Proceedings of the 5th European International Conference on Industrial Engineering and Operations Management
Rome, Italy, July 26-28, 2022

Survival Analysis-Based Delay Prediction for Production Line with Irregular Events

Irene Erlyn Wina Rachmawan, Issei Suemitsu, Matsuba Hiroya

Intelligent Information Research Department
Center for Technology Innovation
Hitachi, Ltd
Japan
irene.rachmawan.fv@hitachi.com, issei.suemitsu.rj@hitachi.com,
hiroya.matsuba.do@hitachi.com

Abstract

Production planning involves scheduling a series of operations in a factory to produce products on time. However, with the current increasing trend of customer demands for customisation where many variations are required, production scheduling is becoming increasingly difficult. High-mix low-volume (HMLV) manufacturing plays significant role to satisfy this trend since it has flexibility through automation to produce many different products in small quantities. Yet, such automated factories are prone to irregular events (e.g. line stoppages, machine breakdowns), making it a challenge to keep production activities on schedule. It is therefore important to consider irregular events as early as possible in the production planning stage to mitigate their impact. This paper proposes a delay prediction based on survival analysis to predict irregular events and calculate the disruptions so that they can be anticipated through the generation of robust production plans which can be executed at any time. This approach can reduce production losses due to production delays caused by irregular events.

Keywords

Irregular Events, Delay Prediction, Robust Production Plan, Production Simulation, Survival Analysis.

1. Introduction

The increasing customer trend for customization accelerates high demands for flexible manufacturing systems (FMS) to produce a large variety of products in small quantities (Qiao, Lu and McLean 2006). Such production systems are known as high-mix low-volume (HMLV) production, and the introduction of an FMS is suitable for increasing efficiency, bottleneck suppression and lead-time reduction (Reinhart and Geiger 2011). However, FMSs are usually complex and prone to irregular events such as unplanned breakdowns and line stoppages (Javaid et al. 2022). In practice, these events can disrupt the production process and the original production plan may not be realised. For example, a cause-and-effect relationship that results in a slightly longer production time may actually cause production delays, resulting in products not being delivered to customers in time. Production planning attempts to optimally schedule all production processes in advance of the production process. However, if irregular events that disrupt the production process are not taken into account, productivity will be low.

In such cases, robust production planning can be implemented to minimise the impact of irregular events on the production process and ensure that the production plan can be implemented at any time. Robust production planning aims to assure the smoothness of production in the face of disturbances and to maintain target performance indicators in an unstable environment (Gyulai et al. 2015). Robustness in production planning involves a sophisticated approach to dealing with predictable or unpredictable changes and disturbances. It can either respond to the occurrence of random events (reactive approach) or protect the performance of the plan by predicting the occurrence of irregular events to some extent (proactive approach) (Campagne et al. 1995). In reactive scheduling, production schedules are not created in advance. Instead, it is usually determined in real time when required, according to various priorities of production requirements, such as minimum delivery time, minimum processing time. Proactive scheduling focuses on creating predictive schedules to meet performance requirements in a dynamic manufacturing environment. In this sense, proactive approach could be more beneficial to produce a robust production plan with acceptable performance, even if unexpected disruptions occur during the execution of the plan. Efficient ways to account for uncertainty and achieve more robust planning through proactive approach include applying probabilistic models, the use of adaptive and collaborative approaches. This is because deterministic models often fail to provide viable plans due to the presence of uncertainty. For this reason, simulation is often used as an essential tool to assist in the creation of robust production plans. The reason for using simulation is the limitation of the computational complexity of the constraints of the simulation process. In the creation of robust production plans, simulation optimisation is mostly applied by repeatedly exploring scenarios that could occur in the actual production process. Based on the results of the simulation experiments, the relevant parameters of the production process are then adjusted until the target values of the performance indicators are reached.

Increasing the robustness of production planning may imply a trade-off with specific production requirements (Gyulai et al. 2015). In such cases, key performance indicators are assessed to measure the quality of the production plan by comparing the intended output of the production plan with the actual output. The most commonly observed directly related to disturbances due to irregular events is the lateness indicator, which measures the difference between the actual production order and the completion of the plan, expressed in terms of time. Lateness measures whether the plan is not sufficiently robust. It can therefore characterise robustness more efficiently, as it is strongly dependent on the execution of the plan. Thus, measuring the cost of robustness can ensure the production plan's quality when implemented in the actual production process. The robustness of the plan is influenced by the ability to encounter irregular events whose probability of occurrence varies widely across production settings, which can lead to inaccuracies in specified mathematical modelling. Hence, the objective of the proposed framework is to drive strategy to generate a robust production plan with data-driven situational awareness of production variability.

This paper proposes a method to predict irregular events that may occur in actual production activities and support the creation of robust production plans to avoid production delays. The paper presents the integration of a novel survival analysis-based machine learning method with discrete event simulation (DES) to develop a delay prediction system that considers disturbances caused by irregular events, thereby creating robust production plans. The proposed method can be used for planning and scheduling manufacturing execution (MES) and enterprise resource planning (ERP) systems to calculate robust production plans. In addition, production planners can use the results to ensure the execution of the production plan at any time. Compared to robust conventional optimisation methods and methods based on iterative simulation, the proposed method maintains the simplicity of the planning algorithm and thus the short execution time, leading to low delays and making it applicable in a real industrial environment.

The remainder of this paper is organised as follows. Section 2 presents a literature survey on traditional methods for predicting production delays, while Section 3 describes the proposed modelling approach based on survival analysis. Section 4 presents the detailed procedure for its implementation. Finally, Section 5 presents the experimental results of the proposed approach, and Section 6 concludes the paper.

2. Literature Review

The accumulation of latency of instructions finishing times due to irregular events causes production delays. To avoid production delays, planners must develop robust production plans to minimise the effects of disturbances caused by irregular events. There are two main approaches to consider these issues: the predictive reactive approach and the proactive approach. The predictive reactive approach has two main steps. The first step is to develop a primary production schedule. The second step is to anticipate potential disturbances and generate a robust plan that can be implemented at any time. Predictive schedules are usually built on knowledge of irregular events. Simulation is a common method for estimating such delays. Simulation models can explicitly represent system variability, interconnectivity and complexity. As a result, simulation enables prediction of system performance, comparison of alternative system models and determination of the impact of alternative policies on system performance. One simulation method that can be used for modelling production processes involving irregular events is the discrete event simulation model. Discrete event simulation models are a simplified representation of a system developed to understand system performance over time and to identify potential improvements.

In the simulation context, the type of distribution such as Gaussian, Weibull, etc., is generally required, as described in previous studies by (Ilar and Powell 2008), (Liu, Gu and Xi 2007). A two-stage hybrid genetic algorithm is proposed to generate predictive schedules for flexible job shops (Hinai and ElMekkawy 2011). In addition, a knowledge-based evolutionary algorithm is studied for handling proactive scheduling of stochastic machine failures due to a deteriorating production environment (Wang et al. 2015). Other research also investigates the use of simulation to analyse the performance of existing scheduling rules and propose a simulation-based real-time scheduling mechanism for dynamic discrete manufacturing generate simulation models at high speed and prepare Negahban and Smith (2014). An integrated software module has been developed that allows simulation experiments to be carried out for rapid analysis of the schedule as introduced by (Grzegorz et al. 2016).

In optimisation, an optimisation-based predictive modelling control was developed that allows the scheduler to solve both constraint-aware production optimisation and in-process inventory management problems at each scheduling instance (Jang et al. 2013)Other research proposes a two-stage particle swarm optimization to solve flexible job shop scheduling problems under machine breakdowns (Nouiri et al. 2017). Despite the stability considerations, a definition of a irregular event distribution is required, this kind of distributions is practically unknown and difficult to model. Estimation of distributions using basic statistics is inadequate because it does not consider covariates of the most recent production situation (e.g. product type, production volume, etc.) Estimating distributions for HMLV production is more difficult because there are so many variations of products, and the availability of historical data is usually limited to all possible variations. This makes it more difficult to estimate the distribution of HMLVs, since the number of HMLVs produced is very large. Therefore, new techniques are needed to correctly estimate the distribution in a practical way.

On the other hand, a current technological trend in both academia and industry is predictive maintenance, where unexpected machine failures are predicted (Cachadav et al. 2018). However, its focus is on preventive measures to actively respond to breakdowns when they occur and does not consider their impact on the actual production process. In this approach, predictive maintenance studies generally do not consider the impact of disturbances caused by those failures, which is an important issue in actual production planning. Another study attempts to use a more flexible approach to robust production planning using non-parametric based survival analysis. This approach does not require any pre-defined distribution for the production process simulation (Łukasz Sobaszek 2017). However, such a non-parametric algorithm always requires historical data for all items, which is insufficient to deal with data not present in the historical data. Therefore, this paper proposes a survival analysis based on CoX regression to model random events independent of the amount of data and estimate any item's distribution from its associated parameters.

3. Proposed Method and Evaluation Metrics

This section discusses the methods for predicting irregular events in the production process and sequentially predicting production delays which employs a simulation-based system, The irregular event prediction module has the means to estimate three factors: the occurrence, start time and duration of the irregular event. In this study, the production delay prediction module relies on a simulation model used for multiple purposes: to assess the viability of the initial production plan when irregular events are expected to occur and provide a realistic response for building reliable production capacity forecasting models. Therefore, the direct link between the simulation model and the physical system is maintained, providing reliable prediction results without direct user interaction.

Figure 1 shows the schematic structure of the methodology of production delay prediction using survival analysis to predict the irregular events to calculate production delays. Our proposed method utilizes the historical event data which collected in various way (e.g., various sensors installed on the machines). The proposed method then learn the historical data to predict three factors of irregular events. In this paper, we employ a machine learning approach based on survival analysis to predict the onset and start time of an event. Meanwhile, a kernel density function is used to predict the duration of events, which samples values from a data distribution. Both methods learn appropriate information from historical data to model irregular events. The combination of these two modules generates a list of irregular events as input to the production simulator. It calculates the risk of production delays based on the disturbances caused by these irregular events. Two matrices are introduced to measure the quality of the forecasting results in this study from a technical and a business perspective.

This approach allows the forecasting results to be evaluated. A discussion of the proposed approach and the evaluation metrics is provided in the following subsections.

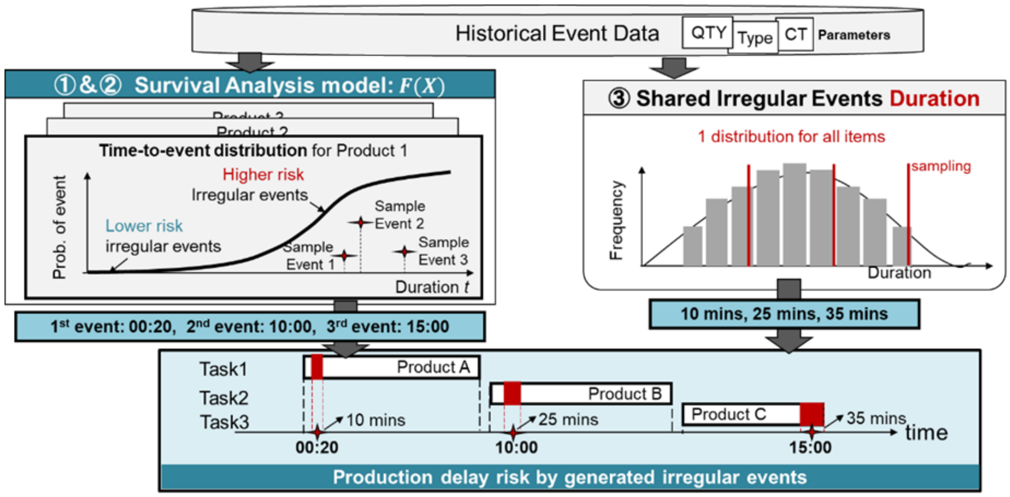

Figure 1. General view of simulation-based production delay prediction

3.1. Predicting production delays

The risk of production delays is estimated based on the disturbance that the predicted irregular event will cause to the original production plan. Irregular events occur arbitrarily, but their probability of occurrence may be influenced by factors such as the amount of production, type of product, and other production parameters. Irregular events comprise three components - occurrence, start time and duration - that must be predicted. Two approaches are proposed to predict these: machine learning based on survival analysis predicts an event's occurrence and start time, while a module based on kernel density predicts the duration. Survival analysis is particularly suited to predicting the time to event occurrence, as it learns how each covariate affects the remaining production time until the event occurs. The duration of each event is modelled by sampling the historical distribution. After predicting irregular events, the production simulation estimates the likelihood of production delays caused by them.

A. Modelling event starting time as time to event distribution using Survival Analysis

As mentioned above, irregular events can occur arbitrarily, but their probability of occurrence is influenced by various production settings. Conventionally, it has been necessary to define probability distributions to predict the occurrence of irregular events. However, modelling such distributions is difficult. Due to limited data, it is difficult to estimate the probability distribution of irregular events in complex production configurations, such as high-mix low-volume production.

Therefore, in order to create probability distributions even with limited data, a machine learning-based survival analysis model is proposed to model the probability of an event occurring, and its start time as a Time-to-Event (TTE) distribution. The TTE distribution is a probability distribution that indicates when an event will occur at a particular time. This combined model allows all data sets to be used without splitting them into separate data. Instead, data points are mapped as TTE data points, which indicate the time span until an event occurs. A TTE distribution is generated from these data points that show the probability distribution of when an event is at high risk of occurring. In other words, survival time analysis models the time until an event occurs. Therefore, the TTE distribution indicates the probability that an event will occur later than a certain specified time.

In the learning process of survival analysis, an initial state to an end state is defined in order to model the distribution. The main observation when modelling the distribution is the length of time it takes for the event of interest to occur during its lifetime. In the case of a production process, this lifetime means the length of time it takes for an instruction to process a certain task, with the initial state being the start time of the instruction and the end state being the end time of the instruction. The event of interest, in this case, is an irregular event.

Equation 1 shows the basis of the survival function: t represents the response variable, where the response is the time to event occurrence; $\boldsymbol{S}(\boldsymbol{t})$ is always one at $\boldsymbol{t} = 0$; meaning that no irregular events occur at time 0. As the probability of an irregular event occurring is correlated with the production setting, it is necessary to consider the production parameters. For this purpose, a semi-parametric model, the so-called CoX Proportional-Hazards, was employed, allowing for flexible data handling.

$$S(t) = Prob\{T > t\} = 1 - F(t) \qquad (1)$$

The CoX Proportional-Hazards method investigates the influence of several variables on time to an irregular event. In this method, no specific survival distribution is assumed. Instead, a baseline function is considered assuming that the effects of the predictor variables on survival are constant over time and additive at one scale. The first baseline function of the TTE distribution is generated according to all data points to represent the time distribution to the general event. This is done by training all data points to create this baseline function. As a result, the Cox proportional hazards regression model can be written as in equation 2:

$$f_i(t) = f_0(t) \times Impact(X_i) \qquad (2)$$

Where $f_o(t)$ is the expected hazard at time t, $f_o(t)$is the baseline hazard, representing the hazard when the predictors (or independent variables) X_1, X_2, X_p are all equal to zero.

Once the baseline has been generated, the TTE distribution for individual production sequences can be constructed by considering the effect of the values on each product feature on the learned baseline function. Learning is done by extracting product features (e.g. production cycle time, quantity, category) and learning an baseline function (X_1). The impact function describes how the event risk per unit of time changes over time at the baseline level of the production setup. To calculate the impact function, the contribution of the likelihood of the target data to the baseline is considered. The parameters of the model can be estimated in a stochastic framework using a regression model. Under this framework, the probability distribution of the target variable is assumed, and then a likelihood function is defined that calculates the probability of observing the outcome given the data points and the baseline function. This function can then be optimised to find a set of parameters such that the sum of the likelihoods for the training data set is maximised.

This baseline function determines the curve of the TTE function for the individual product. Thus, even if there are a limited number of data points with historical data for the product of interest, the TTE function can be estimated because the products have similar characteristics, and the survival analysis assumes that their baseline functions are also similar. The TTE function for another data point is generated by comparison with the most similar product. Furthermore, this closest product's impact function is estimated based on its features. In other words, it is essential to note that the learned survival time method can estimate the TTE function for any item, irrespective of the amount of data, simply by taking the product feature X as input. Thus, the survival time method can overcome the problem of data limitations as it can estimate the TTE from similar product data even for items with zero data volume.

B. Modelling event duration time

In addition to predicting the event's occurrence and start time, the event's duration is another essential factor to be modelled. Once the predicted event starts times have been obtained, the duration of each event needs to be predicted. A sampling process generates event durations from the distribution of historical data. Kernel density estimation (KDE) is employed as the sampling method; KDE is a non-parametric probability density estimator representing the global event duration distribution. Kernel density estimation is closely related to histograms, but smoothness and continuity can be imparted by using appropriate kernels. Kernel density estimation (KDE) is suitable for estimating irregular event duration distributions due to its advantage of being able to estimate the actual density. Kernel density estimation (KDE) produces a distribution that is a positional mixture of kernel distributions. There are two steps to derive a value from the kernel density estimate: (1) derive a value from the kernel density, and (2) sample data points arbitrarily in order to generate the prediction results of events duration time. Thus, a cumulative density function is generated after estimating the density function, and a sample of event durations is drawn. The output of this step is a list of event durations in time units such as seconds.

3.2. Production Simulation

Once the predicted irregularities are obtained, the next step is to estimate the disturbance to the production plan. This disturbance causes delays because the completion time of certain tasks is longer than originally planned. Disturbances caused by irregular events are usually of the type where a single disturbance causes a chain of events in the production process and the cumulative effect is a chain reaction. Therefore, the prediction of production delays by simulation is based on the calculation of the risk of production delays from the accumulation of additional time to complete instructions after an irregular event has occurred. In this regard, two main essential inputs to the simulation are the original production plan and the predicted irregular events. First, the simulation calculates the disturbances and whether they will significantly impact meeting the planned date. The simulation initially reads the original production plan, which contains information on each process's start and end times. It then reads the distribution of TTEs and times for the irregular events and generates the time of occurrence, starts time and time of the event. Scenarios representing several patterns of irregular events are then generated. The more scenarios are generated, the more information can be aggregated to produce better prediction results. Therefore, the following rules are registered to simulate chain effects:

For all events, check whether there is an execution instruction.

- If there are, insert an irregular event and change the end time of the affected execution instructions.
- Check whether any subsequent instructions are scheduled to be executed before the end time predicted in b.
- If so, change the end time of the relevant execution instruction.
- Repeat step c until no more instructions are affected by the preceding delay.
- Repeat step c until the event is no longer present in the list of expected events.

After simulating the production process, including the cascading effects of disturbances in the event of irregular events, the risk of production delays is then calculated. This is done by checking all production end times for each scenario. If simulation results predict that instruction will have a delay in approximately 95% of all scenarios, the instruction is said to be a delayed instruction.

3.3. Evaluation methods

To measure the performance of a model, two aspects are observed. One is the technical aspect, and the other is the effectiveness aspect. The technical metric measures how well our model can predict irregular events, while the second aspect measures how well our model correctly predicts production delays.

A. Evaluation metric 1: Measure performance of proposed method with technical point of view-based metrics

To demonstrate the effectiveness of our technology from the technical point of view, we compare the instruction duration between actual production execution and prediction result. The calculation is carried out by the error ratio equation as shown by equation 3.

$$\varepsilon_{duration} = \frac{1}{N}\sum_{i}^{N}\frac{Pred.Duration - Act.duration}{Plan\ duration} \tag{3}$$

Where:
ε: error rate
N: Total data

Instruction execution time is the total time it takes for an instruction to complete, expressed in working time such as hours or minutes. This duration depends on the smoothness of the instruction execution and thus allows us to observe how irregular events disrupt each instruction. The primary motivation for measuring instruction durations is that it allows us to estimate how accurate our method of generating event occurrences and event durations is. The smaller the difference between the actual instruction execution time and our prediction, the better our method performs. This evaluation metric measures the ratio of inaccurate predicted time to the total time of the instructions evaluated. This metric was used as an evaluation metric to measure the effectiveness of the generated irregular events when tested on unseen data. In general, if the errors are small, one can be confident that the irregular event generation method presented in the previous chapter is working well.

B. Evaluation metric 2: Measure performance of proposed method with business-point of view-based metrics

Assessing correctly predicted production delays can be used to measure the effectiveness of the forecasting model from a business perspective. The assessment was made by comparing the total number of actual production delays with the total number of predicted production delays, as shown in Equation 4.

$$Prediction\ accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Where:
TP: True Predicted Delay
TN: True Non-Delay
FP: False Predicted Delay
FN: False Non-Delay

The actual labels of delayed instructions are constructed by comparing the end time of each instruction in the production plan with the actual production run. For example, an instruction that was scheduled to finish before the delivery date in the original production plan but is now past the delivery date in actual production. In this case, the instruction would be labelled as delayed instruction. The same applies to the labelling of the forecast result. The goal is to predict the total number of delayed instructions correctly. This indicator measures the ratio of correct predictions to the total number of delays evaluated. In this scenario, the total number of products that may be delayed can also be estimated by multiplying the total number of delay instructions by the number of products that one instructions should produce. The additional transport costs required to carry the delayed products can be estimated by converting the total number of delayed instructions into the total quantity of products. Therefore, it is estimated that the total delay-related costs can be reduced by correctly estimating the total delayed product.

4. Data Collection and Experimental Setup

In order to demonstrate the feasibility and robustness of the proposed production planning method, synthetic data is used which is generated from a real case study of an HMLV manufacturing company. The dataset consists of a synthetic production plan and synthetic production history data, including the execution of the production plan and the occurrence of irregular events. The dataset is generated close to the actual production situation by taking into account the insights of essential parameters of the production process, such as production volume, cycle time and product type. Figure 2 shows the production plan instructions generated and the correlated production history data. By using such synthetic data, the performance of the proposed method when implemented in a real production process can be demonstrated while maintaining customer confidentiality of the information. Synthetic production planning maintains information on production instructions based on production parameters. Irregular events were randomly generated. The data obtained was used to refine the standard schedule. Figure 2 shows that irregular events resulted in longer finishing times in the production plan, as recorded in instructions 1 and 5.

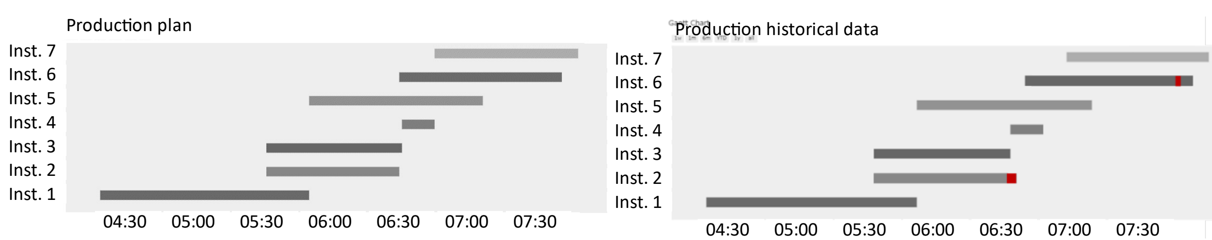

Figure 2. The production plan and associated production historical data with irregular events

After loading the dataset, an experiment was conducted. The initial state is to train a predictive model to estimate irregular events for future production planning. Before the learning process, the dataset has to be normalised. The training data must contain two important values: the execution time of the instructions and the event labels indicating the occurrence of irregular events. Once the training dataset has been built, training the model is quite straightforward.

First, the model is built based on survival analysis using CoX regression as described in the previous section. The fitting process is then performed by passing the dataset and the event columns, which are the labels needed to learn whether irregular events have occurred during production. Once the model has been trained, predictions can be made for unseen datasets based on the training results. Finally, the quality of the model is assessed based on the performance indicators described in the previous section.

5. Results and Discussion

The two main variables in this experiment are two responses: the duration of irregular events and production delays. Therefore, each reaction needs to be analysed separately. As a comparison, a similar analysis was carried out using Monte Carlo forecasts.

5.1. Irregular event generators

The risk of irregular events depends on the baseline and partial hazard. The model learns parameters to obtain a baseline function. The format of the dataset is shown in Figure 3 where the time indicates when the irregular event occurred and is a sequence of described information. The training dataset from the historical dataset of irregular events should be converted to a format for survival analysis. The solid line indicates when the relevant instruction containing the irregular event was under our observation, where the lifetime 0 means the start time of the normal processing of the production process, and the solid dot at the endpoint represents the irregular event that stops the production process.

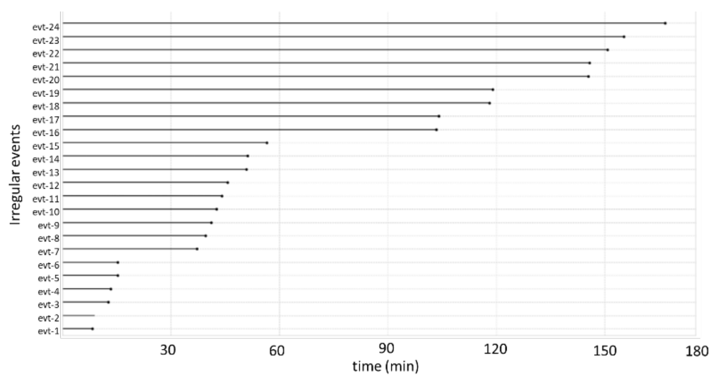

Figure 3. Data training for modelling TTE distribution of irregular events

The survival function from the historical data illustrated above is estimated using the cox method. First, the survival function is calculated for the number of instructions at risk at a time, as in Equation 2. Then, after a learning process, a baseline function is generated. Figure 4 compares the learned baseline function of the model (shown by the dotted line) with what happens when the covariate is varied above the indicated value (coloured line). The results compare the indicative survival function observed when the covariate is varied. The baseline function is equal to the predictive function at the median of the training dataset.

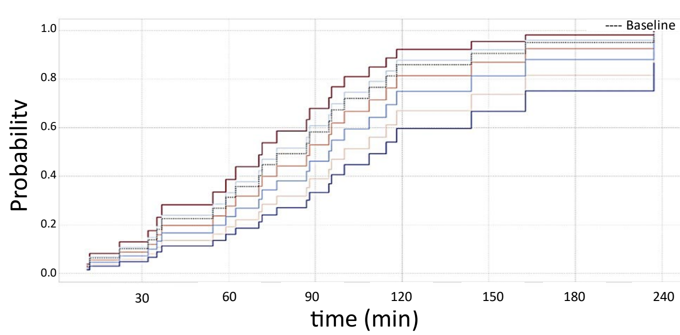

Figure 4. Baseline function (dotted line) and partial functions (solid line)

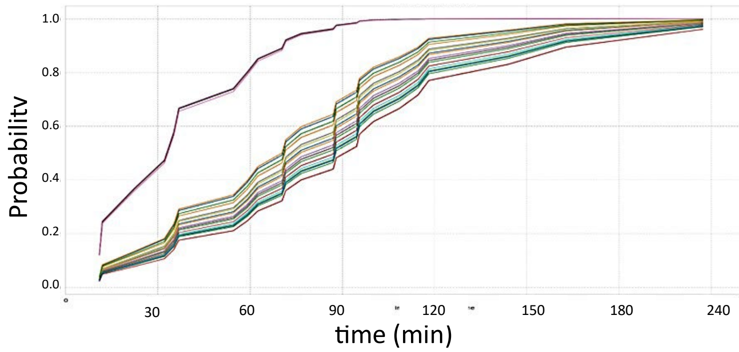

Figure 5. Prediction results of the survival time of each instruction in the test set

After obtaining the baseline function, forecasting can proceed by calculating the partial hazard or log partial hazard, which indicates the risk of irregular events. Figure 5 shows the predictive survival functions for future production plans. Each line represents the survival rate of each instruction. In this meaning, an instruction indicated by the purple line has more probability of experiencing failure much faster than the other transactions, which will have a 50% chance that irregular events will occur in only 30 minutes after the instruction is being executed in a real production situation. Through this analysis, we can obtain the probability of irregular events that will happen in each instruction.

5.2 Model performance

This section presents the experimental results of the proposed idea for one month of synthetic production data. The prediction results are evaluated using two evaluation indicators. First, the model's performance from a technical point of view is assessed by evaluation indicator 1. In the test cases, irregular events are predicted by the above-described methods and compared to the conventional method of monte Carlo. The main results are the total lateness and the objective function value that gives the total delayed instructions. In this case, the executed instruction time is compared with the predicted instruction time. From Figure 6, it can be seen that the average error in the prediction results of the proposed method is 10%. A smaller value for the error rate indicates better performance. The closer the error is zero, the better the proposed method correctly predicts the instruction duration. Finally, a comparison of our method with the original production plan shows that the error rate is 24.5%.

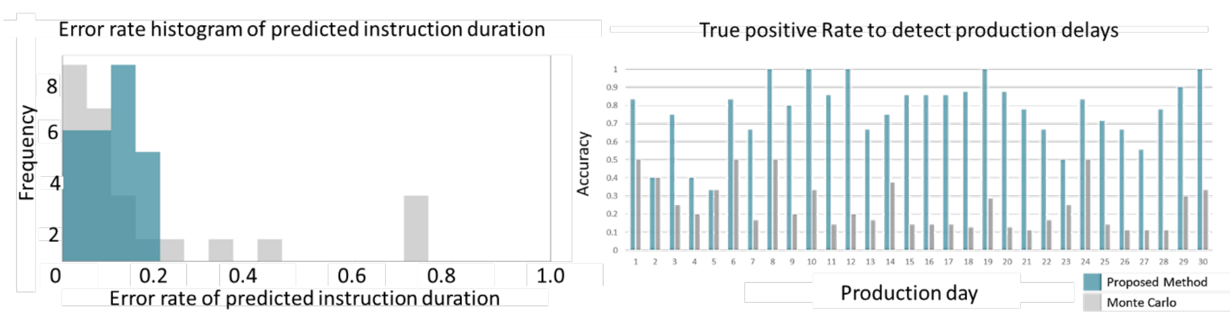

Figure 6. Model performance and result of the benchmark with the Monte Carlo method

We have also measured the daily production delay forecasting performance and found that the proposed method matches the actual delay behaviour more accurately than the Monte Carlo-based simulation. As a result, the proposed method achieved 76.6% accuracy in production delay forecasting. To compare the performance of our method, we run experiment to evaluate the simulations based on the proposed method and the Monte Carlo method with actual production delays, as shown in Figure 7. The results also show higher accuracy than the previous ones on all evaluation dates. In the early stages of the training data, the proposed method can fit the actual delays almost perfectly and shows a slight deviation from the Monte Carlo prediction. The proposed method was able to perform better than the Monte Carlo prediction. The main advantage of the proposed method is that it can be integrated into existing planning workflows without significant modifications to the model. The main prerequisite for applying the proposed robust planning workflow is modelling the simulation, which can be built quickly if a standard rule can be applied.

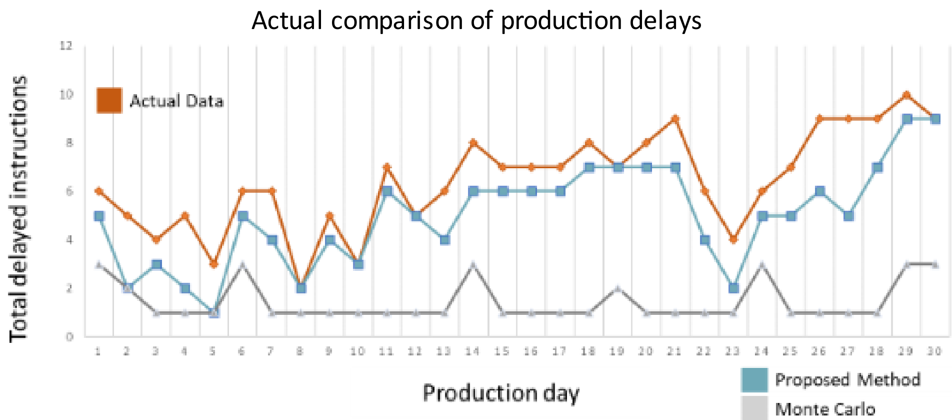

Figure 7. Evaluation metric on total delayed instruction

5.3 Proposed Improvements

For future work, the authors have identified two potential directions worth further research. The first direction is an analysis of the adequacy of the model to handle the latest trends. Survival functions are sensitive to training data, as they depend on patterns from the training data. On the other hand, in high-mix low-volume industries, where new products may be produced quickly, the actual production process will frequently change the probability of irregular events due to the dependency on production parameters. In such cases, the model needs to be brought into contact with the most recent dataset to maintain its ability to predict correctly. Therefore, re-training is vital for the long-term use of this model. The second direction is the analysis of learning requirements, such as appropriate periods that indicate more suitable patterns in the temporal data. The survival function is sensitive to the training data as it depends on patterns from the training data. Therefore, when the model is applied to stream data, the survival function needs to be adjusted to maintain the stability of the model accuracy. To determine the survival function, the choice of data learning is essential. Therefore, a mechanism for selecting appropriate training data is needed.

6. Conclusion and Recommendation

The objective of this paper is to promote strategies for robust production planning in high-mix, low-volume production industries through data-based situational awareness of production variability. Robust production planning requires the prediction of irregular events and estimating the disturbances that can absorb their impact on the production process. This paper proposes a new method to predict production delays due to irregular events. The proposed idea consists of two main stages: predicting irregular events and calculating their impact on the original production plan. A simulation-based optimisation method using survival analysis is deployed to assess the predicted irregular events without increasing the complexity and execution time of the production plan. Through the various parameters tested during the experiments to predict production delays, several conclusions can be drawn about the use of the proposed method. First, the modelling of production line productivity is highly dependent on the type of statistical distribution assigned to each irregular event variable. When using conventional method, obtaining such dataset is very difficult in the High Mix Low Volume manufacturing industry due to data limitation. Therefore, we propose to model the distribution using time to event (TTE) distributions which is effective for modelling irregular events even when there are no data points in the historical data of the target product. This is because survival time analysis assumes that if products have similar characteristics, their impact functions will also be similar. Therefore, even if there are zero data items available from historical data, the Time to Event data distribution can be estimated from similar product data. Thus, the survival time analysis method can overcome the problem of data volume limitations. As a result, the proposed method is able to predict irregular events better than conventional methods, showing reasonable prediction accuracy. Experimental results also show that the proposed method is effective in reducing production delays caused by irregular events and reducing losses associated with production delays. Although the model can be further refined and evaluated when applied to other manufacturing industries, the model developed can be used by production planners to predict production delays due to irregular events, and the method can be applied to other high-mix, low-volume industries.

References

Cachada, Ana, et al. Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture. *International Conference on Emerging Technologies and Factory Automation (ETFA).* Torino, Italy: International Conference on Emerging Technologies and Factory Automation (ETFA), 2018.

Campagne, Jean Pierre, Jacques Henri Jacot, Yannicl Frein, and Gerald Vitry. "A FRAMEWORK TO SPECIFY A REACTIVE AND." *INRIA/IEEE Symposium on Emerging Technologies and Factory Automation. ETFA'95.* Paris, France: IEEE, 1995.

Grzegorz, Kłosowski, Gola Arkadiusz, and Świć Antoni. "Application of Fuzzy Logic in Assigning Workers to Production Tasks." *Advances in Intelligent Systems and Computing* 474 2016.

Gyulai, David, Botond Kadar, and Laszlo Monosotori. "Robust production planning and capacity control for flexible assembly lines." *15th IFAC Symposium onInformation Control Problems inManufacturing.* Ottawa, 2015.

Hinai, NasrAl, and Tarek ElMekkawy. "Robust and stable flexible job shop scheduling with random machine breakdowns using a hybrid genetic algorithm." *International Journal of Production Economics*, 2011.

Ilar, Torbjörn, and John Powell. "Simulation of production lines–The importance of breakdown statistics and the effect of machine position." *International Journal of Simulation Modelling*, 2008.

Jang, Hong, Tae Yeong Jung, Kevin Yeh, and Jay Lee. "A model predictive control approach for fab-wide schedulling in semiconductor manufacturing system." *Management and Control of Production and Logistics.* Brazil: The International Federation of Automatic Control, 2013.

Javaid, Mohd, Abid Haleem, Ravi Pratap Singh, and Rajiv Suman. "Enabling flexible manufacturing system (FMS) through the applications of." *Internet of Things and Cyber-Physical Systems*, 2022.

Liu, Lin, Han yu Gu, and Yu geng Xi. "Robust and stable scheduling of a single machine with random machine breakdowns." *The International Journal of Advanced Manufacturing Technology* 31, no. 7 2007.

Łukasz Sobaszek, Arkadiusz Gola, Edward Kozłowski. "Application of survival function in robust scheduling of production jobs." *Conference: 2017 Federated Conference on Computer Science and Information Systems.* 2017.

Negahban, Ashkan, and Jeffrey S Smith. "Simulation for manufacturing system design and operation: Literature review and analysis." *Journal of Manufacturing Systems* 33, no. 2 2014.

Nouiri, Maroua, Abdelghani Bekrar, Abderrazzak Jemai, Damien Trentesaux, and Ahmed Chiheb Ammari. "Two stage particle swarm optimization to solve the flexible job shop predictive schedulling probblem considering possible machine breakdowns." *Computers & Industrial Engineering* 112 2017.

Qiao, Guixiu, Roberto Lu, and Charles McLean. "Flexible Manufacturing System for Mass Customization." *International Journal of Mass Customisation*, 2006.

Reinhart, Gunther, and Florian Geiger. "Adaptive scheduling by means of product-specific emergence data." *IEEE International Conference on Industrial Engineering and Engineering Management.* Singapore: IEEE, 2011.

Wang, Dujuan, Feng Liu, Yan Zhang Wang, and Yaochu Jin. "A knowledge-based evolutionary proactive scheduling approach in the presence of machine breakdown and deterioration effect." *Knowledge-Based System*, 2015.

Biographies

Irene Erlyn Wina Rachmawan is a researcher in Intelligent Innovation Research of Center for Technology innovation Department in the Central Research Laboratory at Hitachi Ltd. She received a doctoral degree in cyber infomatic from Keio University. Her current field of research is AI for global supply chain. She is interested in AI for industry, AI for remote sensing and AI for social innovation.

Issei Suemitsu is a senior researcher in Intelligent Innovation Research of Center for Technology innovation Department in the Central Research Laboratory at Hitachi Ltd. He received a master defree from Kyoto University and currently a PhD candidate in Kyoto University. His research focus is Global Supply Chain and Optimization technology.

Matsuba Hiroya is a chief researcher in Intelligent Innovation Research of Center for Technology innovation Department in the Central Research Laboratory at Hitachi Ltd. He received a doctoral degree in Information Science. His research field is High Performance Computing.