Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

ENABLING KNOWLEDGE DISCOVERY FROM SIMULATION-BASED MULTI-OBJECTIVE OPTIMIZATION IN RECONFIGURABLE MANUFACTURING SYSTEMS

Carlos Alberto Barrera Diaz
Henrik Smedberg
Sunith Bandaru
Amos H.C. Ng

Intelligent Production Systems Division
University of Skövde
Box 408 Högskolevägen
Skövde, 54128, SWEDEN

ABSTRACT

Due to the nature of today's manufacturing industry, where enterprises are subjected to frequent changes and volatile markets, reconfigurable manufacturing systems (RMS) are crucial when addressing ramp-up and ramp-down scenarios derived from, among other challenges, increasingly shortened product lifecycles. Applying simulation-based optimization techniques to their designs under different production volume scenarios has become valuable when RMS becomes more complex. Apart from proposing the optimal solutions subject to various production volume changes, decision-makers can extract propositional knowledge to better understand the RMS design and support their decision-making through a knowledge discovery method by combining simulation-based optimization and data mining techniques. In particular, this study applies a novel flexible pattern mining algorithm to conduct post-optimality analysis on multi-dimensional, multi-objective optimization datasets from an industrial-inspired application to discover the rules regarding how the tasks are assigned to the workstations constitute reasonable solutions for scalable RMS.

1 INTRODUCTION

In the current competitive market, manufacturing companies are frequently challenged by demand variations, and therefore they often need to address fluctuating production volumes. Consequently, the efficiency of a manufacturing system in reacting and adjusting its capacities and functionalities to cope with the volumes and demands changes constitutes a critical challenge for production organizations (Dou et al. 2021; Koren et al. 2017; Koren and Shpitalni 2010). To tackle, among other challenges, those caused by the demand and volume changes, Koren et al. 1999 first introduced the concept of reconfigurable manufacturing systems (RMS). RMS are production systems capable of adding/removing resources and modifying their capabilities to efficiently cope with expected or unexpected market shifts (Diaz et al. 2020; Koren et al. 2018). In such a manner, RMS are responsive manufacturing systems that, cost-effectively through reconfigurations such as the arrangement of machines or the process plan, can provide the required functionalities for several demand periods (Diaz et al. 2021). In a nutshell, RMS are essential to the current manufacturing industry to achieve high flexibility, fluctuating production volumes, flexible batches, and the required short life cycles for today's competitive market (Bortolini et al. 2018). Studies suggest that this type of system provides better performance in terms of scalability, productivity, responsiveness, and cost when compared to traditional production systems (Freiheit et al. 2003; Gu 2017).

978-1-6654-7661-4/22/$31.00 ©2022 IEEE

For decades, simulation techniques have successfully been applied within the manufacturing industry. Simulation has successfully become a powerful tool for designing, analyzing, and optimizing manufacturing systems (Mourtzis et al. 2014; Pehrsson et al. 2015). The complexity found in the dynamics scenarios of manufacturing systems can be assessed, modeled, and supported using simulation technology. When a simulation model is constructed, stakeholders and decision-makers better understand the systems (Mourtzis, 2020). Simulation tools are used to understand the behavior and performance of production systems under different scenarios defined by a set of input variables. However, as the complexity of manufacturing systems grows, the amount of input variables and feasible combinations increases exponentially. Against this backdrop, simulation-based optimization allows engineers and decision-makers to search in ample decision space for an optimal or near-optimal combination of input variables (Niño-Pérez et al. 2018; Xu et al. 2016). Simulation-based multi-objective optimization (SMO) is used when several conflicting objectives are pursued. However, the use of SMO reveals a large amount of data, including the impact of the input variables on the studied scenarios. Most studies usually focus on finding the optimal solution for a specific case. Still, as the size of the manufacturing system and the number of scenarios that need to be considered increases, the number of input variables combination grows exponentially. A relatively recent research area within SMO is knowledge discovery, wherein data mining methods are applied to the optimization dataset to reveal underlying information about what constitutes a satisfactory solution according to different criteria from the generated Pareto-optimal front. These data mining methods support the extraction of patterns that can provide the decision-makers with a better understanding of solving the problem under different scenarios (e.g., production volumes) (Bandaru et al. 2017).

Regarding the design of manufacturing systems, knowledge discovery methods have been successfully applied to extract patterns and rules between the capabilities of the system and product features. In (Kou and Xi 2018), it has been shown how association rules extracted from historical datasets can significantly impact production development effectiveness. Still, as commented recently in (Tripathi et al. 2021), data mining and knowledge discovery methods are challenging areas for future research and the evolution of manufacturing systems. RMS have a significant impact and act as an enabler on today's so-needed changeable and reconfigurable intelligent manufacturing systems (ElMaraghy et al. 2021). Due to the intrinsic complexity of RMS, an increasing number of researchers underscore the importance of applying SMO to tackle the design problems of RMS (Barrera Diaz et al. 2021; Bensmaine et al. 2013; Bortolini et al. 2018; Diaz et al. 2020; Renzi et al. 2014). As today's product lifecycles are becoming shorter and shorter, manufacturing systems evolve to become not only more flexible to be re-configurated more frequently but also more complex. Therefore, optimization problems generate an ever-increasing amount of data, making knowledge capturing and decision-making a more complicated task (Algeddawy and Elmaraghy 2011; Diaz et al. 2021). Accordingly, due to the nature of today's manufacturing industry being subjected to frequent changes and volatile markets, data mining and knowledge discovery methods have become even more crucial for RMS applications (ElMaraghy et al. 2021; Koren et al. 2018). Knowledge discovery methods are necessary for supporting engineers and decision-makers for two main reasons. Firstly, because many decision variables need to be considered when optimizing RMS, setting up a new optimization scenario can be considerably time-consuming. In this regard, gathering knowledge from previous optimizations can support understanding new scenarios without running further optimizations. Secondly, the analysis of SMO results can be simplified by the rules extracted from the optimization. These rules can be used to reveal how different decision variables affect the overall performance of the system.

One of the main challenging areas of RMS is the process planning and how to reconfigure the tasks to the workstations (WSs) under different scenarios. To this extent, this study uses the multi-objective optimization datasets from an industrial-inspired application to apply data mining methods and discover how the tasks assigned to WSs constitute reasonable solutions for a scalable RMS that need to be analyzed under different production volumes scenarios. This knowledge will be extracted and represented in rules from the best trade-off solutions between the optimization objectives, namely throughput (THP) and Leadtime (LT). Consequently, this study aims at proposing a methodology to conduct post-optimality analysis on multi-dimensional, multi-objective optimization datasets by applying a novel flexible pattern mining algorithm in an RMS application. The remaining of the paper is structured as follows. Section 2 describes the understanding of the main RMS design challenges, SMO, and Knowledge Discovery. Section 3 presents the methods of this study and the studied case. Section 4 presents the findings and discusses the insight that facilitated the proposed process. Conclusions and future work can be found in Section 5.

2 LITERATURE REVIEW

2.1 Reconfigurable Manufacturing Systems Challenges

RMS can be explained as the capability of a production system to change and reallocate its components effectively and efficiently to fulfill several new predictable or unpredictable restrictions/conditions of a system as many times as required (Goyal et al. 2012). Researchers support the idea that manufacturing organizations should be equipped with RMS that can rapidly be reconfigured to cope with changing markets and customers' needs to respond to a volatile market (Koren and Shpitalni 2010; Rösiö and Säfsten 2013). However, the RMS design and management are underdeveloped and involve crucial challenges in the research community (Andersen et al. 2018). These aspects include three main areas named the system configuration, the components of the system, and the process planning (Koren et al. 2018). The system configuration tackles the physical arrangement of machines and affects the overall performance and the functionality and scalability aspects of the system (Diaz et al. 2020; Koren et al. 1998). Most of the research focused on this area deals with the machine assignment to WSs. The components of the system refer to the required type and number of resources, e.g., machines and buffers, to achieve the desired production capacity and are considered a critical aspect for capacity planning and scalability (Koren et al. 2018). The majority of the research in this area focuses on optimizing the number of machines in the system. Lastly, the process planning area refers to how tasks are balanced throughout the RMS and assigned to the WSs, having a significant impact on the reconfigurations of the system to handle fluctuating production volumes (Azab et al. 2007; ElMaraghy 2007). Previous research in this area mainly focuses on optimizing the task assignment to WSs. Consequently, for an RMS to cope with fluctuating production volumes, it needs to address the previously explained areas and change its configuration by re-assigning, adding, or removing resources and rebalancing the tasks in the system (Koren et al. 2017; Wang and Koren 2012). These challenges constitute complex NP-hard problems that can be aided by simulation and optimization techniques (Bortolini et al. 2018; Delorme et al. 2016; Diaz et al. 2021; Michalos et al. 2012).

2.2 Simulation-Based Multi-Objective Optimization and Knowledge Discovery

2.2.1 Simulation-Based Multi-Objective Optimization and Multi-Criteria Decision Making

Multi-Objective Optimization (MOO) is a widely known research area focused on optimizing several conflicting objectives. Scalarization and *posteriori* are some of the most used MOO methods. Scalarization methods include ε-constraint and weighted sum method, among others. *Posteriori* methods generate a set of trade-off or non-dominated solutions which form the Pareto-optimal front, representing that optimizing one of the objectives degrades another. The rest of the solutions found in the objective space are known as dominated solutions as they are inferior to the non-dominated solutions in all the considered objectives. Thus, decision-makers often need support to simplify the selection of the best choice among all the available alternatives. Multi-Objective Evolutionary Algorithms (MOEAs) are among the most commonly used algorithms for generating the Pareto-optimal front (Deb 2014; Touzout and Benyoucef 2018). Accordingly, SMO can be seen as a combination of simulation and optimization. The intersection of these two powerful techniques has shown advantages compared to analytical optimizations or applied separately (Barrera Diaz et al. 2021; Jian and Henderson 2015). From a simulation perspective, SMO considers the variability and randomness found in RMS, avoiding simplifying the problem, which might result in inaccurate solutions. From an optimization perspective, SMO allows for solving more complex or impractical issues that would not be attainable by only simulation techniques (Carson and Maria 1997). The general representation of an SMO problem consists of several conflicting objectives defined by the objective function, possibly subjected to several equality and inequality constraints.

In practical scenarios, a single solution has to be chosen among all the near Pareto-optimal solutions generated by MOEAs. This task is not trivial since it often involves considering qualitative aspects of candidate solutions as perceived by one or more domain experts (decision-makers). Several methods have been developed within Multi-Criteria Decision Making (MCDM) to handle this task. These can broadly be classified into *a priori*, *posteriori* and interactive methods (Agrawal and Srikant 1995). *A priori* methods assume that the decision maker's preference is available before the optimization run, which is incorporated in the optimization algorithm to focus the searching process. While this drastically reduces the computational effort needed to find the most preferred solution, the decision-makers cannot always be expected to know what they want at the outset, especially in optimization scenarios that are unfamiliar to them. On the other hand, *posteriori* methods generate and present a representative Pareto-optimal front to the decision-makers. Its advantage over a priori methods is that the decision-makers can get a complete picture of all the available trade-off solutions in the objective space. Interactive methods aim to balance finding multiple trade-off solutions and the computational cost needed to find the complete Pareto front by interacting with the decision-maker who can provide his/her preference to guide the search and then narrow down the number of solutions to be considered during the optimization process. This implicitly entails an iterative process as the decision-maker is allowed to change and update the preferences during the optimization. While such approaches are more practical than *a priori* methods, they also increase the cognitive load on the decision-makers, which is a vital aspect to consider for practical decision-making activities.

2.2.2 Knowledge Discovery and Flexible Pattern Mining

Much of the literature on multi-criteria decision support is focused on assisting the decision-maker in visualizing and navigating the objective space. In practical optimization scenarios, it can be argued that a more informed decision requires not only finding a solution that conforms to the decision-maker's preferences but also knowledge about how those preferences affect the decision variables. Specifically, the decision-maker may be interested in knowing how different variable values change with preferences and what are the most important variables within a given region of preferences. Such knowledge can be extracted from solutions of an MOEA using various data mining and machine learning techniques (Bandaru et al. 2017). While some traditional data mining methods can be used directly with MOO data, they often need to be customized to consider that MOO data consists of two distinct spaces, the objective space and the decision space. The decision-maker provides preference in the objective space, while the knowledge about the variables exists in the decision space. Moreover, knowledge discovery in MOO also requires the methods to be interactive so that the decision-makers can realize the impact of changing their preferences.

Flexible Pattern Mining (FPM) is a recent interactive data mining method designed with the goal of knowledge discovery in MOO and multi-criteria decision support. FPM uses the popular Apriori algorithm (Agrawal and Srikant 1995) to discover *discriminative decision rules* in the MOO data that distinguish between a chosen *selected set* (typically preferred solutions) and an *unselected set* (typically all other solutions). The Apriori algorithm was developed for extracting frequent itemsets (items that are often bought together) from market basket data. Hence, it treats all variables as categorical and extracts knowledge in the form of patterns. The main difference in FPM is the way the MOO data is processed to convert continuous and discrete variables into categorical features that can then be handled by the Apriori algorithm. Thus, FPM is able to extract complex decision rules formed by singular rules of the form $x_{\mathrm{t}} > c$, $x_{\mathrm{t}} < c$ or $x_{\mathrm{t}} = c$, where x_{t} can be any of the variables in the MOO data and c is a value for that variable from the data. Each decision rule is associated with a selected significance (sig) and an unselected significance ($unsig$), indicating the percentage of solutions that follow the decision rule in the selected and unselected sets. Thus, a highly discriminative rule should have a very high sig value, and a meager $unsig$

value. An example of such a discriminative rule is $[x_1 > 4.2 \wedge x_5 < 10.5 \wedge x_6 = 2]$ with $sig = 90\%$ and $unsig = 5\%$. It indicates that 90% of the solutions in the selected set have x_1, x_5 and x_6 values as specified by the rule, while this is true for only 5% of solutions in the unselected set. If the selected set consists of preferred solutions, this decision rule informs the decision-maker that variables x_1, x_5, and x_6 are critical to determining whether a solution is desirable or undesirable.

3 METHOD AND PROBLEM FORMULATION

3.1 An Interactive and Iterative Methodology

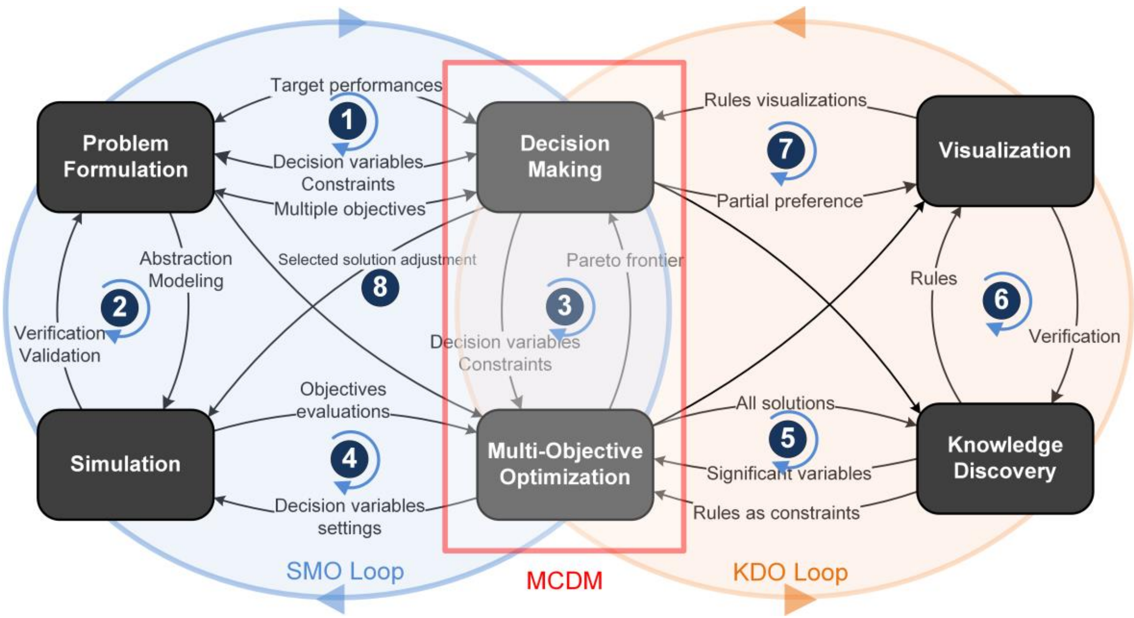

Figure 1: SMO-KDO Methodology.

The unique concept proposed in this research is Knowledge-Driven Optimization (KDO) – instead of merely capturing knowledge from experiences (e.g., in the form of rules of thumb) or experiments by using physical equipment, extracting knowledge for decision support is achieved through systematically exploring and analyzing (e.g., using data mining and visualization techniques) multiple optimal solutions (designs/configurations/settings) generated via MOO on simulation models. The conceptual framework of the double-loop SMO-KDO methodology is schematically described in Figure 1. Within this methodological framework, there are six processes: (1) Problem Formulation; (2) Simulation; (3) MOO; (4) Knowledge Discovery; (5) Visualization, and (6) Decision-Making. The numbered labels in Figure 1 are purposed to identify the nine interactive and iterative loops among the processes, but their number order describes a typical workflow of such a methodology. By interactive, it means that there are always some users, like a decision-maker (or a group of decision-makers) with a team of simulation/production engineers, who interact with each other and control these processes. On the other hand, iterative means that any process and any loop can be re-visited and run multiple times until the user is satisfied with the results that can be fed into the other processes or to support the final decision.

3.2 Problem Formulation

The datasets used to show the applicability of the SMO-KDO methodology come from an industrial application study in the Swedish automotive industry. The case is based on a 4-cylinder crankshaft production that manufactures two product families, part 1 (4cylP) and part 2 (4cyLD). The production line consists of 18 WSs, wherein 3 are reconfigurable WSs placed in the bottleneck of the line. Unlike in the rest of the line, these 3 consecutive reconfigurable WSs can add, remove and relocate machines to cope with production changes up to a maximum of 5 machines per WS. The considered multi-part flow line (MPFL) needs to produce two parts at specific volumes. As the demand fluctuates, the system configuration, components of the system, and process planning of the reconfigurable WSs change to meet the new production scenario. These changes affect not only the layout and the total number of machines needed in the aforementioned WSs, but also the assignment of tasks to them. Due to the company's interest in specific scenarios, the SMO was applied to a total of 12 scenarios considering different production proportions and the number of machines in the reconfigurable WSs. The scenarios consisted of modifying the production proportions of part 1 and part 2 from 80/20 (80 % part 1 and 20 % part 2), the opposite case 20/80, and considering two more cases in between, i.e., 60/40 and 40/60, respectively. In addition, each of these proportions needed to be studied for 7, 8, and 9 machines in the reconfigurable WSs, making 12 different scenarios. Figure 2 shows the precedence of the tasks for both parts in the reconfigurable WSs.

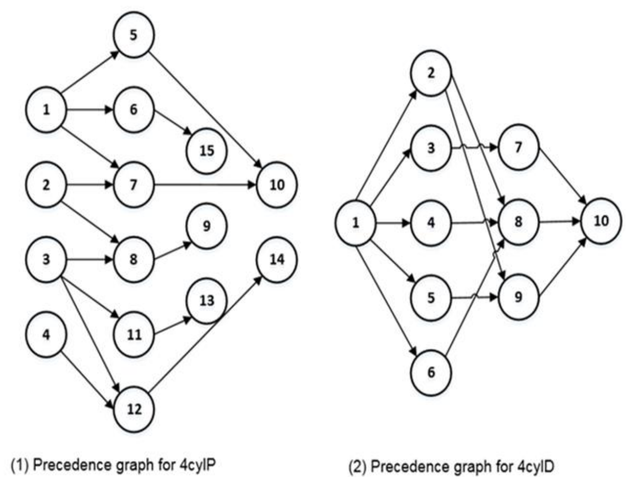

Figure 2: Precedency graphs.

Each scenario was optimized using NSGA II and 15000 iterations. An SMO software called Facts Analyzer (Ng et al. 2011), which integrates a discrete-event simulation (DES) engine and several optimization algorithms, was used to model the studied manufacturing line and run the optimizations. Two conflicting optimization objectives were used:

$$Maximize\ f1 = TH(x){:}\ Throughput\ (jobs\ per\ hour)\ , \tag{1}$$

$$Minimize\ f2 = \ LT(x)\ :\ Lead\ Time. \tag{2}$$

Accordingly, in this study, we have applied FPM method to the decision space of the generated datasets to find patterns and rules in the tasks assignment that could support decision-makers in understanding how they impact the studied scenarios and how this could help setting-up future optimizations scenarios. To extract the knowledge about the influence and interaction of the tasks under different scenarios, we compare the Pareto-optimal solutions to the rest. This comparison was applied to eight different sets of Pareto-optimal solutions according to various criteria such as the number of machines or production proportion, as shown in the results sections.

4 RESULTS

This section explains the knowledge discovery process, followed by the presentation and discussion of the results. The formulation of the simulation model resulted in many duplicated solutions; therefore, to not bias the resulting knowledge, it is first necessary to process the data and remove all duplicate solutions. Further, due to the problem formulation of task allocation as Boolean values, the task allocation variables were combined into one integer value per task, representing the station the task was assigned to. The values 1, 2, or 3 represent if the study was performed in the first, second, or third reconfigurable WS on the line. Finally, non-dominated sorting is applied to find the Pareto-optimal solutions. These three data processing steps were performed for each of the twelve scenarios before combining them into a single dataset. This study's data processing and analysis were performed using the openly available decision support tool KDO-Mimer[1], which facilitates knowledge discovery in MOO.

The combined results from the optimizations can be seen in Figure 3. The left-hand side of the Figure shows all twelve datasets combined in a 2D scatter plot where the axes represent the objectives of the optimizations. The dark red, green, and blue represent the non-dominated optimal solutions for 7, 8, and 9 machines, respectively. The 3D scatter plot presented on the right-hand side shows a better visualization of the twelve scenarios where the axes represent the optimization objectives (THP), the total numbers of machines used in the reconfigurable WSs, and the different proportions studied.

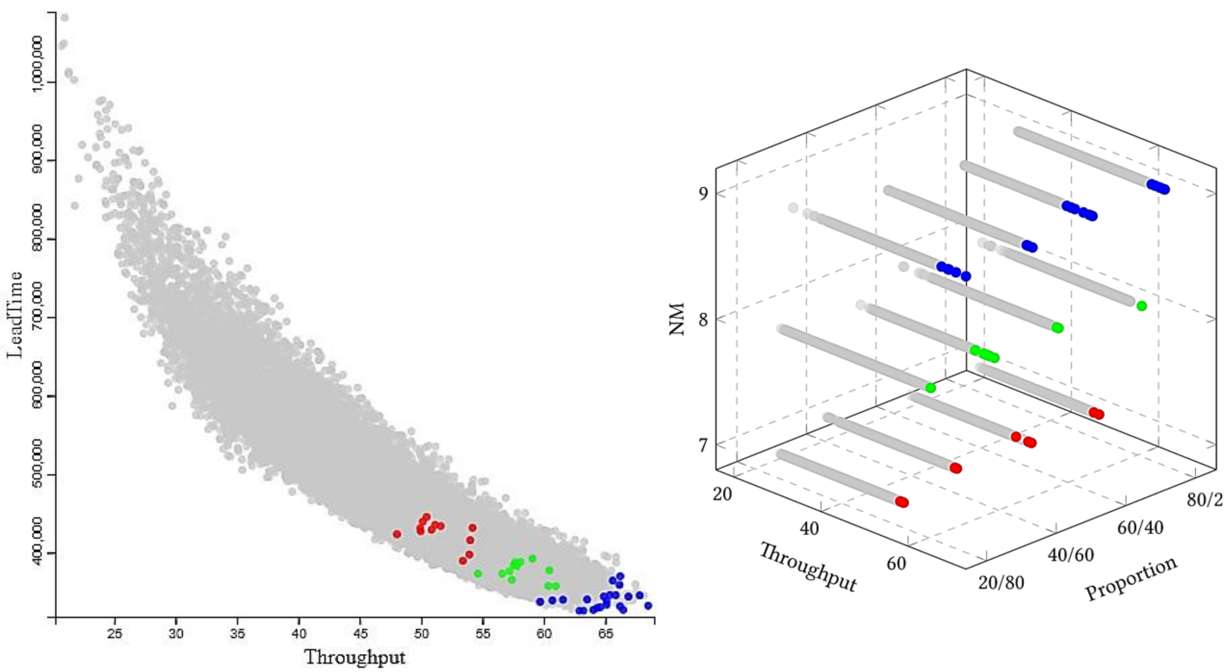

Figure 3: Combined datasets and the studied scenarios.

Figure 4 illustrates how the Pareto-optimal solutions P are grouped into eight subsets to be compared with the dominated solutions D according to different proportions and numbers of machines. The FPM procedure requires a selected and an unselected set of solutions. We combine all other cases in the first considered case and use the Pareto-optimal solutions from all scenarios as the selected set and all remaining dominated solutions as the unselected set. Next, we consider the three cases for the different number of machines. We use the Pareto-optimal solutions from all scenarios where the number of machines used matches the case as the selected set and the remaining solutions as the unselected set. For the last four cases,

[1] https://assar.his.se/mimer/html

we consider the scenarios where the proportions are the same, in the same way. Figure 4 demonstrates the selected set (gray) and unselected set (white) for the eight different scenarios. By dividing the different scenarios this way, we hope to find specific rules related to the specific proportions and numbers of machines.

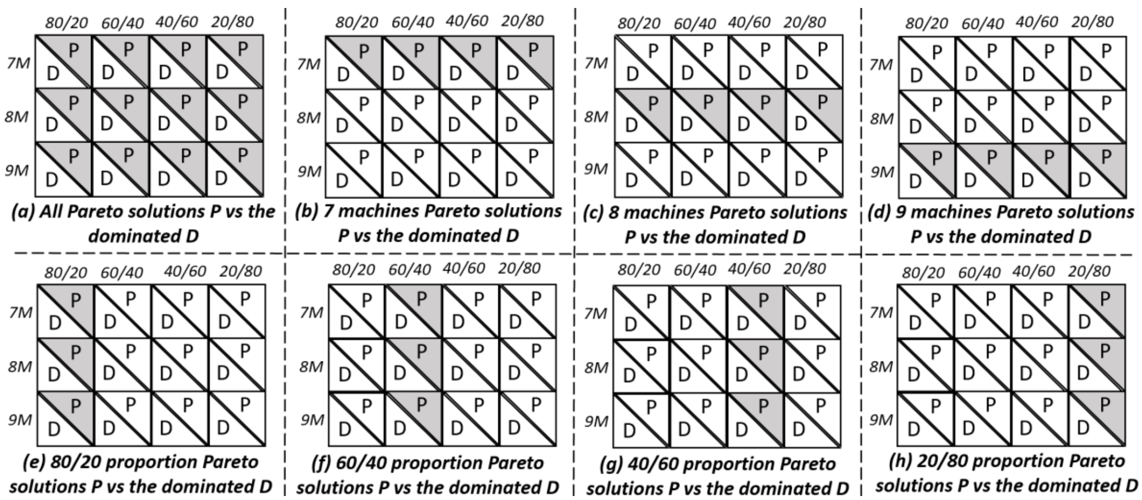

Figure 4: FPM selected and unselected set of solutions.

While the rules generated separately from these different cases can give insights into optimal task allocations for each case, it is interesting to note that the combined case can generate more general rules that apply invariantly to the system, no matter the number of machines or the proportion used. The three cases for the different number of machines will generate rules describing the optimal task allocation for a specific number of machines, regardless of the proportion. Finally, the four cases for the different proportions will generate rules relating to the optimal task allocation for a particular proportion, irrespective of the number of machines.

To generate the rules, we ran the FPM procedure with a minimum allowed significance of 90% to find only rules that accurately describe the selected set. Then the rules were filtered using the sliders to the right of the graph-interface to find the single rule-interaction with the highest ratio between the significance and unselected significance values. In this way, the rules accurately describe the difference between the selected and unselected sets. The resulting rules from each case can be found in Table 1. In this table, the "T" represents tasks from part 1 while "P" represents tasks from part 2.

Table 1: Resulting Rules.

Selection	Rules	Sig.	Unsig.
Fig. 4 (a)	$T5 \neq 3 \land T8 = 2 \land T12 \neq 3 \land T15 \neq 3 \land P8 \neq 1$	91.3%	38.5%
Fig. 4 (b)	$T9 \neq 3 \land T12 \neq 3 \land T15 \neq 3 \land P6 \neq 2 \land P8 = 3$	91.67%	22.29%
Fig. 4 (c)	$T5 \neq 3 \land T12 \neq 3 \land T15 \neq 3 \land P6 \neq 3 \land P7 \neq 2$	91.67%	21.4%
Fig. 4 (d)	$T3 = 1 \land T4 = 1 \land T9 \neq 3 \land T12 \neq 3 \land T15 \neq 3 \land P8 \neq 1$	90.91%	21.23%
Fig. 4 (e)	$T4 = 1 \land T5 \neq 3 \land T9 = 2 \land T15 \neq 3 \land P7 = 3$	90.0%	16.11%
Fig. 4 (f)	$T4 = 1 \land T5 \neq 3 \land T9 = 2 \land T12 \neq 3 \land T15 \neq 3 \land P7 = 3 \land P8 \neq 1$	90.0%	12.79%
Fig. 4 (g)	$T5 = 1 \land T8 = 2 \land T12 \neq 3 \land T15 \neq 3 \land P8 \neq 1$	92.31%	21.95%
Fig. 4 (h)	$T4 = 1 \land T6 = 1 \land T9 = 2 \land T12 \neq 3 \land P2 = 2 \land P3 \neq 3 \land P8 = 3$	90.91%	10.78%

We used the openly available decision support software KDO-Mimer to generate the results. A snapshot from KDO-Mimer showing the graph-interface for filtering FPM rules can be found in Figure 5. The upper part of Figure 5 shows the selection from Figure 4 (d), and the bottom part shows the selection from Figure 4 (f). The graph-interface in KDO-Mimer offers the decision-makers a holistic overview of the rules and allows them to more easily compare the results compared to solely presenting the rules in a table.

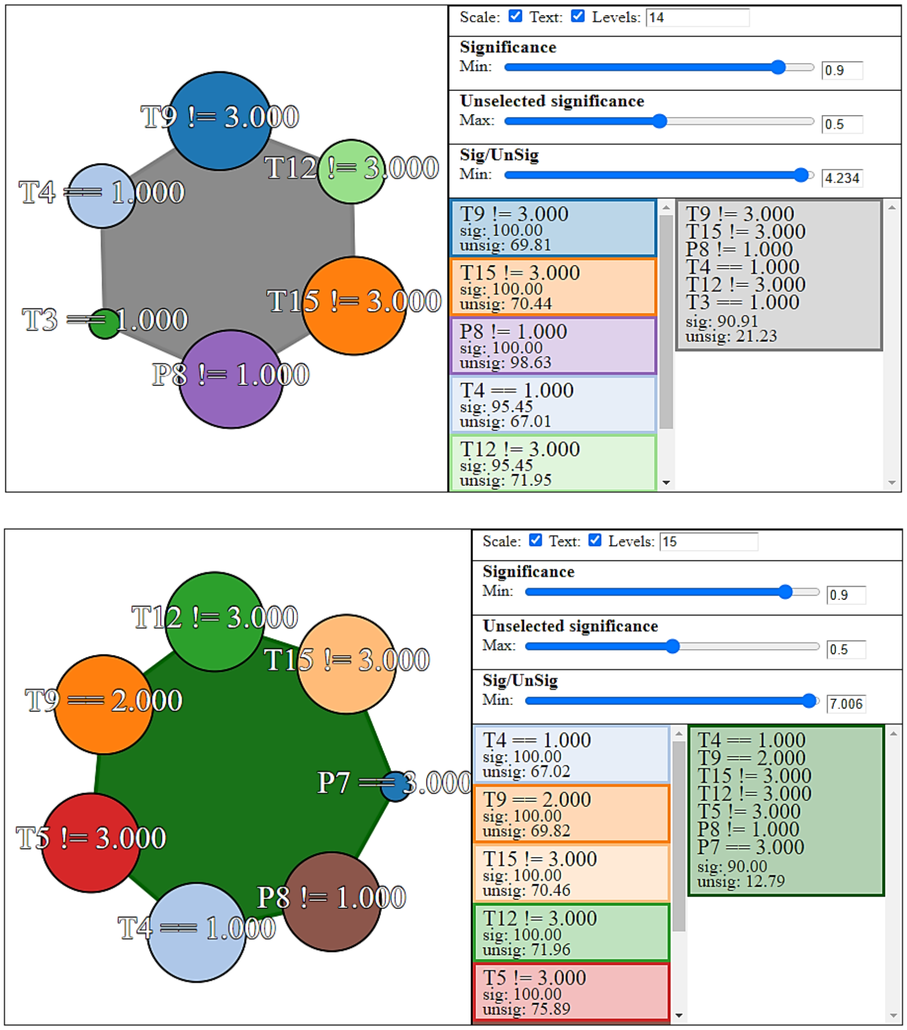

Figure 5: KDO-Mimer interface for filtering FPM rules.

4.1 Knowledge Discussion

In this study, we aimed to use data mining in twelve MOO datasets of an RMS to discover knowledge that could lead to a better understanding of the systems and therefore be applied in future scenarios. Looking at the general rules in Table 1, we can appreciate that all found rules have more conditions applied to part 1 than part 2. This fact can be explained due to part 1 having more tasks to be performed than part 2 or due to the nature of the precedence tree being more restrictive and leaving less freedom when assigning tasks to WSs. This implies that engineers could prioritize assigning the tasks of part 1 over part 2 when studying new scenarios. However, as expected, we can see that the higher number of part 2 tasks involved in a rule would be when producing 20% of part 1 and 80% of part 2. However, even for such an extreme scenario, we can see a higher involvement of tasks of part 1 than part 2.

One aspect not considered in the data mining but important to focus on when discussing the rules is the arrangement of the machines in the reconfigurable WSs. The machines configuration was another aspect included with task allocation as input variables in all optimization scenarios. The optimal configurations obtained in the reconfigurable WSs starting from 80/20 to the opposite proportion were: 1-2-4; 1-2-4; 2-1-4; 2-2-3 for seven machines; 1-2-5; 1-2-5; 2-2-4; 2-3-3 for eight machines; and 2-2-5; 2-2-5; 2-2-5; 2-3-4 for nine machines. Having this information, we can observe the implication of the rules on the different configurations due to the dependency between the number of tasks assigned to each WSs and the number of machines in it. One example is the number of tasks assigned to the first reconfigurable WS (tasks equal to 1 in the rules). There are either 2, 1, or 0 tasks assigned to the first reconfigurable WS in all found rules. However, two machines were placed in the first reconfigurable WS in all the scenarios where nine machines were employed. Consequently, the greatest number of tasks possible found in the rules (two tasks) were assigned there.

Another interesting aspect extracted from this method is the importance of some tasks. It can be observed in Table 1 that most of the tasks included in the extracted rules are often found in many of the rules implying the relevance of these tasks for the overall system. In other words, decision-makers can use these rules to understand which tasks are more critical to the overall performance of the RMS.

5 CONCLUSIONS AND FUTURE WORK

Due to the nature of today's manufacturing industry, where enterprises are subjected to frequent changes and volatile markets, RMS are becoming crucial and more sophisticated. Consequently, the optimization of RMS generates an ever-increasing amount of output data, making knowledge capturing a challenging task for engineers and decision-makers. Thus, knowledge discovery and data mining techniques have become important for RMS applications. Therefore, this study used the MOO datasets from an industrial-inspired application to apply data mining methods and discover how the tasks assigned to WSs constitute reasonable solutions for a scalable RMS that need to be analyzed under different production volume scenarios. The use of SMO to optimize RMS explores the search space seeking feasible solutions to find the optimal system configuration avoiding a manual and time-consuming trial and error process. The presented method showed how data mining and FPM could be applied to the generated data sets and support a better understanding of the behaviors of the RMS and its variables under different scenarios providing the decision-makers with critical factors to improve and understand the system. Using this method in an MPFL identified which product to prioritize when deciding on the tasks allocation and which tasks are more critical to the overall performance of the system when optimizing THP and LT. The applicability of the presented method is not limited to RMS; it should nevertheless be beneficial to be applied to many other applications where MOO generates large data sets regarding changing scenarios. The extracted knowledge can be applied to future scenarios and save time and effort by reducing the search space of the optimization. Therefore, the authors propose using the gained knowledge in future related optimization scenarios as future work. Such a process of utilizing the extracted knowledge within SMO in future optimization is known as off-line KDO. By expanding this study and applying off-line KDO, it is interesting to further investigate how the performance of the MOO algorithms can be boosted in RMS applications.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Knowledge Foundation (KK-Stiftelsen), Sweden, for their support and provision of research funding through the research profile Virtual Factories Knowledge-Driven Optimization (VF-KDO) at the University of Skövde, Sweden, in which this work is a part of it.

REFERENCES

Agrawal, R., and R. Srikant. 1995. “Mining sequential patterns”. In *Proceedings of the International Conference on Data Engineering*, March 6th-10th, Taipei, Taiwan, 3–14.

Algeddawy, T., and H. Elmaraghy. 2011. “Manufacturing systems synthesis using knowledge discovery”. *CIRP Annals* 60(1):437–440.

Andersen, A.-L., H. ElMaraghy, W. ElMaraghy, T. D. Brunoe, and K. Nielsen. 2018. “A participatory systems design methodology for changeable manufacturing systems”. *International Journal of Production Research* 56(8):2769–2787.

Azab, A., G. Perusi, H. A. ElMaraghy, and J. Urbanic. 2007. “Semi-generative macro-process planning for reconfigurable manufacturing”. *In Proceedings of Digital Enterprise Technology*, edited by P.F. Cunha, P.G. Maropoulos, 251–258. Boston, MA: Springer.

Bandaru, S., A. H. C. Ng, and K. Deb. 2017. “Data mining methods for knowledge discovery in multi-objective optimization: Part A - Survey”. *Expert Systems with Applications* 70:139–159.

Barrera Diaz, C. A., M. Fathi, T. Aslam, and A. H. C. Ng. 2021. “Optimizing reconfigurable manufacturing systems: A Simulation-based Multi-objective Optimization approach”. *Procedia CIRP* 104:1837–1842.

Bensmaine, A., M. Dahane, and L. Benyoucef. 2013. “A Simulation-based Genetic Algorithm Approach for Process Plans Selection in Uncertain Reconfigurable Environment”. *IFAC Proceedings Volumes* 46(9):1961–1966.

Bortolini, M., F. G. Galizia, and C. Mora. 2018. “Reconfigurable manufacturing systems: Literature review and research trend”. *Journal of Manufacturing Systems*, Elsevier B.V. 49:93–106.

Carson, Y., and A. Maria. 1997. “Simulation optimization: methods and applications”. In *Proceedings of the 1997 Winter Simulation*, edited by S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, 118–126. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Deb, K. 2014. “Multi-objective optimization”. *Search Methodologies*, Springer, 403–449.

Delorme, X., S. Malyutin, and A. Dolgui. 2016. “A multi-objective approach for design of reconfigurable transfer lines”. *IFAC-PapersOnLine* 49(12):509–514.

Diaz, C. A. B., T. Aslam, and A. H. C. Ng. 2021. “Optimizing Reconfigurable Manufacturing Systems for Fluctuating Production Volumes: A Simulation-based Multi-Objective Approach”. *IEEE Access* 9:144195-144210.

Diaz, C. A. B., T. Aslam, A. H. C. Ng, E. Flores-Garcia, and M. Wiktorsson. 2020. “Simulation-Based Multi-Objective Optimization for Reconfigurable Manufacturing System Configurations Analysis”. In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 1527–1538. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Dou, J., J. Li, D. Xia, and X. Zhao. 2021. “A multi-objective particle swarm optimisation for integrated configuration design and scheduling in reconfigurable manufacturing system”. *International Journal of Production Research* 59(13):3975–3995.

ElMaraghy, H. A. 2007. “Reconfigurable process plans for responsive manufacturing systems”. In *Proceedings of the Digital Enterprise Technology*, edited by P.F. Cunha, P.G. Maropoulos, 35–44. Boston: Springer.

ElMaraghy, H., L. Monostori, G. Schuh, and W. ElMaraghy. 2021. “Evolution and future of manufacturing systems”. *CIRP Annals* 70(2):635–658.

Freiheit, T., M. Shpitalni, S. J. Hu, and Y. Koren. 2003. “Designing productive manufacturing systems without buffers”. *CIRP Annals* 52(1):105–108.

Goyal, K. K., P. K. Jain, and M. Jain. 2012. “Optimal configuration selection for reconfigurable manufacturing system using NSGA II and TOPSIS”. *International Journal of Production Research* 50(15):4175–4191.

Gu, X. 2017. “The impact of maintainability on the manufacturing system architecture”. *International Journal of Production Research* 55(15):4392–4410.

Jian, N., and S. G. Henderson. 2015. “An introduction to simulation optimization”. In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 1780–1794. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Koren, Y., X. Gu, and W. Guo. 2018. “Reconfigurable manufacturing systems: Principles, design, and future trends”. *Frontiers of Mechanical Engineering* 13(2):121–136.

Koren, Y., U. Heisel, F. Jovane, H. Van Brussel. 1999. “Reconfigurable manufacturing systems”. *CIRP Annals* 48(2):527–540.

Koren, Y., S. J. Hu, and T. W. Weber. 1998. “Impact of manufacturing system configuration on performance”. *CIRP Annals* 47(1):369–372.

Koren, Y., and M. Shpitalni. 2010. “Design of reconfigurable manufacturing systems”. *Journal of Manufacturing Systems* 29(4):130–141.

Koren, Y., W. Wang, and X. Gu. 2017. “Value creation through design for scalability of reconfigurable manufacturing systems”. *International Journal of Production Research* 55(5):1227–1242.

Kou, Z., and L. Xi. 2018. “Binary particle swarm optimization-based association rule mining for discovering relationships between machine capabilities and product features”. *Mathematical Problems in Engineering* 2018:1-16.

Michalos, G., S. Makris, and D. Mourtzis. 2012. “An intelligent search algorithm-based method to derive assembly line design alternatives". *International Journal of Computer Integrated Manufacturing* 25(3):211–229.

Mourtzis, D., M. Doukas, and D. Bernidaki. 2014. "Simulation in manufacturing: Review and challenges". *Procedia CIRP* 25:213–229.

Ng, A. H. C., J. Bernedixen, M. U. Moris, and M. Jägstam. 2011. "Factory flow design and analysis using internet-enabled simulation-based optimization and automatic model generation". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, 2181–2193. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Niño-Pérez, E., Y. M. Méndez-Vázquez, D. E. Arias-González, and M. Cabrera-Ríos. 2018. "A simulation–optimization strategy to deal simultaneously with tens of decision variables and multiple performance measures in manufacturing". *Journal of Simulation* 12(3):258–270.

Pehrsson, L., M. Frantzén, T. Aslam, and A. H. C. Ng. 2015. "Aggregated line modeling for simulation and optimization of manufacturing systems". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W K. V. Chan, I. Moon, T. M K. Roeder, C. Macal, and M D. Rossetti, 3632–3643. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Renzi, C., F. Leali, M. Cavazzuti, and A. O. Andrisano. 2014. "A review on artificial intelligence applications to the optimal design of dedicated and reconfigurable manufacturing systems". *The International Journal of Advanced Manufacturing Technology* 72(1):403–418.

Rösiö, C., and K. Säfsten. 2013. "Reconfigurable production system design–theoretical and practical challenges". *Journal of Manufacturing Technology Management* 24(7):998–1018.

Touzout, F. A., and L. Benyoucef. 2018. "Multi-objective sustainable process plan generation in a reconfigurable manufacturing environment: exact and adapted evolutionary approaches". *International Journal of Production Research* 57(8): 2531–2547.

Tripathi, S., D. Muhr, M. Brunner, H. Jodlbauer, M. Dehmer, and F. Emmert-Streib. 2021. "Ensuring the Robustness and Reliability of Data-Driven Knowledge Discovery Models in Production and Manufacturing". *Frontiers in Artificial Intelligence* 4:22.

Wang, W., and Y. Koren. 2012. "Scalability planning for reconfigurable manufacturing systems". *Journal of Manufacturing Systems* 31(2):83–91.

Xu, J., E. Huang, L. Hsieh, L. H. Lee, Q. S. Jia, and C. H. Chen. 2016. "Simulation optimization in the era of Industrial 4.0 and the Industrial Internet". *Journal of Simulation* 10(4):310–320.

AUTHOR BIOGRAPHIES

CARLOS ALBERTO BARRERA DIAZ received a B.Sc. degree in Electrical Engineering from the University of Malaga, Spain, a B.Sc. degree in Automation Engineering, and his Master's degree in Industrial Systems Engineering from the University of Skövde, Sweden, where he is currently pursuing his Ph.D. degree with the Intelligent Production Systems Division. His research interests include design, modeling, simulation, and multi-objective optimization to analyze manufacturing systems. His email address is carlos.alberto.barrera.diaz@his.se.

HENRIK SMEDBERG received his B.Sc. degree in Informatics and Master's degree in Automation Engineering from the University of Skövde, Sweden, where he is currently pursuing a Ph.D. within the Division of Intelligent Production Systems. His research interests include the application of knowledge discovery techniques both for decision support in multi-objective optimization and in algorithms for knowledge-driven optimization. His email address is henrik.smedberg@his.se.

SUNITH BANDARU is a Senior Member, IEEE, and an Associate Professor at the Division of Intelligent Production Systems, University of Skövde, Sweden. He received his Bachelor's degree (2006) in Mechanical Engineering from Jawaharlal Nehru Technological University, India, and Masters (2008) and Ph.D. (2013) degrees from IIT Kanpur, India. His research interests lie at the intersection of evolutionary computation, machine learning, and data mining. His email address is sunith.bandaru@his.se.

AMOS HC NG is a Professor of Automation Engineering at the Division of Intelligent Production Systems, University of Skövde, Sweden. He is also a Visiting Professor at the Division of Industrial Engineering and Management, Uppsala University, Sweden, and the CEO of Evoma AB. His research interests include the applications of simulation, multi-objective optimization, and artificial intelligence for the design, analysis, and improvement of production/health-care/business systems. His email address is amos.ng@his.se.