alpha/2 = \alpha_j$ or $1 - \alpha/2 = \alpha_k$ of the cumulated binomial distribution $B(n, 0.5)$ [38]. Since the binomial distribution is a discrete distribution, the limits of the confidence interval for the median are additionally linearly interpolated. In this way, the required confidence probability of $1 - \alpha$ is achieved. However, if identical values occur at the neighboring ranks of j and k to the respective ranks themselves, an interpolation may lead to false results. In this case, the confidence interval without interpolation is used.

If the minimum of one is defined as the intervention limit for the threshold value l , the iteration ends at this point. If the confidence interval was able to maintain the required accuracy, a reliable forecast value is available. If the intervention limit is set as higher than one, this leads to a further loop-like check. From all permissible observations of the current MTF class, the most recent observation is removed in l iterations and the verification of the confidence interval for the required accuracy is repeated. The purpose of this is to secure the forecast against potential outliers whose

number u is below the threshold value l . At the same time, the volatility of the confidence interval in stochastic observations is compensated because the randomness of MTF durations does not permit a strictly monotonously decreasing confidence interval with an increasing number of observations. Thus, even with a high intervention limit and the ranking-based confidence interval around the median, the quality of prognosis is assured.

4 Experimental evaluation

First, the functionality of the developed methodology for prognosis of MTF durations and, in particular, its practical applicability is tested. For this purpose, it is examined whether the method fulfils the following characteristics:

1. *Quantity of data* The prediction of MTF duration should be able to deliver trustworthy results even with small quantities of data, so that a fast adoption for planning and scheduling processes is made possible.
2. *Data quality* Against the background of stochastic effects of a real production, the forecast should be able to identify and process incorrect values. At the same time, the system should be able to adequately process atypical behavior.
3. *Interaction* Interaction with the user should be reduced to a minimum to save resources. However, this must not contradict data quality. Atypical behavior must trigger an interaction quickly.

Subsequently, the effect of the more precise prognosis options on corporate key figures is examined using an exemplary production scenario.

4.1 Model functionality and practical applicability

For the first part of the evaluation, different configurations of the model are considered as part of sensitivity analyses. Effect direction of parameter changes becomes apparent and problematic configurations are identified. In addition, various generated distributions are used.

4.1.1 Model behavior with small amounts of data

Particularly with small amounts of data, the choice of an appropriate error probability α is subject to a conflict of objectives. On the one hand, a high probability of error makes it possible to obtain a forecast value for small amounts of data. On the other hand, an error probability that is too high contradicts the objective of obtaining reliable values for the forecast. The aim of first investigations is therefore to clarify whether a situation that influences the

choice of the confidence probability $1 - \alpha$ exists or enables specific recommendations for action for its choice. Therefore, the mean and median value of a normal distribution as a function of standard deviation σ , error probability α and number of observations n are investigated. The detailed test configurations are shown in Table 1.

The investigations of the confidence interval around the mean value for normal distributions with different scattering initially show only trivial correlations, which confirm the basic functionality of the method.

For the confidence interval around the median, the case without interpolation was focused on because here the actual confidence probability does not necessarily correspond to the required one. The test results presented in Table 2 show the actual confidence probability $\alpha_k - \alpha_j$, with the lower rank j and the upper rank k of $x(i)$, which represent the limits of the confidence interval. The number of observations n_α must be available for all four probabilities of error examined to ensure that they do not fall below the required confidence probability $1 - \alpha$. These are indicated by the dotted lines in Table 2. To ensure the functionality of the prognosis model during the evaluation part (cf. Fig. 1 and Sect. 3.4), n_α has to be corrected by the expression $(l - 1)$. As a result, the

required minimum number of observations n_{\min} is calculated as follows:

$$n_{\min} = n_\alpha + l - 1 \tag{9}$$

n_{\min} is the minimum number of observations; n_α is the number of observations to ensure confidence probability without interpolation.

The graphical comparison of confidence and error probability as well as the number of observations in Fig. 3 demonstrates that this problem can be disregarded by the method from $\alpha = 0.125$, since at least four observations are inevitable for a forecast.

However, there is no specific restriction of the use of the method based on the facts presented because the method also ensures compliance with the confidence probability for $\alpha < 0.125$. It is only necessary to individually examine whether a deviation from the standard value $\alpha = 0.05 - 0.0625$ is logical as the minimum number of observations can then be reduced by one.

4.1.2 Interaction effort with different data quality

The choice of the intervention threshold l and the distribution function of the MTF have a decisive influence on the interaction effort. The latter depends on the considered machine tool as well as other framework conditions of a company (e.g. capacities for maintenance) and can therefore not be directly influenced. On the other hand, the threshold value l represents a configuration variable of the forecast model. Therefore, the effect of different threshold values on the resulting interaction effort is investigated. Detailed information on the experimental settings are shown in Table 3.

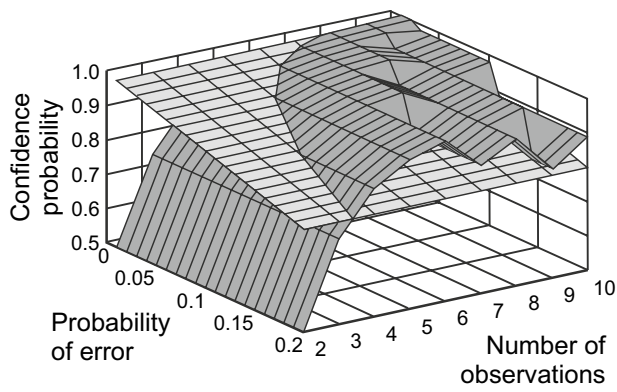
The main experimental results are summarized in Fig. 4. The probability for none, one and two manual interventions is shown as a function of the threshold value l .

Table 1 Test configurations with small amount of data

Para-meter	Configurations	Description
n	2, ..., 10	Number of observations
α	0.01	Very low probability of error
	0.05	Standard probability of error
	0.10	High probability of error
	0.20	Very high probability of error
σ	2	Small scattering
	4	Medium scattering
	8	Large scattering

Table 2 Ranks and confidence probabilities for the confidence interval around the median without interpolation

N	Significance level α											
	0.01			0.05			0.10			0.20		
	j	k	$\alpha_k - \alpha_j$	j	k	$\alpha_k - \alpha_j$	j	k	$\alpha_k - \alpha_j$	j	k	$\alpha_k - \alpha_j$
2	1	2	0.5000	1	2	0.5000	1	2	0.5000	1	2	0.5000
3	1	3	0.7500	1	3	0.7500	1	3	0.7500	1	3	0.7500
4	1	4	0.8750	1	4	0.8750	1	4	0.8750	1	4	0.8750
5	1	5	0.9375	1	5	0.9375	1	5	0.9375	1	5	0.9375
6	1	6	0.9688	1	6	0.9688	1	6	0.9688	1	6	0.9688
7	1	7	0.9844	1	7	0.9844	1	7	0.9844	2	6	0.8750
8	1	8	0.9922	1	8	0.9922	2	7	0.9297	2	7	0.9297
9	1	9	0.9961	2	8	0.9609	2	8	0.9609	3	7	0.8203
10	1	10	0.9980	2	9	0.9785	2	9	0.9785	3	8	0.8906
	$n_{\min} = 8 + l - 1$			$n_{\min} = 6 + l - 1$			$n_{\min} = 5 + l - 1$			$n_{\min} = 4 + l - 1$		



Type of distribution

Light grey: Required confidence probability
 Dark grey: Minimum, actual confidence probability without interpolation

Fig. 3 Influence of the probability of error on the minimum number of observations

Since all distributions are free of actual outliers, ideally no interventions should be necessary. Regardless of the very different number of manual interventions for different scatters of the normal distributions (4007 for $\sigma = 4$, 4053 for $\sigma = 8$ and 4845 for $\sigma = 2$), the probability of no necessary manual intervention for all three distributions examined is highest from $l = 5$. This also implies that, in the worst case, outliers are only detected if five extreme values are already present. The choice of $l = 3$ as threshold value is therefore more practical. Here, the relative frequency for a single manual intervention is highest for the first time.

The model behavior at Weibull distributed input values is similar to that at normal distributed values. However, for the logarithmic normal distribution, the limit value for no manual intervention is $l = 6$ and the threshold value for one manual intervention is $l = 4$. Since the adjusted box plot takes into account the skewness of the distribution, this cannot be the only cause for the slightly deviating

model behavior. The extent to which the scattering behavior could be the cause is still to be examined.

Lastly, it should be noted designing data sets that can lead to undesired, but mathematically correct model behavior is possible. Thus, the median for a data set with two perfectly alternating growing centers, which differ exclusively in the mean value, lies exactly in the middle of the centers at every second iteration. This corresponds to the mean value between the two centers and does not allow adequate differentiation. In addition to the accuracy e , the threshold value l also has a significant influence on the prevention of this problem. For $l > 3$, the two distributions add up enough observations to cancel out the detection of skewness with further uniform growth. In this case, the box plot is not able to provide sufficient differentiation via the parameters used.

To investigate the extent to which these theoretical limitations influence the practicability of the methodology and to estimate its practical potential, the model behavior is analyzed in a use case in the next step. According to the results above, for this α is set to 0.0625 and l is set to 3.

4.2 Exemplary application of the model

The added value of a more accurate prediction possibility of MTF durations is investigated on the basis of the production of a sample component. Details on the use case are given in Fig. 5. Oladokun’s classification into the three fault classes “short term” (≤ 0.5 h), “medium term” (0.5–2 h) and “long term” (≥ 2 h), which has been available so far as prediction model, is used as reference for evaluating the newly introduced methodology [14].

To evaluate the planning quality, a labelled training data set of MTF reports for six exemplary MTF classes (cf. Table 4) with 1000 fault events each is generated and randomly mixed. This data set is manipulated in such a way that the number of correct messages (assignment of MTF class to MTF duration) is approximately 50%. This simulates the poor quality of manual production data collection (PDC) inputs which frequently occurs in practice.

Table 3 Test configuration for different distribution functions

Parameter/distribution	Configurations	Description
Normal (rounded up, with seed)	$\sigma = 2$	Small scattering
	$\sigma = 4$	Medium scattering
	$\sigma = 8$	Large scattering
Log-normal	$\mu = 0$	Non-normal distribution
	$\sigma = 0.5$	
Weibull	Scale parameter = 128 Form parameter = 2	Investigation without failure class
α	0.05	Standard probability of error
n	100	Number of observations
Data sets	2500	Number of examinations per distribution

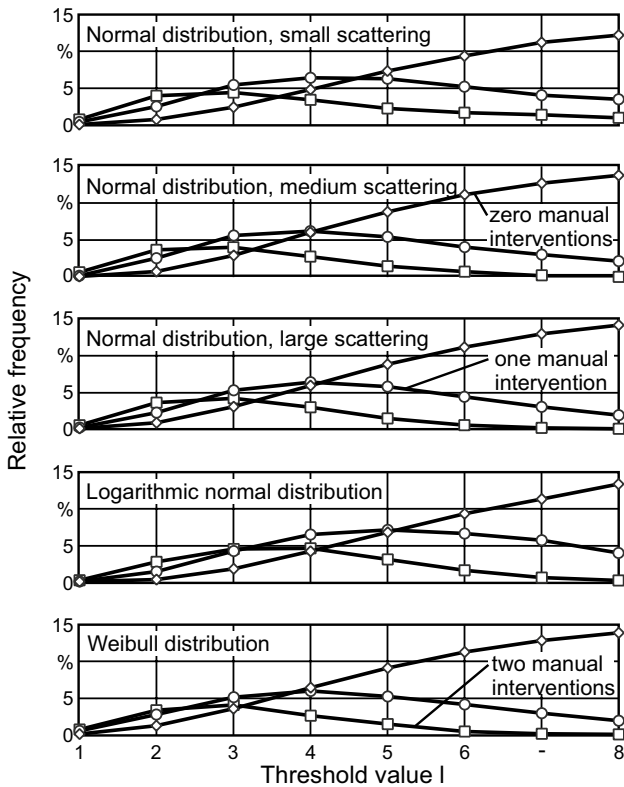


Fig. 4 Relative frequencies for different numbers of manual interventions

The prognosis model is then trained with the aid of the manipulated data set by entering the disturbance events one by one. The class assignment is corrected in the case of an automatic error detection. In the next step, the resulting forecast values x_{prog} are stored for planning in

a Manufacturing Execution System (MES) as a one-time MTF for the second milling machine. An exemplary production program with 40 orders of different batch sizes is created and the resulting lead times (LT) are calculated. In addition, the mean values of the disturbance classes according to Oladokun are used for further three planning scenarios.

To analyze the planning quality of all considered scenarios, production is simulated with the help of a material flow simulation model in Tecnomatix Plant Simulation. This digital twin of the considered production system takes into account the correct distribution function $s(x)$ for the six disturbance scenarios (cf. Table 4). To statistically secure the simulated LT, five simulation runs per scenario are performed.

The results show that an underestimated MTF duration due to a lack of accuracy of prognosis can quickly lead to large deviations between planned and actual LT. The analyses indicate that, with the new methodology, the disturbance durations used for forecasting are a good approximation of the real disturbance distribution, even if the quality of the manual PDC messages is medium. The real mean values are predicted with a quality of at least 97.3%. In comparison, the Olandokun model only achieves a quality of 70% with an ideal input database. This aspect was neglected in the comparison in Table 4, which is why the gained accuracy represents a minimum value.

Due to the very good interference duration prognosis, the predicted LT deviates from the real one by a maximum of 0.12% (absolute 0.4 h). With the old Olandokun model, the maximum deviation is 1.69% and 5.97 h is absolute. As more than one single fault event usually occurs during the production of an order in the production system, this reduction can already have a strong influence on the PPC performance of a company.

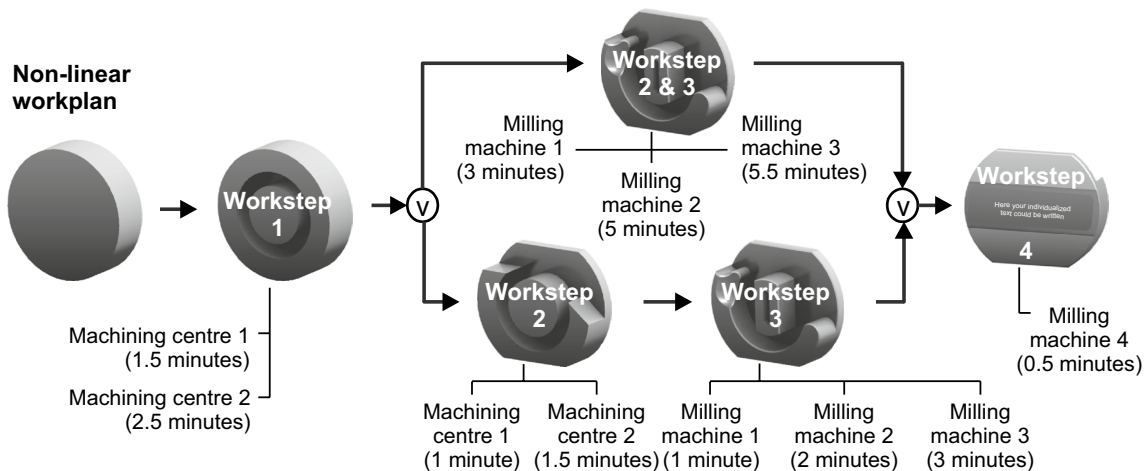
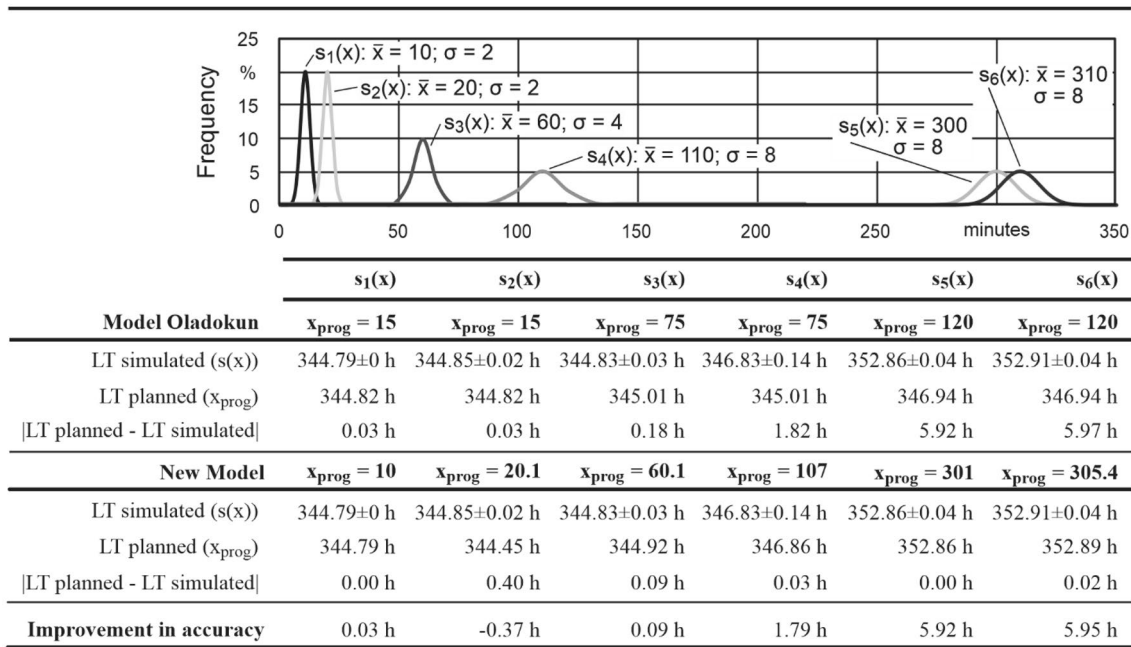


Fig. 5 Non-linear workplan of sample component production

Table 4 Improvement in lead time forecast by new MTF duration prognosis methodology



Wii/91159 ©IFW

5 Conclusion

The model developed within the scope of this work on the basis of the CRISP-DM concept is used for the reliable prognosis of MTF durations. Depending on the underlying MTF duration distribution, the median or mean value is used for the prediction. An input-parallel check of PDC inputs helps to avoid incorrect reporting and allows forecast accuracy of over 97%.

For practical validation and research, the prognosis model was tested with different configurations and different data sets. It has been proven that the model can handle small amounts of data as well as poor data quality. Additionally, it has been shown that $\alpha = 0.0625$ and $l = 3$ are suitable as default settings for the forecast model. An additional case study could prove that the presented methodology is applicable for LT prediction as the maximum variance between target and actual LT was 0.12% and thus clearly below the value of an alternative MTF forecast model.

The aim of further research is to examine the performance of the approach in the context of a more complex failure class distribution. The ability of the prognosis to cope with a higher number and closely arranged MTF classes is of significant interest. In addition, the required amount of input data should be an object of investigation. Moreover, the extent to which the model correctly recognizes error classes that occur in rare cases must be examined. Finally, the transferability of the method should

be investigated. The universal approach should be able to handle any kind of interruption (e.g. shortages of material) in the production.

Acknowledgements Open Access funding provided by Projekt DEAL. The results presented in this paper were obtained within in the Collaborative Research Center 653 “Gentelligent Components in Their Lifecycle”. The authors would like to thank the German Research Foundation (DFG) for its financial support (Grant SFB653 T09).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Denkena B, Dittrich MA, Wilmsmeier S (2019) Automated production data feedback for adaptive work planning and production control. International Conference on Changeable, Agile, Reconfigurable and Virtual Production. Procedia Manuf 28:18–23. <https://doi.org/10.1016/j.promfg.2018.12.004>
- Abele E, Elzenheimer J, Liebeck T, Meyer T (2006) Globalization and decentralization of manufacturing. In: Dashchenko AI (ed)

- Reconfigurable manufacturing systems. Springer, Berlin, pp 3–13. https://doi.org/10.1007/3-540-29397-3_1
3. Nyhuis P, Wulf S, Klemke T, Benjamin B (2010) Integrative factory, technology, and product planning—systemizing the information transfer on the operational level. *Prod Eng Res Dev* 4:231–237. <https://doi.org/10.1007/s11740-010-0225-6>
 4. Mehrabi MG, Ulsoy AG, Koren Y, Heytler P (2002) Trends and perspectives in flexible and reconfigurable manufacturing systems. *J Intell Manuf* 13:135–146. <https://doi.org/10.1023/A:1014536330551>
 5. Isermann R (2006) *Fault-diagnosis systems*. Springer, Berlin. <https://doi.org/10.1007/3-540-30368-5>
 6. Swanson L (2001) Linking maintenance strategies to performance. *Int J Prod Econ* 70(3):237–244. [https://doi.org/10.1016/S0925-5273\(00\)00067-0](https://doi.org/10.1016/S0925-5273(00)00067-0)
 7. Sauer H (1987) *Mengen- und ablauforientierte Kapazitätssplanungen von Montagesystemen*. Springer, Berlin. <https://doi.org/10.1007/978-3-642-83100-3>
 8. Vieira GE, Herrmann JW, Lin E (2000) Predicting the performance of rescheduling strategies for parallel machine systems. *J Manuf Syst* 4:256–266. [https://doi.org/10.1016/S0278-6125\(01\)80005-4](https://doi.org/10.1016/S0278-6125(01)80005-4)
 9. Al Khateeb A (2014) *Planung und Optimierung von Fließmontagesystemen*. Technische Universität Ilmenau, Dissertation
 10. Heinz K (1969) *Mathematisch-statistische Untersuchungen über die Erlang-Verteilung*. Westdt Verl, Köln
 11. Bergmann S (2013) *Automatische Generierung adaptiver Modelle zur Simulation von Produktionssystemen*. Technische Universität Ilmenau, Dissertation
 12. Moohialdin ASM, Hadidi LA (2017) Effect of failure type on downtime duration for a manufacturing facility. *J Perform Constr Facil* 3:4016114. [https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0000983](https://doi.org/10.1061/(ASCE)CF.1943-5509.0000983)
 13. Cassady CR, Kutanoğlu E (2005) Integrating preventive maintenance planning and production scheduling for a single machine. *IEEE Trans Reliab* 54(2):304–309. <https://doi.org/10.1109/TR.2005.845967>
 14. Oladokun VO, Charles-Owaba OE, Nwaouzru CS (2006) An application of artificial neural network to maintenance management. *J Ind Eng Int* 2(3):19–26. <http://ir.library.ui.edu.ng/handle/123456789/1774>
 15. Schenk M (2010) *Instandhaltung technischer systeme*. Springer, Berlin. <https://doi.org/10.1007/978-3-642-03949-2>
 16. El-Thalji I, Jantunen E (2015) A summary of fault modelling and predictive health monitoring of rolling element bearings. *Mech Syst Signal Process* 60–61:252–272. <https://doi.org/10.1016/j.ymssp.2015.02.008>
 17. Ambani S, Li L, Ni J (2009) Condition-based maintenance decision-making for multiple machine systems. *J Manuf Sci Eng*. <https://doi.org/10.1115/1.3123339>
 18. Ji W, Wang L (2017) Big data analytics based fault prediction for shop floor scheduling. *J Manuf Syst* 43(1):187–194. <https://doi.org/10.1016/j.jmsy.2017.03.008>
 19. Abu-Samah A, Shahzad MK, Zamai E (2017) Bayesian based methodology for the extraction and validation of time bound failure signatures for online failure prediction. *Reliab Eng Syst Saf* 167:616–628. <https://doi.org/10.1016/j.res.2017.04.016>
 20. Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech Syst Signal Process* 20(7):1483–1510. <https://doi.org/10.1016/j.ymssp.2005.09.012>
 21. Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D (2014) Prognostics and health management design for rotary machinery systems - Reviews, methodology and applications. *Mech Syst Signal Process* 42(1–2):314–334. <https://doi.org/10.1016/j.ymssp.2013.06.004>
 22. Teti R, Jemielniak K, O'Donnell G, Dornfeld D (2010) Advanced monitoring of machining operations. *CIRP Ann* 59(2):717–739. <https://doi.org/10.1016/j.cirp.2010.05.010>
 23. Martin KF (1994) A review by discussion of condition monitoring and fault diagnosis in machine tools. *Int J Mach Tools Manuf* 34(4):527–551. [https://doi.org/10.1016/0890-6955\(94\)90083-3](https://doi.org/10.1016/0890-6955(94)90083-3)
 24. Peng Y, Dong M, Zuo MJ (2010) Current status of machine prognostics in condition-based maintenance—a review. *Int J Adv Manuf Technol* 50(1–4):297–313. <https://doi.org/10.1007/s00170-009-2482-0>
 25. Goyal D, Pabla BS (2015) Condition based maintenance of machine tools—a review. *CIRP J Manuf Sci Technol* 10:24–35. <https://doi.org/10.1016/j.cirpj.2015.05.004>
 26. Valenti G, Lelli M, Cucina D (2010) A comparative study of models for the incident duration prediction. *Eur Transp Res Rev* 2(2):103–111. <https://doi.org/10.1007/s12544-010-0031-4>
 27. Hachesu PR, Ahmadi M, Alizadeh S, Sadoughi F (2013) Use of data mining techniques to determine and predict length of stay of cardiac patients. *Healthc Inform Res* 19(2):121–129. <https://doi.org/10.4258/hir.2013.19.2.121>
 28. Tsai PF, Chen PC, Chen YY, Song HY, Lin HM, Lin FM, Huang QP (2016) Length of hospital stay prediction at the admission stage for cardiologic patients using artificial neural network. *J Healthc Eng*. <https://doi.org/10.1155/2016/7035463>
 29. Master N, Scheinker D, Bambos N (2016) Predicting Pediatric Surgical Durations. *ArXiv e-prints*. <http://arxiv.org/pdf/1605.04574v2#>. Accessed 11 Nov 2019
 30. Kuhnle A, Kuttler M, Dümpelmann M, Lanza G (2017) *Intelligente Produktionsplanung und -steuerung - Erlernen optimaler Entscheidungen*, vol 107. Springer-VDI Verlag GmbH & Co. KG, Düsseldorf, pp 625–629
 31. Hubert M, Vandervieren E (2008) An adjusted boxplot for skewed distributions. *Comput Stat Data Anal* 52(12):5186–5201. <https://doi.org/10.1016/J.CSDA.2007.11.008>
 32. Brys G, Hubert M, Struyf A (2004) A Robust Measure of Skewness. *J Comput Graph Stat* 4:996–1017. <https://doi.org/10.1198/106186004X12632>
 33. Schmidt KD (2011) *Maß und Wahrscheinlichkeit*. Springer, Berlin
 34. Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52(3–4):591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
 35. Thode HC (2002) *Testing for normality*. Dekker, New York
 36. von Hippel PT (2005) Mean, median, and skew—correcting a textbook rule. *J Stat Educ*. <https://doi.org/10.1080/10691898.2005.11910556>
 37. Kröning S (2014) *Integrierte Produktions- und Instandhaltungssplanung und -steuerung mittels Simulationstechnik*. Universität Hannover, Dissertation
 38. Conover WJ (1999) *Practical nonparametric statistics*. Wiley, New York
 39. Shearer C (2000) The CRISP-DM model: the new blueprint for data mining. *J Data Warehous* 5(4):13–22

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.