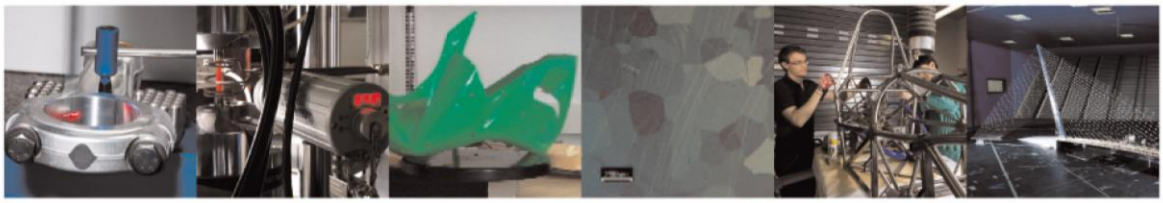

Analysis of the near flatness phenomenon for multi-loop closed manufacturing systems

Loffredo, Alberto; Zhang, Mengyi; Lin, Ziwei; Matta, Andrea

This is an Accepted Manuscript of an article published by Taylor & Francis in INTERNATIONAL JOURNAL OF COMPUTER INTEGRATED MANUFACTURING on 15 Mar 2023, available online: https://doi.org/10.1080/0951192X.2023.2189307.

Analysis of the Near Flatness Phenomenon for Multi-loop Closed Manufacturing Systems

Alberto Loffredo[a], Mengyi Zhang[a], Ziwei Lin[a,b] and Andrea Matta[a]

[a] Department of Mechanical Engineering, Politecnico di Milano, via La Masa 1, 20156 Milano, Italy
[b] Department of Industrial Engineering and Management, School of Mechanical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Rd, 200240 Minhang, Shanghai, China

ABSTRACT
With the recent development of the electrical vehicle (EV) industry, the study of manufacturing systems producing key components in this sector is becoming increasingly important. Multi-loop closed manufacturing systems (MCMS), whose operation and control are rarely studied in the literature, are widely used in the EV industry. This work provides innovative guidelines for MCMS operation also valid in a general context and not necessarily limited to the EV field. The main focus is on a specific topic of MCMS operation, namely the *near flatness* phenomenon. The near flatness indicates how system throughput is influenced by its population, i.e. the number of items circulating in the MCMS, especially in high-throughput conditions. The study of near flatness aims at enhancing MCMS flexibility in terms of population control and handling while guaranteeing high system throughput. In this work, a new indicator quantitatively describing the near flatness is provided. Numerical studies are conducted to analyze the effect of machine efficiency in isolation, mean times to repair, and buffer capacities on the near flatness. Experiments are also carried out on a real case of MCMS in the EV field. Based on the experimental results, practical specifications to improve MCMS design and management are provided.

KEYWORDS
Multi-loop Closed Manufacturing Systems; Electrical Vehicle Industry; Near Flatness; Throughput.

1. Introduction

1.1. *Motivation*

The rapid growth of the electric vehicle (EV) field has been a major trend in recent years for the automotive industry: the need to reduce fuel costs and to decrease the environmental impact of the car mobility sector makes the EV an environmental-friendly alternative to a conventional vehicle (Milas, Mourtzis, and Tatakis 2020). In addition, the expansion and assimilation of Industry 4.0 Key Enabling Technologies by companies is definitely facilitating the overcoming of challenges related to the performance of an EV (Athanasopoulou et al. 2022); examples of such challenges are increasing battery performance, reducing battery charging time, and so on. Therefore, the EV sector is expected to deal with a significant increase in production and sales of this

CONTACT Alberto Loffredo Email: alberto.loffredo@polimi.it

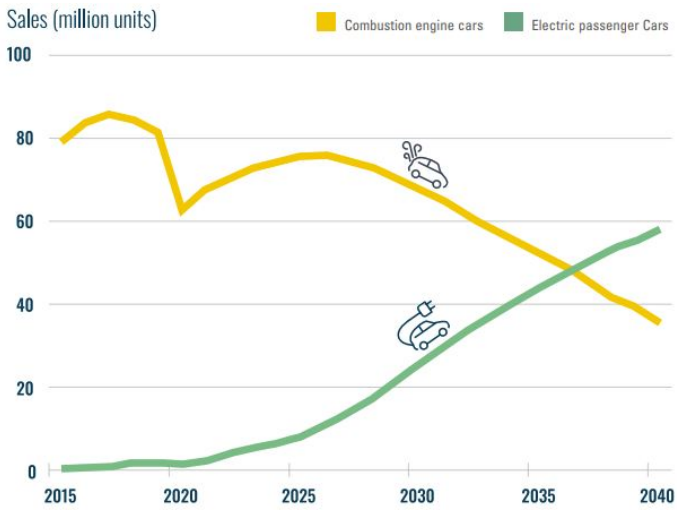

Figure 1. Expected sales of EVs and traditional combustion-engine vehicles in the next 20 years (Parajuly, Ternald, and Kuehr 2020).

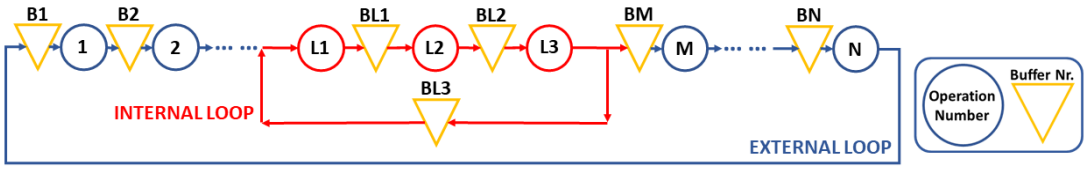

Figure 2. Example of MCMS composed of an external and an internal loop.

kind of vehicle in the upcoming years (see Figure 1 from Parajuly, Ternald, and Kuehr (2020)).

The motor of the electric drive system, composed of the rotor and the stator, plays a key role in the EV industry. Over the last years, the predominant configuration for the stator is the so-called *hairpin stator*. The latter is an alternative to the traditional wound-coils-based technology (Fleischer, Haag, and Hofmann 2017; Kim 2017), but with a lower size for equal power and a superior production rate (Ishigami, Tanaka, and Homma 2014, 2015). Each hairpin stator is characterized by the presence of more semicircular fixtures stacked on top of each other; all these elements are placed on the stator's upper part. In the manufacturing practice, each of this set of components is called a *layer*. The operations to assemble each layer with the stator core (i.e. layer insertion, crimping, and welding) must be repeated for a certain number of times. Since the aforementioned operations are identical for each layer, the repetition can be achieved with an internal closed loop, in which each pallet carrying one stator must circulate for a fixed number of times. Operations that are not related to the layer assembly are executed in an external loop connected with the internal one. The architecture of such a system is shown in Figure 2, and similar to common automatic mass production systems, buffers are allocated between operations to mitigate the negative effect of uncertainty on production rate. The system belongs to the multi-loop closed manufacturing systems (MCMS) class (Levantesi 2001).

The goal of this work is to develop a comprehensive study on MCMS characterization and operation. EV stator production is the reference application but not the only one. Other industrial MCMS applications can be indeed found in other fields.

For instance, the semiconductor production process is characterized by a repetition of operations to integrate different circuits on the same silicon wafer to produce a single semiconductor; this integration is performed multiple times with the same machines (i.e. with an internal closed loop) and the overall production system becomes an MCMS (Hatch and Mowery 1998). Another example regards the automotive sector, where all the car doors must be cleaned, painted, coated, and inspected for quality issues; all these processes can be carried out more times in the same workstations, leading to an internal loop in the system and to a consequent MCMS architecture (Resano Lazaro and Luis Perez 2009). One last example is represented by assembly/disassembly lines: these are characterized by one internal loop opened by disassembly machines and closed by assembly machines that is normally introduced to allow different sub-assemblies to receive different treatments before being joined for common steps (Levantesi 2001); being parts carried on pallets circulating over and over in the internal loop, also assembly/disassembly lines are characterized by MCMS layout.

1.2. *Scientific Background*

Literature for MCMS is not wide and only few works are present. In (Resano Lazaro and Luis Perez 2009; Ferreira et al. 2011, 2012) the authors developed an analytical model to evaluate the performance of a four-loop closed system used in the automotive sector for automobile assembly and preassembly lines. However, the overall methodology and all the results were limited to the application to that specific MCMS only. A different approach can be found considering MCMS for semiconductor assembly lines. In this case, the effect of different parameters (pallet number, failure rates, processing times, and buffer capacities) on system throughput was evaluated through simulation on a two-loop closed manufacturing system (Li, Zheng, and Li 2007, 2009). Nevertheless, the latter was characterized by a peculiar layout: only three machines were present, one in each loop plus a third shared machine connecting the two loops. For this reason, all the extracted properties and results were limited to that specific two-loop system under study and cannot be applied to other MCMS layouts. It is also possible to find approximate analytical methods for MCMS performance evaluation (Levantesi 2001; Zhang 2006). However, these models have strong limitations, such as identical processing time for all the machines and exponentially distributed failure times: their results cannot be extended to general cases of MCMS. Finally, it is possible to find a considerable amount of literature for production systems with one or more rework loops that might recall the MCMS architecture (recent studies in (Li et al. 2009; Zhou and Lin 2020; Zhu, Chang, and Arinez 2020)). Within the rework loops, the defective parts are repaired and sent back to the production line for reprocessing so that the production quality is improved (Li 2004). However, unlike MCMS, only a small percentage of the total produced items passes through one of the rework loops; hence, the amount of parts in the rework loops is not as relevant as in MCMS to evaluate system performance. Secondly, the number of items going into the rework loops cannot be controlled since it depends on the production scrap rate, while in MCMS this number must be controlled in order to optimize system throughput. For this reason, these systems are different from the class of MCMS analyzed in this work. Finally, the practice of EV stator production, in which each job must pass through the internal loop for a fixed number of times, does not fall into the scope of any of the above-mentioned works, hence, no characterization and operation insight on the MCMS studied in this work have ever been addressed in the literature.

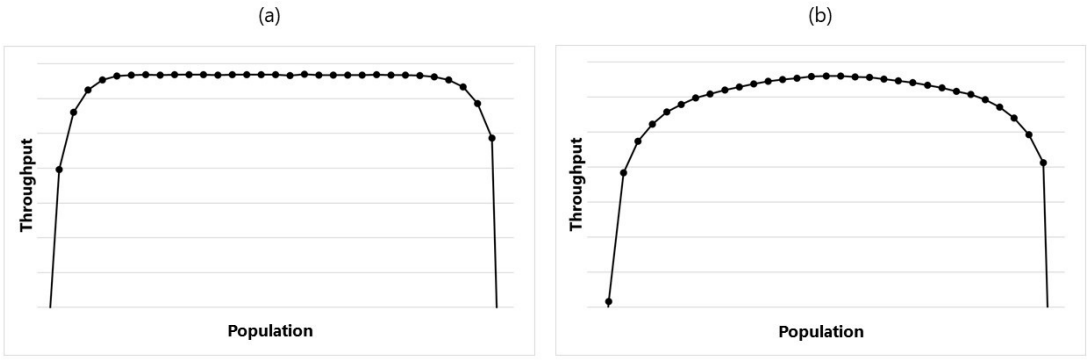

Figure 3. Two examples of high (a) and low (b) near flatness.

A very important phenomenon concerning closed loop systems is the *near flatness* (NF), described firstly by (Gershwin and Werner 2007) and analyzed deeper by (Loffredo et al. 2020). The NF is characterized by the relationship between loop population and throughput. For single loops composed by Bernoulli machines, the throughput is non-monotonic concave on population, which is proved with an exact analytical performance evaluation approach for two-machine-two-buffer loops and demonstrated with numerical experiments in (Biller et al. 2008). Being non-monotonic concave, the first-order derivative of the throughput decreases from positive to negative, and the throughput reaches the optimum and suboptimum when the derivative is equal to and close to zero, where the NF can be observed. For single loops composed by machines with phase-type distributed processing time and block-before-service assumptions, the throughput is proved to be symmetric on population which ranges from zero to the total buffer capacity (Dallery and Towsley 1991). The symmetry is also shown numerically in (Gershwin and Werner 2007; Shi and Gershwin 2014), with approximated analytical methods developed for multi-stage loops composed by unreliable machines with geometrically distributed failure and repair time. Combining the non-monotonic concave and symmetric properties, it can be expected that the range of NF appears in the center of the throughput-population curve for single closed loops. For different systems, the range of NF can be wide or narrow, which are defined as high (Figure 3-(a)) or low NF (Figure 3-(b)), respectively. In (Gershwin and Werner 2007) there is a first and qualitative definition of NF, studying the phenomenon for single-loop closed systems. In (Loffredo et al. 2020) the NF phenomenon was analyzed for MCMS. Both of the works found that balanced systems, in terms of buffer distribution or processing time, usually have low NF. In (Loffredo et al. 2020), the authors also found out that, for unreliable machines, increasing machine efficiency leads to increasing NF. In this work also a first quantitative definition of NF is provided, but this measure was dependent on parameters selected by the experimenters and this may lead to non-general results. Therefore, there are no works in the literature providing a quantitative indicator to properly describe the system NF without any user-defined parameters. Moreover, all the results of (Gershwin and Werner 2007) were limited to the study of single loops, without taking into account MCMS and, lastly, in (Loffredo et al. 2020) the only parameters analyzed as leading factors to the MCMS-NF are the efficiency of machines and balancing in capacities allocation of the buffers. However, the authors did not provide any practical insight on how to apply these relationships in the management and design of an MCMS.

1.3. *Problem*

In this work, the focus is on MCMS characterized by an external and an internal loop, as in Figure 2. In this type of production system, parts are carried on pallets. Each part is processed once by the workstations in the external loop, while the operations performed in the internal loop are repeated for a fixed number of times.

Without a proper control policy to manage the load and release of pallets for the internal loop, the system may be subject to deadlock, i.e. a complete block of the movement of pallets in the internal loop and a possible stop of the production. To solve this problem, it is required to maintain the internal loop population lower than a specific threshold value, guaranteeing the movement of items in the internal loop. However, this threshold must be carefully selected because it affects the system throughput. A high threshold will lead to an excessive population and high blocking in the internal loop; on the opposite side, a low threshold will lead to low population size and bring high starvation in the internal loop: both situations lead to a throughput decrease. In the in-between conditions, there will be a range of optimal thresholds leading to optimal population values able to guarantee the highest (or nearly highest) throughput, which is directly related to the system NF. Furthermore, it is also worthwhile to study how the system NF is affected by various system parameters: understanding how to handle the NF means understanding how to handle flexibility in terms of threshold choice. Thus, the main goal of this work is to provide new prospects, intended as practical guidelines, on MCMS operation starting from the NF phenomenon.

1.4. *Contribution*

The study of MCMS is becoming increasingly important, given the industrial relevance of this system architecture. However, literature on MCMS is not deeply developed and the study of the NF phenomenon is limited to non-extensive analysis reported in few works. To fill this gap, this work focuses on the NF phenomenon for MCMS and it provides the following contributions:

(i) A novel non-parametric NF indicator, called *Near Flatness Indicator* or *NFI*, is proposed. *NFI* represents a quantitative measure that is able to describe the system NF. This measure is independent of any user-defined parameters. Hence, *NFI* leads to general and consistent results.

(ii) Properties of *NFI* are studied through numerical experiments using an analytical method on a two-machine-two-buffer single loop and discrete event simulation on MCMS for performance evaluation. Several leading factors to NF are analyzed: efficiency in isolation of the unreliable machines (Li and Meerkov 2008), the capacity of the buffers, and machine mean times to repair ($MTTR$). The analysis is firstly performed on test cases of single-loop closed systems and MCMS (more details in sections 3.1 and 4.1). Then, the same analysis is executed on an industrial case of MCMS (more details in section 4.2), to confirm the results for a real MCMS.

(iii) Practical guidelines for MCMS operation are provided based on the experiment results. Despite this work being inspired by the EV field, results are general and can be applied to other fields such as semiconductors, assembly and disassembly lines, and the automotive sector.

It must be noted that the experimental campaign leading to the second and third contributions provided by this work is partly carried out by running simulation ex-

periments. Simulation is extensively being used as a tool aiming to better design and manage operations of a manufacturing system, with the goal to increase productivity, quality, energy efficiency, safety, and so on. Furthermore, with the advent of Industry 4.0, digitalization has played a key role in the creation of factories of the future, and this also improved simulation effectiveness: the simulation-based technologies constitute a focal point of digital manufacturing solutions (Mourtzis 2020).

In detail, the work is organized into six sections. In section 2, *NFI*, the new quantitative indicator describing the NF phenomenon, is presented. In section 3 the analysis for the single-loop system is presented, pointing out NFI properties that can be used as a reference point for the MCMS study. In section 4, the MCMS experiments are reported, studying possible leading factors to NF for MCMS. Different practical guidelines for MCMS operation are presented in section 5. Conclusions and further developments are discussed in section 6.

2. Near Flatness

2.1. *Considerations about Near Flatness Causes*

In this work, the type of MCMS system considered is always characterized by an internal and an external loop. This structure is general enough to represent many realistic situations in the industry. Moreover, the slowest processes are related to the internal loop, since the associated operations must be repeated many times. Consequently, the throughput is more affected by the internal other than the external loop. Hence, the main focus of this work is the internal loop of the MCMS and, in particular, the throughput-threshold connection leading to the system NF (as described in section 1.3).

To qualitatively show the NF phenomenon, a graphical representation of the throughput trend according to the threshold value is needed. This is called Throughput-over-Threshold curve, or T-o-T curve. An example of T-o-T curve, extracted for the MCMS test case (system layout and parameters in section 4.1), is reported in Figure 4-(a), along with the detail of the average starvation and blocking probabilities for the internal loop machines in Figure 4-(b). In the left section of the T-o-T curve, the throughput increases significantly as the threshold increases, since the population increases and the starvation of the internal loop machines is reduced. Then, with a further threshold increase, there is a range of thresholds leading to optimal population values able to guarantee the highest (or nearly highest) throughput. This range represents the system NF. In the right section of the curve, high thresholds lead to increasing population values and, consequently, increasing machines blocking: for this reason, the throughput decreases. However, when the buffer downstream of the internal loop is not full, parts can always be released to the station downstream of the internal loop. Hence, the blocking due to high internal population does not occur as frequently as starvation due to the low population. Therefore, the T-o-T curve might be not symmetric: the first part has an increasing trend while the second half might show a less evident curvature. The asymmetry is different from what is observed in single-loop closed systems (Dallery and Towsley 1991).

Being the T-o-T curve strictly related to the population influence on machines starvation and blocking, also the NF is connected to this population effect. Systems characterized by high starvation propagation will show low NF: to slow down the starvation propagation (and get to the highest throughput), higher thresholds leading to

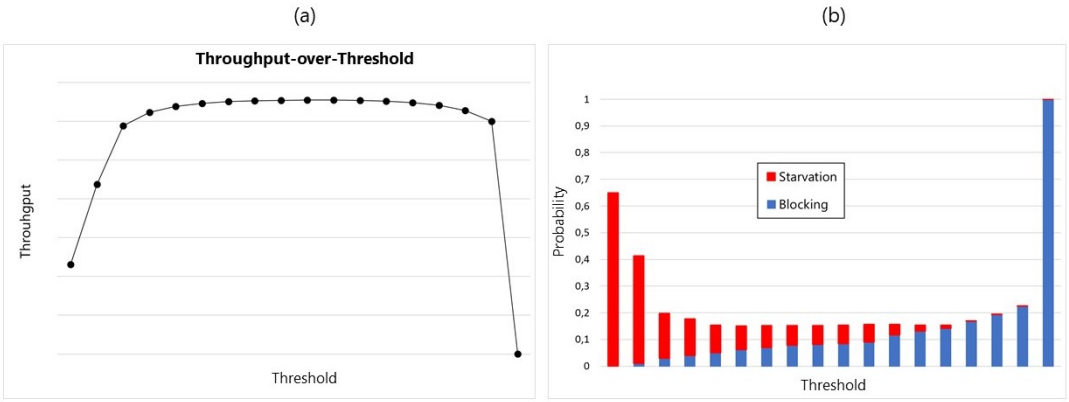

Figure 4. Example of T-o-T curve (a) and detail of the blocking and starvation probabilities of the internal loop machines (b).

higher population values are required. Thus, in this case, the left part of the T-o-T curve has sharp curvature and this leads to low NF. Moreover, also systems characterized by high blocking propagation will show low NF: if the population size is high, blocking occurs easily in the system. Consequently, for high thresholds leading to high population values, the throughput strongly decreases due to machines blocking. Hence, in this case, the right part of the T-o-T curve has sharp curvature and this leads to low NF. On the other hand, systems where it is difficult for blocking and starvation to propagate will show high NF: since starvation and/or blocking propagation is low, adding or removing pallets does not lead to significant throughput changes. Thus, the T-o-T curve will show a wide range of thresholds leading to the highest (or nearly highest) throughput and NF will be high.

2.2. *Near Flatness Indicator (NFI)*

The only work defining a quantitative indicator to describe the NF in the literature is (Loffredo et al. 2020). Their measure was dependent on parameters selected by the experimenters and this may lead to non-general results. For example, their indicator was based on a so-called *indifference zone*, i.e. a zone of the T-o-T curve where the value of the throughput is considered "optimal" if it is greater or equal to 95% of the maximum throughput. Through this zone, they evaluated the NF of a loop system. However, the choice of considering a throughput optimal if it is even 5% lower than the maximum value might not be suitable for extremely high-productivity sectors such as semiconductors, where a 5% production drop cannot be accepted. Hence, this 95% parameter depends on experimenters' choices and might generate a lack of generality. Moreover, the parameter from (Loffredo et al. 2020) also focused mainly on the right section of the T-o-T curve, looking at blocking phenomena and not properly including in the indicator-computation the left section and associated starvation. For all these reasons, it is considered highly important in the MCMS research field to have a quantitative indicator for the NF, not depending on any user-defined parameters, and properly considering all the phenomena affecting the T-o-T curve. The new indicator proposed in this work, hence, does not rely on any parameter defined by the experimenters and can be extracted directly from the T-o-T curve. Moreover, this new indicator also takes into account the whole T-o-T curve, considering both starvation and blocking effects on the NF. This indicator is referred to as *Near Flatness Indicator*

or *NFI*, and it is defined as follows:

$$NFI = \frac{1}{N} \sum_{i=1}^{N} \frac{TH_i}{TH_{max}} \tag{1}$$

where N denotes the maximum possible threshold, equal to the total internal loop buffer capacities plus the number of machines in the loop, TH_i represents the throughput when the threshold is equal to i, TH_{max} represents the highest throughput that can be achieved when the threshold is optimally chosen (the maximum throughput in the T-o-T curve). According to Equation (1), NFI is between zero and one. With an extremely low NF, throughput will be significantly reduced if the threshold is not optimally chosen. On the contrary, the maximum value of NFI can be reached in an ideal perfectly flat T-o-T curve, where the throughput does not change for the different thresholds. In conclusion, given this definition of NFI, the higher the value of NFI, the higher the NF of the MCMS.

It must be noted that this definition of *NFI* is required because it also allows identifying quantitative, general, and straightforward relationships between NF phenomenon and system parameters as reported in the analysis of sections 3 and 4.

3. Near Flatness Analysis for Single-loop Systems

In this section, single-loop closed systems are the focus, understanding how NF is affected by: (i) the efficiency in isolation of the machines referred to as η, (ii) the capacity of the buffers, and (iii) the $MTTR$ of the machines. In particular, η is defined as a function of machine "Mean Time To Repair" ($MTTR$) and "Mean Time To Failure" ($MTTF$) :

$$\eta = \frac{MTTF}{MTTR + MTTF} \tag{2}$$

Experiments are performed on a two-machine-two-buffer single-loop test case (section 3.1). Throughput in the various scenarios is evaluated with the analytical method proposed in literature (Gershwin 1994). By using an analytical approach not affected by sampling noises, single-loop results can be used as a guaranteed benchmark for all the experiments on MCMS subsequently performed (section 4).

3.1. *Single-loop Test Case*

The first test case is a single-loop closed manufacturing system, composed of two identical machines (M1 and M2) and two buffers with identical finite capacities (B1 and B2) shown in Figure 5. Processing times are deterministic, failure time and repair time are exponentially distributed. System parameters can be found in Table 1.

Being the system a single loop, there is no threshold to select: the parameter to be optimally selected to guarantee the highest throughput is the population itself. Thus, similarly to the T-o-T curve, for single-loop systems it is possible to extract a Throughput-over-Population curve (or T-o-P curve). From the latter, *NFI* can be

computed with Equation (3) where: P denotes the maximum possible population, equal to the total loop buffer capacities plus the number of loop machines, TH_i represents the throughput when the population is equal to i and TH_{max} represents the maximum throughput in the T-o-P curve. It must be noticed that the definition of NFI for single-loop systems and MCMS is the same. Hence, the only difference among equations (1) and (3) is the parameter to be selected: the population or the threshold.

$$NFI = \frac{1}{P} \sum_{i=1}^{P} \frac{TH_i}{TH_{max}} \tag{3}$$

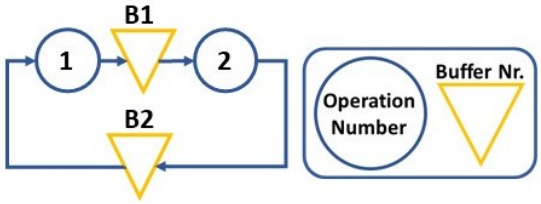

Figure 5. Layout of the first test case: the single loop.

Table 1. Single-loop test case machine processing times, time to failure (TTF), time to repair (TTR), and buffer capacities.

Op. Number	Processing Time [s]	TTF [s]	TTR [s]	Buff. Number	Capacity [Units]
1	60	Exp(3240)	Exp(360)	1	5
2	60	Exp(3240)	Exp(360)	2	5

3.2. *NFI Analysis and Experimental Results*

The experimental results presented below are extracted with the literature algorithm (Gershwin 1994) implemented in *Matlab R2020a (Mathworks)* software and results have been obtained utilizing a laptop with 4.90GHz i7 Intel Core and 16GB RAM. One single experiment with this method had a duration of about 20 seconds. For each studied configuration, the parameter to be measured is NFI: in order to compute it, it is required to extract the overall T-o-P curve associated with the configuration itself. A detailed framework of the experiments performed with the single loop test case is visible in Figure 6. For each factor to analyze, i.e. η, capacity of the buffers or $MTTR$, different levels at which varying it are selected; each of these levels corresponds to a different system configuration where the associated NFI must be identified. Hence, for each configuration, different experiments are carried out. In each experiment, a different population value i is selected and the resulting TH_i is computed with the literature algorithm (Gershwin 1994). This is repeated for each i value suitable for the system configuration under study: at this point, the T-o-P curve and, consequently, the NFI associated with that specific configuration is extracted with Equation (3). This process must be repeated for each level of each factor, i.e. for each configuration

to be studied. In total, 52 configurations are tested, corresponding to 3588 experiments and an overall computational time of about 20 hours.

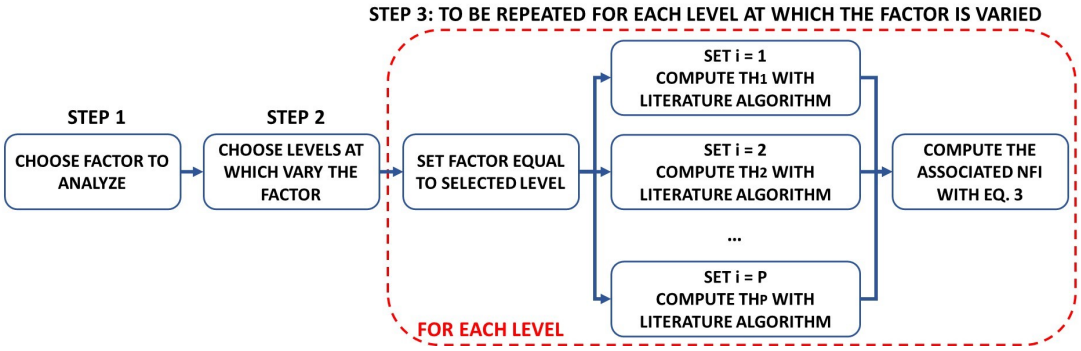

Figure 6. Detailed framework of the experiments performed with the single loop test case.

3.2.1. Efficiency

In this analysis, all the machines have the same efficiency, varied from 0.05 to 0.95, and the same $MTTR$. When machine efficiency is varied, the $MTTR$ is kept constant, and $MTTF$ is computed according to Equation (2). Machine processing times and buffer capacities are not modified. In Figure 7-(a) the corresponding NFI values are displayed: NFI increases continuously as the efficiency also increases. In Figure 7-(b) three examples of resulting T-o-P curves are reported, showing how higher efficiency leads to higher NF in the curve.

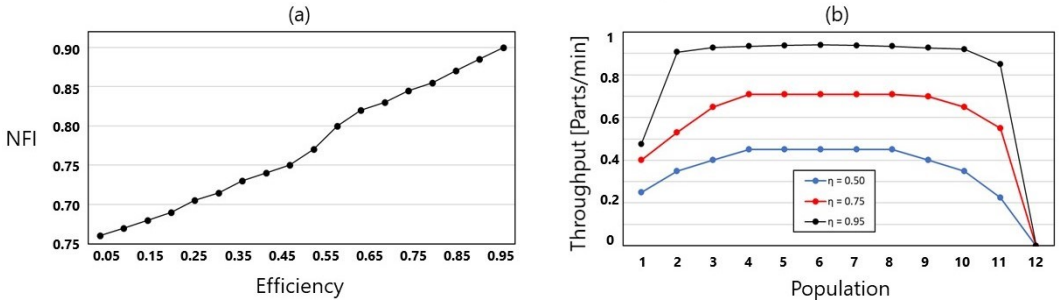

Figure 7. Machines efficiency variation effect on the NF for the single-loop test case: NFI (a) and examples of T-o-P curves (b).

3.2.2. Capacity of the buffers

The second analysis is performed with varying buffer capacities from 1 to 100, while all the other parameters are not modified. In Figure 8-(a) the corresponding NFI values are displayed: NFI increases rapidly in the first part and slowly in the second part, even though the capacity continues to enhance. It is important to notice that varying the buffer capacities means also varying the maximum population that can circulate in the loop, as visible in Figure 8-(b) with three examples of resulting T-o-P curves. However, given the definition of NFI, this variation does not affect the NF comparison among the different systems.

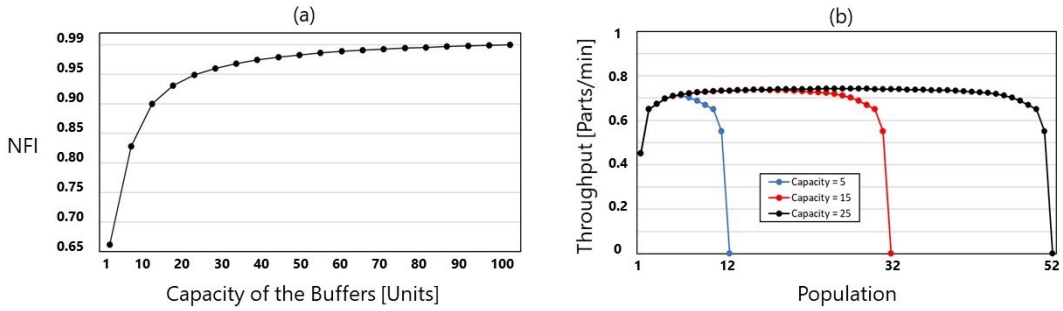

Figure 8. Buffer capacities variation effect on the NF for the single-loop test case: NFI (a) and examples of T-o-P curves (b).

3.2.3. Mean Time To Repair

In this analysis, all the machines have the same efficiency value, equal to 0.90. $MTTR$ is varied from 30 to 3120 seconds, while $MTTF$ is computed according to Equation (2): increasing $MTTR$ leads to increasing $MTTF$, i.e. decreasing failure-frequency. In order to generalize the results, it is possible to compare the selected $MTTR$ values with the machine processing time (PT), equal to 60 seconds. In this case, the $MTTR$/PT ratio varies from about 0.5 to 50. Machine processing times and buffer capacities remain unchanged. In Figure 9-(a) the corresponding values of NFI are displayed: NFI initially decreases and reaches its minimum value, then it grows again for large values of $MTTR$. This behavior is confirmed by the three examples of resulting T-o-P curves shown in Figure 9-(b): curves with high and low $MTTR$ show higher NF than the curve with intermediate $MTTR$.

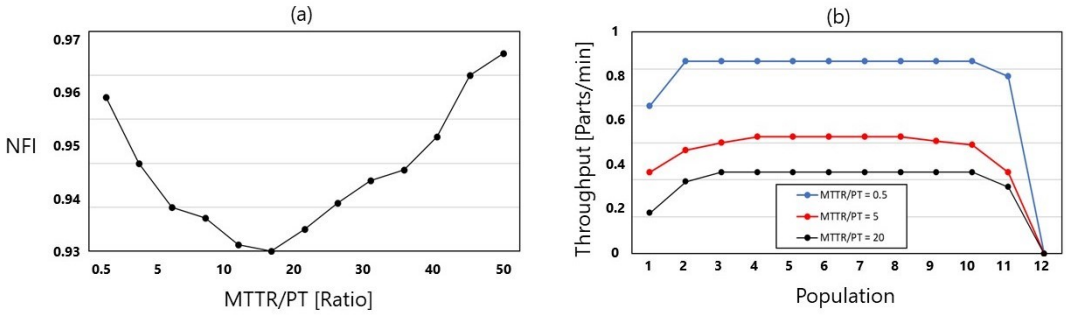

Figure 9. Machines $MTTR$ variation effect on the NF for the single-loop test case: NFI (a) and examples of T-o-P curves(b). The $MTTR$ variation is always shown using the $MTTR$/PT (processing time) ratio.

3.2.4. Discussion

The experimental results for the single-loop test case lead to three main results that can be used as a reference point for the subsequent MCMS analysis.

The first conclusion is that when machines' efficiency increases, also NFI increases. With low efficiency, indeed, machines have more failures during production time. When a machine fails, if the population is low, starvation propagates easily to other machines in the system. On the other hand, when a machine fails, if the population is high, the blocking propagates easily to other machines in the system. Thus, with low values of efficiency, the system is characterized by high starvation/blocking propagation: the

T-o-P curve shows low NF and this leads to a low NFI. With high efficiency, instead, machines are less prone to fail. For this reason, it is difficult for blocking and starvation to propagate: the T-o-P curve shows high NF, and NFI is high too.

The second result is related to the capacity of the buffers: when this increases, NFI initially increases rapidly and then slowly, even though the capacity continues to enhance. When a machine fails, its downstream machine can continue processing parts that are accumulated in the buffer downstream of the failed device. Thus, with high buffer capacity, the downstream machine can process a high number of parts accumulated in this buffer and its starvation probability is reduced. On the other hand, while the failed machine is under repair, its upstream machine continues processing and releasing parts to the in-between buffer until the latter is not full. Thus, with a high buffer capacity, the upstream machine can release a high number of parts during this time and its blocking probability is reduced. As conclusion, increasing the capacity of the buffers means increasing the difficulty for blocking and starvation to propagate: the T-o-P curve is characterized by higher NF, and NFI is higher too. When buffer capacity reaches a certain value, it is actually almost equivalent to infinite capacity, and continuing to increase buffer capacity will have almost no effect on the system performances, including NF. For this reason, increasing further the capacity does not lead to significant effects for the NF and NFI.

The last outcome is related to the $MTTR$: when machines $MTTR$ increases, NFI shows a *tick* pattern, i.e., first decreasing and then increasing. Very low values of $MTTR$ lead to short downtime of the machines. Even though the failure frequency in this case is high, being $MTTF$ also low, machines are failed for a "short time": it is difficult for blocking and starvation to propagate. Hence, the T-o-P curve shows high NF and high NFI. On the opposite side, high values of $MTTR$ lead to high values of $MTTF$: low failure frequency. However, for high $MTTR$ values every time a failure occurs the loop is completely stopped. This occurs because the time to repair is really long: the number of pallets has no influence on starvation and blocking propagation, since it is impossible to reduce them, being the system completely blocked. Hence, there is a high NF in the T-o-P curve and thus a high NFI. Finally, in the in-between situations, each failure causes starvation and blocking propagation: the T-o-P curve has increasing curvature and this leads to low NF and a low NFI.

4. Near Flatness Analysis for MCMS

Similarly to single-loop systems, in this section, the throughput-threshold relationship is studied, analyzing how NF for MCMS is affected by: machine efficiency in isolation, mean times to repair, and buffer capacities. Two systems are studied: a test case of MCMS (described in section 4.1) and a real MCMS from the industrial sector (described in section 4.2). Numerical experiments are obtained with discrete event simulation, under the assumptions presented in section 4.3. Simulation experiments are performed first on the test case to demonstrate the relations between NF and the studied factors for general cases of MCMS. Then, the same analysis is executed on the industrial case, to confirm the obtained results for the real system under investigation. All the results are reported in section 4.4. Lastly, in Section 4.5, it is assessed how relevant each of the three factors is in affecting system NF, and consequently, NFI; this analysis is carried out only on the MCMS test case.

4.1. *MCMS Test Case*

The MCMS test case is composed of an internal and an external loop (Figure 10). The latter has two single machines, two buffers of finite capacity (B1 and B4), and buffer B5 of infinite capacity. The operations in the internal loop are repeated 4 times. This loop is made of three machines and three buffers of finite capacity (B2, B3, and B6). All the machines have deterministic processing times as reported in Table 2. The processing times for Op.2, 3, and 4 are referred to one single repetition. Since those operations are repeated 4 times and represent the slowest processes of the overall system. All machines are unreliable. Time to failure and time to repair distributions are also reported in Table 2. Finally, details about test case buffer capacities are reported in Table 3.

Table 2. MCMS test case machine processing times, time to failure (TTF), and time to repair (TTR) distributions. "Triang" stands for Triangular distribution.

Op. Number	Processing Time [s]	TTF [s]	TTR [s]
1	46	Triang(20,30,45)	Triang(54000,67500,81000)
2	84	Weibull(1.0437; 41339.0)	Weibull(1.9206; 375.5)
3	90	Weibull(1.1127; 45849.0)	Weibull(1.4411; 566.2)
4	88	Weibull(1.4371; 12849.0)	Weibull(1.4371; 12849.0)
5	110	Triang(120,180,300)	Triang(2700,5400,27000)

Table 3. MCMS test case buffer capacities.

Buff. Number	Capacity [Units]	Buff. Number	Capacity [Units]
B1	2	B4	5
B2	6	B5	∞
B3	6	B6	6

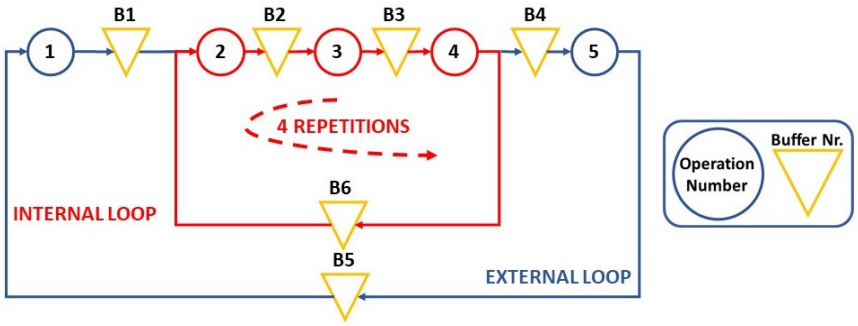

Figure 10. Layout of the second test case: the MCMS.

4.2. *Real Case*

The real case analyzed is an MCMS with an internal and an external loop (Figure 11). This system is used to produce hairpin stators in the EV field. In the manufacturing system, stators are carried on pallets, and pallets are carried by a conveyor

connecting all the workstations. The production process consists of 18 automated operations and 1 manual operation (Op.17). All the automated machines are unreliable and have deterministic processing times, while Op.17, since it is manually executed, has a stochastic behavior. Details about processing times are reported in Table 4. For which concerns the automated machines, all the processing times are provided by the company owning the industrial system under study. Op.4 is performed by three identical machines (Op.4.1, 4.2, and 4.3). Processing times reported for Op.10, 11, and 12 are referred to one layer. Since the produced hairpin stator has 4 layers, those operations are repeated 4 times and they represent the slowest processes of the overall system. Op.17 is a manual operation and has a stochastic processing time fitted from real data provided by the company (data-size equal to 20 and *p-value* for *Kolmogorov-Smirnov* test (Massey and Frank 1951) equal to 0.59). Each machine can work on one piece at a certain time. All the buffers have finite capacity, as reported in Table 5: also these data are provided by the company. Further details about system parameters are reported in Appendix A.

In Op.1, stator cores are placed on pallets (loading station) and the production process begins. Operations from Op.2 to Op.9 regard the insertion of insulation paper in the stator slots, the hairpin insertion (performed in parallel by three machines during Op.4), and the hairpin enlarging, twisting, cutting, and welding, followed by a series of checks after the welding is performed. The following three processes (Op.10, Op.11, and Op.12) are repeated four times, because of the presence of four layers (as explained in section 1.1). In the internal loop, each layer is inserted, crimped, and finally welded respectively in Op.10, Op.11, and Op.12. Afterwards, operations from Op.13 to Op.16 regard the blowing operation, removing the remaining dust on the stator, and the resining operation, where stators are covered with insulating gel and resin. Then, a final visual test is performed (Op.17) by an operator. Parts rejected from the test are sent to a repairing station (Op.18). Completed stators with no defects, instead, are unloaded from the pallets by a robot (Op.19). Finally, empty pallets are sent again to Op.1, ready to reiterate this path.

Table 4. Real case machine processing times.

Op. Number	**Processing Time [s]**	**Op. Number**	**Processing Time [s]**
1	150	11	90
2	38	12	88
3	136	13	110
4.1, 4.2, 4.3	144	14	92
5	150	15	156
6	144	16	86
7	144	17	Unif(50,90)
8	122	18	60
9	46	19	50
10	84		

4.3. *Simulation Models Assumptions and Experimental Framework*

The simulation models are consistent with the systems described in sections 4.1 and 4.2. Depending on the considered experiment, buffer capacities and failure data (*MTTR*, *MTTF* and efficiency) can vary. In particular, when machines "Time To Failure" (*TTF*) and/or "Time To Repair" (*TTR*) vary, both parameters are imposed

Table 5. Real case buffer capacities.

Buff. Number	Capacity [Units]	Buff. Number	Capacity [Units]
B1	25	B12	2
B2	4	B13	13
B3	7	B14	10
B4	2	B15	15
B5	2	B16	5
B6	2	B17	5
B7	10	B18	20
B8	26	B19	15
B9	5	B20	4
B10	11	B21	9
B11	2		

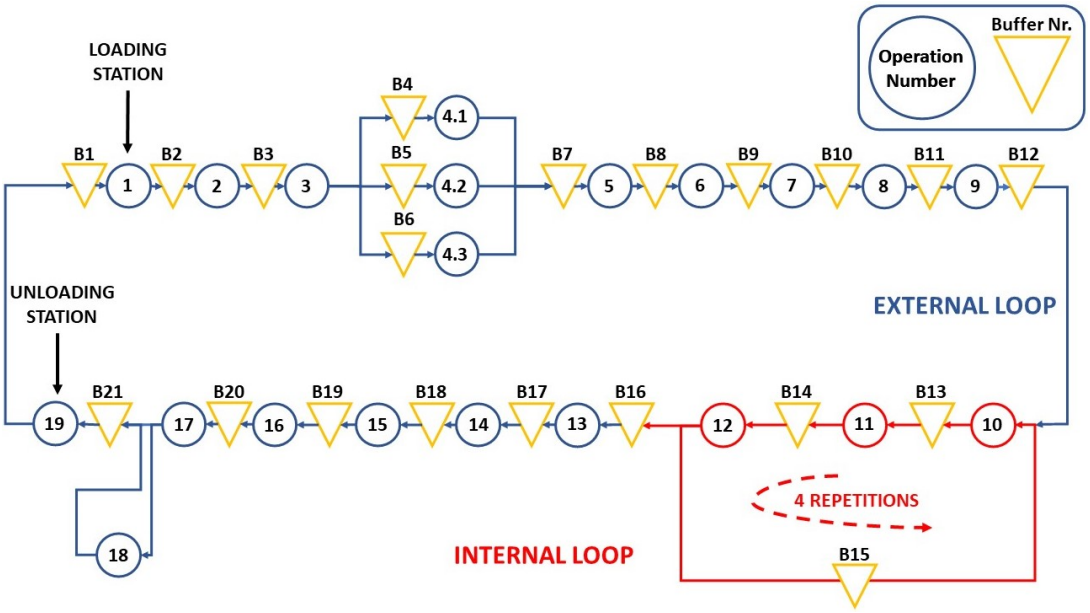

Figure 11. Real case layout.

following a Weibull distribution with shape factor equal to 1.5 and variable scale factor. The shape factor is selected higher than 1 so that the failure rate increases with time as in an aging process: machines that are more likely to fail as time goes on (Pham and Lai 2007). The selected $MTTR$ and $MTTF$ correspond to the expected values of the distributions of TTR and TTF. First come first serve and blocking after service rules are applied. All the experiments are performed with a simulation length of 365 days (one year of production time). This length is obtained considering 8 hours per shift, and three shifts per day. The same warm-up period is imposed for each experiment, which is equal to 30 days of production period: this represents an overestimation, for computational-accuracy reasons, of the transient period identified with the Welch method (Welch (1983)).

For all the analyzed MCMS configurations, the respective T-o-T curves are extracted. A detailed framework of the experiments performed with both the test and real cases of MCMS is reported in Figure 12. For each factor to analyze, different levels at which varying it are selected; each of these levels corresponds to a different system configuration where the associated NFI must be identified. Hence, for each configuration different experiments are carried out. In each experiment, a different threshold i is chosen and the corresponding TH_i is computed. This is repeated for each i-value suitable for the system under study. In this way, the T-o-T curve and the NFI related to the system configuration under investigation are computed. It must be noticed that, for the MCMS-related experiments, NFI is a random variable. Indeed, it derives from experimental results obtained through computer simulation on a stochastic system. Thus, every single experiment is replicated for a fixed number of times, equal to 10, and NFI is extracted with a 95% confidence level on the mean value. A total of 40 configurations are tested for the test case, corresponding to 14730 experiments; while for the real case, 32 configurations (i.e. 15920 experiments) are analyzed.

Arena 16.0 (Rockwell Automation Technologies) software is used to develop the simulation models for the two MCMS systems. The simulation experiments are carried out with *Arena* software using a laptop with 4.90GHz i7 Intel Core and 16GB RAM. One single simulation experiment for the MCMS test case had a duration of less than 1 minute, while for the real case this period was higher: 2 to 3 minutes. The overall computational time for the test case was about 240 hours, i.e. 10 days, while for the real case was around 530 hours, i.e. 22 days.

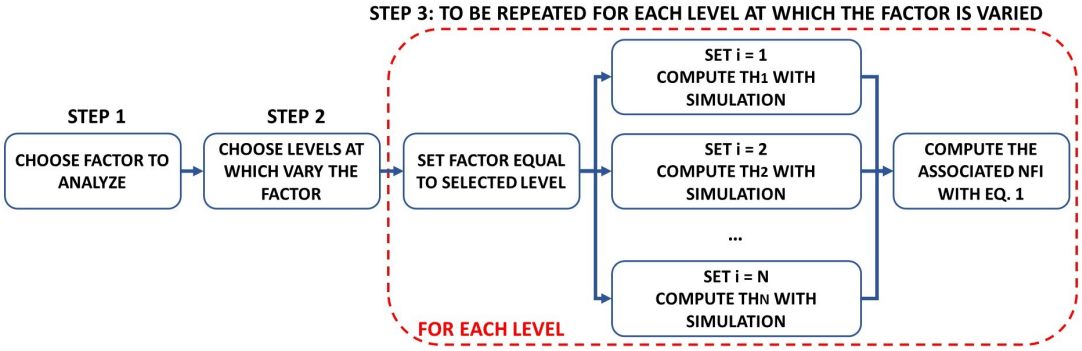

Figure 12. Detailed framework of the experiments performed with the MCMS, both the test and the real cases.

4.4. *NFI Analysis and Experimental Results*

4.4.1. Efficiency

At first, the MCMS test case is studied. All the internal loop machines have the same efficiency and the same $MTTR$. Machine processing times and buffer capacities remain unchanged. The efficiency of all the internal loop machines is varied between 0.75 and 0.99. In Figure 13-(a) the NFI values are shown. The results are consistent with the single-loop case studied with analytical methods, and this holds also for the T-o-T curves, as visible in Figure 13-(b) with three examples: higher efficiency leads to higher NF in the curve.

Moreover, similar analyses are repeated for the real case. To vary the efficiency in isolation of the internal loop devices, their real failure data are removed and substituted. Thus, the three internal loop machines have an identical imposed efficiency that varies from 0.75 to 0.99. Machine processing times and buffer capacities remain unchanged. Results of this analysis are shown in Figure 13-(c), with three examples of resulting T-o-T curves visible in Figure 13-(d), and they are consistent with the test cases.

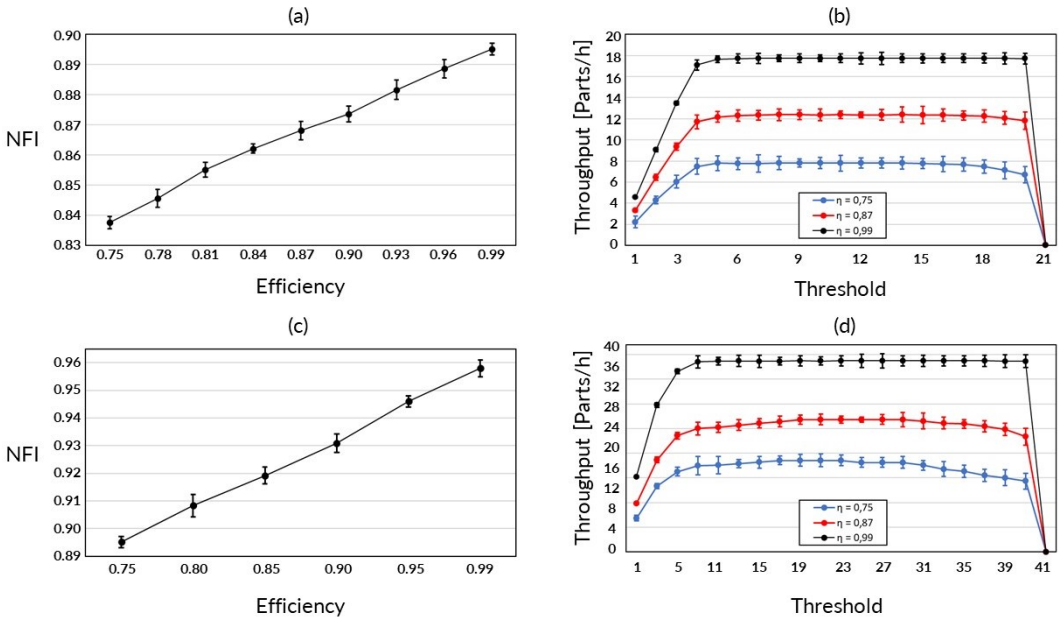

Figure 13. Machines efficiency variation effect on the NF for MCMS: NFI for the test case (a) along with examples of T-o-T curves (b); NFI for the real case (c) along with examples of T-o-T curves (d).

4.4.2. Capacity of the buffers

The first analysis is performed on the MCMS test case, all the system parameters are not changed except for the capacity of the internal loop buffers, which is identical for each buffer and varied from 2 to 50. The corresponding NFI values are reported in Figure 14-(a) and they confirm the behavior obtained analyzing the single-loop with analytical methods: when the capacity is enhanced, NFI initially increases rapidly and then slowly, even though the capacity continues to increase. Figure 14-(b) shows 3 examples of resulting T-o-T curves, still consistent with the single-loop test case.

For the real system, the parameters remain unchanged from the real ones, except for

the capacity of the three internal loop buffers: it is varied from 13 to 31. Experimental results are consistent with the test cases (Figures 14-(c) and 14-(d)).

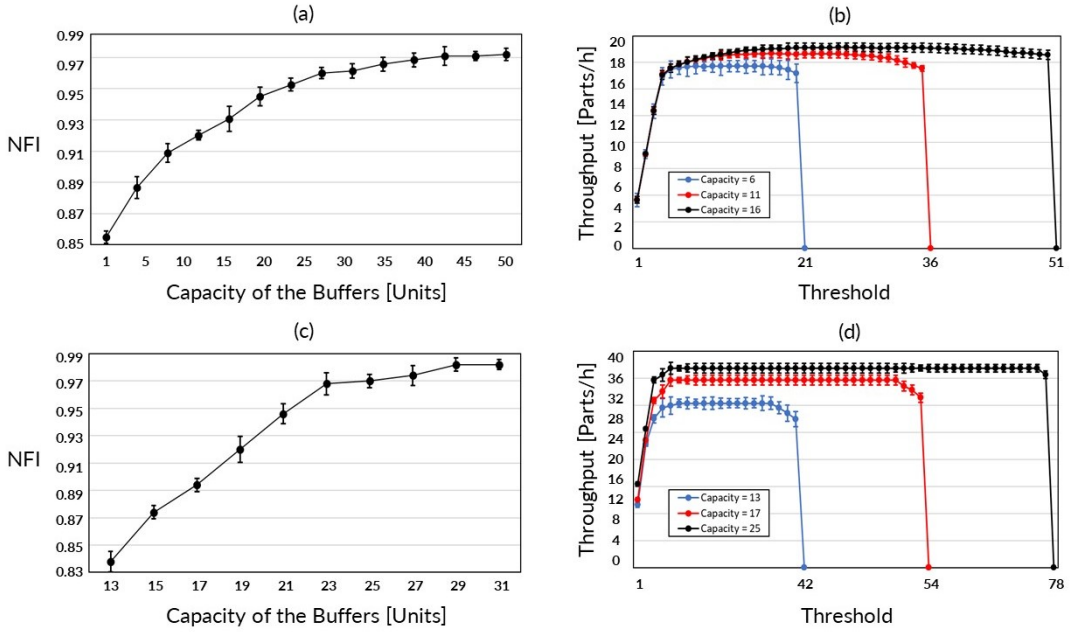

Figure 14. Buffer capacities variation effect on the NF for MCMS: NFI for the test case (a) along with examples of T-o-T curves (b); NFI for the real case (c) along with examples of T-o-T curves (d).

4.4.3. Mean Time To Repair

In this analysis, for the MCMS test case, all the internal loop machines have identical efficiency equal to 0.90. Machine processing times and buffer capacities remain unchanged, while $MTTR$ is varied from 50 to 3500 seconds. Considering the slowest internal loop machine processing time, equal to 90 seconds for Op.3, the $MTTR$/PT ratio varies from about 0.5 to about 38.5. Figures 15-(a) and 15-(b) confirm that for MCMS the results are consistent with the single-loop case, showing the same *tick* pattern.

For the real case, the real failure data of the internal loop machines are removed and substituted. All the machines have the same efficiency, equal to 0.90. Furthermore, the three internal loop machines have an identical imposed varying $MTTR$: from a minimum of 100 seconds to a maximum of 2000 seconds. Considering the slowest internal loop machine processing time, equal to 90 seconds for Op.11, the $MTTR$/PT ratio varies from about 0.5 to about 22.5. Machine processing times and buffer capacities do not vary. Results of this analysis are shown in Figures 15-(c) and 15-(d) and they are consistent with the test cases.

4.5. *Factors Relevance for NF*

An *Analysis of Variance* (ANOVA) based on a full factorial design with three factors at three levels (Montgomery 2017), is carried out to assess the relevance of the three studied factors on NFI and, consequently, the system NF. The analysis is performed on the MCMS test case; the three factors are machine efficiency in isolation, mean

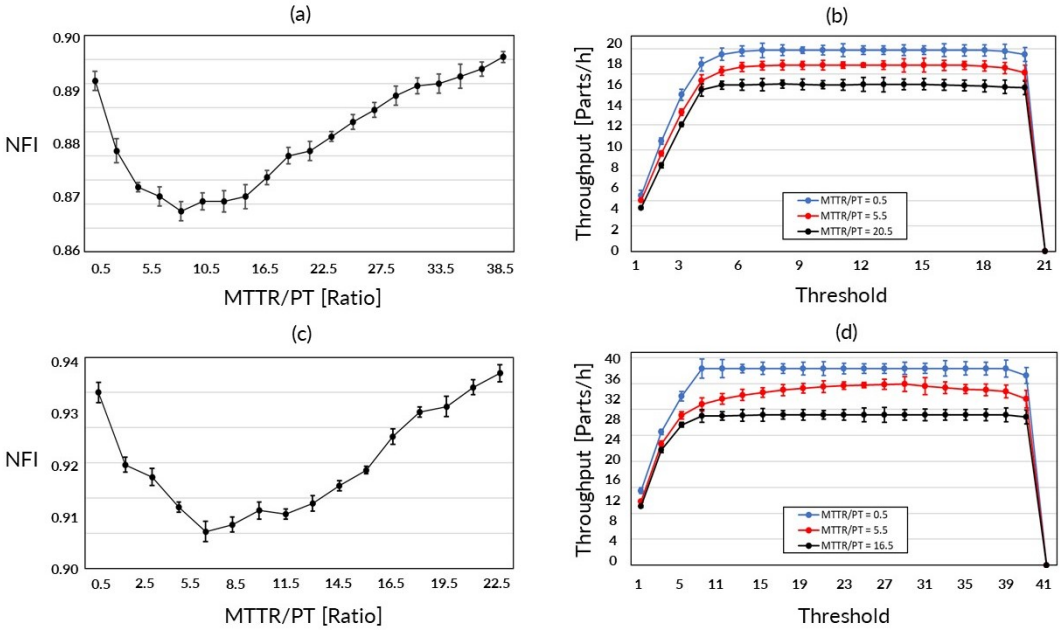

Figure 15. Machines $MTTR$ variation effect on the NF for MCMS: NFI for the test case (a) along with examples of T-o-T curves (b); NFI for the real case (c) along with examples of T-o-T curves (d). The $MTTR$ variation is always shown using the $MTTR$/PT (processing time) ratio.

times to repair, and buffer capacities while the corresponding levels are reported in Table 6. It must be noted that accordingly to the results of Section 4.4.3, $MTTR$ is taken into account through the $MTTR$/PT ratio. Furthermore, the three $MTTR$/PT values belong to three different zones of the *tick* pattern of Figure 15-(a). A total of 27 new system configurations are tested and for each of them, a NFI value is identified.

Table 6. Factors and levels for the full factorial design.

Factor	Efficiency	Capacity of the Buffers	$MTTR$/PT
Low Level	0.80	5	0.5
Medium Level	0.90	15	8.5
High Level	0.99	25	33.5

Table 7 reports the resulting p-values and F-values for the ANOVA test. The low p-values for the single factors confirm that all of them are significant and therefore they affect system NF and NFI. At the same time, the high p-values for all the second-grade interactions indicate that these are not significant for the system NF. For which concerns the single factors, the F-value of $MTTR$/PT is slightly higher and this indicates that this factor appears to be slightly more relevant than the others. Especially with really high or really low $MTTR$, indeed, system NF is high and it is really difficult to change this $MTTR$-effect on starvation/blocking propagation by varying system efficiency or buffer capacities. This leads to a slightly higher relevance of this factor. Moreover, the F-value of buffer capacities is marginally higher than the efficiency one. Indeed, especially with really high buffer capacities, there is an increasing difficulty for blocking and starvation to propagate, i.e. system has high NF: in these conditions, it is really complicated to change this effect by varying system efficiency and this leads to a slightly higher relevance of this factor in respect to

the efficiency of machines. However, all the F-values for the three single factors have the same order of magnitude: this indicates that there is not a predominant factor regarding their relative relevance and that it is not possible to extract a clearly defined order of factor priority when assessing the NF of a system. For this reason, all three factors must be taken into account when evaluating the NF of an MCMS.

Table 7. ANOVA test results.

Factor	Sum of Squares	Mean Squares	F-value	p-value
Efficiency	0.011	0.006	52.46	0.00
Capacity	0.013	0.007	62.80	0.00
***MTTR*/PT**	0.014	0.007	66.09	0.00
Efficiency*Capacity	0.001	0.000	2.19	0.16
Efficiency**MTTR*/PT**	0.001	0.000	1.75	0.23
Capacity**MTTR*/PT**	0.001	0.000	1.54	0.28

5. Discussion

The T-o-T curves in all the MCMS experiments show that the throughput first increases and then decreases as the threshold increases, which is consistent with single-loop closed systems (Li and Meerkov 2008; Gershwin and Werner 2007). Besides throughput, manufacturing system operations also aim at leanness, waste reduction, and quick response. Specifically for loop systems, on one hand, decreasing the population can reduce the WIP. On the other hand, due to near flatness, on the plateau of the T-o-T curves, throughput almost keeps constant, and the system time is proportional to the population, according to Little's Law (Little and Graves 2008). Thus, keeping a low level of population can reduce the production lead time and achieve a quick response. Hence, roughly speaking, in an MCMS the internal population threshold should be chosen on the left section of the T-o-T curve for sake of leanness; in particular, according to this perspective, the optimal threshold choice is represented by the leftmost value on the plateau of the T-o-T curve: highest (or nearly-highest) throughput achieved along with reduced WIP and system time. The study of NF provides a hint to pursue leanness and guarantee the highest throughput in the meantime.

The relations between the studied factors and NFI from empirical studies in section 4.4 are summarized in Table 8. These relations are consistent with the experimental results for the single-loop case reported in section 3.2.4 along with the discussion of the causes leading to these correspondences. The following management insights are based on Table 8.

NFI increases linearly with efficiency and, in the efficiency-related experiments, $MTTF$ increases while $MTTR$ is fixed. The change of $MTTF$ usually occurs with machine degradation or preventive maintenance, hence, the control parameter, i.e., internal loop threshold, should be re-optimized accordingly. It is possible to show an

Table 8. Summary of experiment results.

Analyzed factor	Varied factor	Fixed factors	Factor-NFI relation
Efficiency	MTTF	PT, MTTR, Buffer	Linearly increasing
Buffer	Buffer	MTTR, MTTF, PT	Concavely increasing
MTTR	MTTR, MTTF	Efficiency, PT, Buffer	Tick-phenomenon

example of this situation, analyzing the MCMS test case in three situations: with original system parameters, varying only the internal loop machines $MTTF$, i.e. increasing it due to preventive maintenance, and, again, varying only the internal loop machines $MTTF$ but in this case, decreasing it due to degradation. In Figure 16 the three T-o-T curves representing the aforementioned cases are reported. Specifically, machine degradation leads to the decrease of $MTTF$ and NF; as a consequence, the throughput becomes more sensitive to the population threshold, and the threshold may need to be increased. In the case of preventive maintenance, $MTTF$ and NFI are both increased, thus there is the possibility to decrease the threshold. According to what was stated before, in all three cases, the optimal threshold choice is represented by the leftmost value on the plateau of the T-o-T curve.

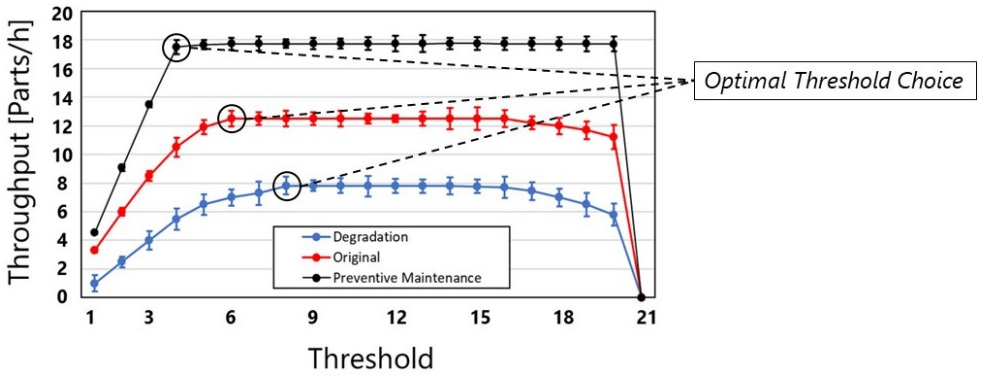

Figure 16. Degradation and preventive maintenance effect on the T-o-T curve: examples of optimal threshold choice in the various cases.

NFI increases with buffer space. Instead of operational factors, buffer space is usually decided in the system design phase and is difficult to change frequently. Allocating large buffer space in the design phase, and consequently increasing *NFI*, will facilitate the operation of loop systems. Another benefit of large buffer space in loops is that given a fixed loop population, the larger the buffer space, the higher the throughput, as in Figures 14-(b) and 14-(d). Furthermore, in loop systems, increasing buffer space without changing the loop population will not increase work-in-process. To conclude, on contrary to common sense that a smaller buffer is more favorable in saving space and reducing WIP in serial lines, increasing buffer space within the space limit in loop systems can be beneficial to system operation and throughput.

The relation between $MTTR$ and NFI shows a *tick* pattern, i.e., first decreasing and then increasing. Thus, for systems with micro failures, which are frequent but require short repair time, usually called *micro stoppages* in the industrial context, and major failures, which are rare but require long repair time, the threshold can be lower. However, the criteria for categorizing a failure as micro stoppage, major failure, or medium failure is not addressed. It is recommended to analyze the effect of $MTTR$ on NFI of each individual system.

It must be noticed that in all the analyses the parameters of the external loop are not varied. This choice is motivated by the non-relevant effect of the external loop on the overall system throughput since the internal loop is the slowest section of the MCMS. Otherwise, if the slowest machines were in the external loop, the system throughput would depend mainly on the slowest machine positioned outside the internal loop and the analysis of the latter would not be useful to understand how to improve system performance. Thus, additional analysis with different parameters of the external loop

is not interesting to characterize the NF of an MCMS. Moreover, the utilized random distributions, such as weibull, lognormal, triangular, and so on, are statistically plausible to model the stochastic time distributions, i.e. $MTTR$, $MTTF$ and processing times, characterizing a manufacturing system (Trietsch and Baker 2012). In conclusion, the presented results can be considered valid and suitable for general cases of MCMS without the need for further experiments.

All the results have been extracted considering an MCMS operating in the EV field and, therefore, they are directly applicable to any similar MCMS used in this expanding field. However, internal loops and consequent MCMS layouts are also present in other significantly wide industrial sectors, i.e. semiconductors, automotive, and assembly/disassembly lines. This means that the extracted insights can also be transferred to these sectors. The threshold choice for the internal loop is a key passage for the design, control, and management of MCMS in any of the aforementioned fields. Therefore, being able to assess what affects the NF of the respective T-o-T curves has a great impact from a practitioner's point of view in a large number of production systems used in these industrial sectors.

6. Conclusions and Further Developments

In this work, the investigation of a phenomenon called *near flatness* (or NF) related to the T-o-T curve in MCMS is performed. In particular, it is analyzed how the NF is affected by different system parameters. First, a general quantitative indicator called NFI able to describe the NF of the T-o-T curve is introduced. In addition, a deep analysis to evaluate the NF dependency on system parameters is performed. Machines efficiency in isolation, the capacity of the buffers, and machines $MTTR$ are discovered to be leading factors to NF.

Based on the numerical results, insights are addressed for MCMS design, control, and management. In the case of machine degradation and preventive maintenance, the threshold may need to be increased and decreased, respectively. Large buffer spaces, within the space limit, are recommended to improve system throughput and facilitate system operation. The relation of $MTTR$ and NFI shows a *tick* pattern, and the analysis of each individual system is recommended to recognize the bottom of the tick, where NFI is small and the threshold should be set with a higher value.

This work studies the NF regarding the internal loop population threshold in MCMS commonly seen in the EV industry and also in other industrial applications such as the semiconductor field, assembly and disassembly lines, and the automotive sector. In this architecture, there are two parameters related to the loop population, which are the internal loop population threshold studied in this work and the population circulating in the entire system. To improve the operation of such systems, the effect, relation, and interaction of the two population-related parameters raise future research interest.

As for the design of MCMS, the experiment has unveiled that allocating more buffers can increase the NF and hence flexibility in system operation. However, the distribution of buffer spaces in the loop is another important factor in system performance. The distribution of buffer spaces has combinatorial complexity, thus, the experiment to study its effect is not trivial. Instead, it is more feasible to solve the buffer allocation as an optimization problem. Since the interaction of buffer allocation and the population is not negligible, jointly allocating buffer and pallet/threshold is proved to be relevant in manufacturing system design, but seldom addressed in the literature.

Experiments of this work have qualitatively shown that the *MTTF* and *MTTR* have complicated effects of NF and optimal internal loop population threshold value. Furthermore, the *MTTR* and *MTTF* quite depend on the state of machines and maintenance activity and could vary frequently. How to decide the internal loop population threshold dynamically is another interesting research topic.

References

Athanasopoulou, Lydia, Harry Bikas, Alexios Papacharalampopoulos, Panos Stavropoulos, and George Chryssolouris. 2022. "An industry 4.0 approach to electric vehicles." *International Journal of Computer Integrated Manufacturing* 1–15.

Biller, Stephan, Samuel P Marin, Semyon M Meerkov, and Liang Zhang. 2008. "Closed bernoulli production lines: analysis, continuous improvement, and leanness." *IEEE Transactions on Automation Science and Engineering* 6 (1): 168–180.

Dallery, Yves, and Don Towsley. 1991. "Symmetry property of the throughput in closed tandem queueing networks with finite buffers." *Operations Research Letters* 10 (9): 541–547.

Ferreira, Luis Pinto, Enrique Ares Gómez, Gustavo C.Peláez Lourido, José Diéguez Quintas, and Benny Tjahjono. 2012. "Analysis and optimisation of a network of closed-loop automobile assembly line using simulation." *The International Journal of Advanced Manufacturing Technology* 59 (1-4): 351–366.

Ferreira, Luis Pinto, Enrique Ares Gomez, Gustavo Pelaez Lourido, Marina Salgado, and Jose Dieguez Quintas. 2011. "Analysis on the influence of the number of pallets circulating on an automobile closed-loops assembly line." *manufacturing systems* 14: 15.

Fleischer, J, S Haag, and J Hofmann. 2017. "Quo vadis winding technology." *A study on state of the art and research on future trends in automotive engineering* .

Gershwin, Stanley B. 1994. "Manufacturing systems engineering." *PTR Prentice Hall Englewood Cliffs* .

Gershwin, Stanley B., and Loren M. Werner. 2007. "An approximate analytical method for evaluating the performance of closed-loop flow systems with unreliable machines and finite buffers." *International Journal of Production Research* 45 (14): 3085–3111.

Hatch, Nile W, and David C Mowery. 1998. "Process innovation and learning by doing in semiconductor manufacturing." *Management Science* 44 (11-part-1): 1461–1477.

Ishigami, Takashi, Yuichiro Tanaka, and Hiroshi Homma. 2014. "Development of Motor Stator with Rectangular-Wire Lap Winding and an Automatic Process for Its Production." *Electrical Engineering in Japan* 187 (4): 51–59.

Ishigami, Takashi, Yuichiro Tanaka, and Hiroshi Homma. 2015. "Motor stator with thick rectangular wire lap winding for HEVs." *IEEE Transactions on Industry Applications* 51 (4): 2917–2923.

Kim, Sang-Hoon. 2017. *Electric motor control: Dc, ac, and bldc motors*. Elsevier.

Levantesi, R. 2001. "Analysis of multiple loop assembly/disassembly networks." .

Li, Jingshan. 2004. "Performance analysis of production systems with rework loops." *IIE Transactions* 36 (8): 755–765.

Li, Jingshan, Dennis E. Blumenfeld, Ningjian Huang, and Jeffrey M. Alden. 2009. "Throughput analysis of production systems: recent advances and future topics." *International Journal of Production Research* 47 (14): 3823–3851.

Li, Jingshan, and Semyon M Meerkov. 2008. *Production systems engineering*. Springer Science & Business Media.

Li, Na, Li Zheng, and Quan-Lin Li. 2007. "Semiconductor system with multiple closed-loops constrains." In *Automation Science and Engineering, 2007. CASE 2007. IEEE International Conference on*, 484–488. IEEE.

Li, Na, Li Zheng, and Quan-Lin Li. 2009. "Performance analysis of two-loop closed production systems." *Computers & Operations Research* 36 (1): 119–134.

Little, John DC, and Stephen C Graves. 2008. "Little's law." In *Building intuition*, 81–100. Springer.

Loffredo, Alberto, Mengyi Zhang, Ziwei Lin, and Andrea Matta. 2020. "Throughput Sensitivity Analysis in Closed Loop Manufacturing Systems for Hairpin Stator Production." *Procedia Manufacturing* 51: 1515–1522.

Massey, Jr, and J Frank. 1951. "The Kolmogorov-Smirnov test for goodness of fit." *Journal of the American statistical Association* 46 (253): 68–78.

Milas, Nikolaos, Dimitris Mourtzis, and Emmanuel Tatakis. 2020. "A decision-making framework for the smart charging of electric vehicles considering the priorities of the driver." *Energies* 13 (22): 6120.

Montgomery, Douglas C. 2017. *Design and analysis of experiments*. John wiley & sons.

Mourtzis, Dimitris. 2020. "Simulation in the design and operation of manufacturing systems: state of the art and new trends." *International Journal of Production Research* 58 (7): 1927–1949.

Parajuly, Keshav, Daniel Ternald, and Ruediger Kuehr. 2020. "The Future of Electric Vehicles and Material Resources: A Foresight Brief." .

Pham, Hoang, and Chin-Diew Lai. 2007. "On recent generalizations of the Weibull distribution." *IEEE transactions on reliability* 56 (3): 454–458.

Resano Lazaro, A., and CJ. Luis Perez. 2009. "Dynamic analysis of an automobile assembly line considering starving and blocking." *Robotics and Computer-Integrated Manufacturing* 25 (2): 271–279.

Shi, Chuan, and Stanley B Gershwin. 2014. "Improvement of the evaluation of closed-loop production systems with unreliable machines and finite buffers." *Computers & Industrial Engineering* 75: 239–256.

Trietsch, Dan, and Kenneth R Baker. 2012. "PERT 21: Fitting PERT/CPM for use in the 21st century." *International journal of project management* 30 (4): 490–502.

Welch, Peter D. 1983. "The statistical analysis of simulation results." *The computer performance modeling handbook* 22: 268–328.

Zhang, Zhenyu. 2006. "Analysis and design of manufacturing systems with multiple-loop structures." PhD diss., Massachusetts Institute of Technology.

Zhou, Binghai, and Song Lin. 2020. "A novel analysis model for optimization performance of Bernoulli serial production system considering rework processes." *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 234 (1-2): 295–309.

Zhu, Cheng, Qing Chang, and Jorge Arinez. 2020. "Data-enabled modeling and analysis of multistage manufacturing systems with quality rework loops." *Journal of Manufacturing Systems* 56: 573–584.

Appendix A. Industrial System Parameters

Details about industrial system machines time to failure and time to repair distributions are reported in Table A1 and Table A2, respectively. Data are taken from the industrial system over a processing period of three months. Data are fitted with *EasyFit (Mathwave)* software and *p-values* are reported according to the *Kolmogorov-Smirnov* test. In addition, all the machines are affected by microstoppages: failures with a duration lower than ten minutes. The *TTF* and *TTR* distributions are estimated through direct interviews with the line responsible of the company. Triangular distributions are utilized based on the interview, considering the worst case – most frequent case – best case experienced by the operators (Table A3).

Table A1. Industrial system machines time to failure (TTF) distributions. *P-values* are computed according to the *Kolmogorov-Smirnov* test.

Op. Number	TTF [s]	P-Value	Nr. Data
1	Lognormal(10.389; 1.1137)	0.786	19
3	Weibull(3.0934; 15000)	0.998	30
4.1, 4.2, 4.3	Normal(21825; 9977.8)	0.409	34
6	Gamma(0.68939; 71621)	0.475	18
8	Weibul(9.3335; 1.0169)	0.135	37
10	Weibull(1.0437; 41339.0)	0.386	13
11	Weibull(1.1127; 45849.0)	0.566	16
12	Weibull(1.4371; 12849.0)	0.786	18
15	Lognormal(8.6138, 0.71277)	0.159	35
18	Weibull(1.9237; 43047)	0.720	8

Table A2. Industrial system machines time to repair (TTR) distributions. *P-values* are computed according to the *Kolmogorov-Smirnov* test.

Op. Number	TTR [s]	P-Value	Nr. Data
1	Lognormal(7.2942; 0.5958)	0.47	19
3	Weibull(2.1811; 590.31)	0.667	30
4.1, 4.2, 4.3	Erlang(6; 168.73)	0.484	34
6	Lognormal(6.8368; 0.7548)	0.605	18
8	Gamma(1.3401; 711.86)	0.145	37
10	Weibull(1.9206; 375.5)	0.932	13
11	Weibull(1.4411; 566.2)	0.884	16
12	Weibull(1.1237; 734.86)	0.717	18
15	Lognormal(6.6131; 0.741)	0.137	35
18	Lognormal(7.7656; 1.2947)	0.897	8

Table A3. Industrial system machines *micro stoppages*.

Op. Number	TTR [s]	TTF [s]
3	Triang(480,600,720)	Triang(12000,13500,16000)
4.1, 4.2, 4.3	Triang(30,60,120)	Triang(2700,5400,27000)
5	Triang(180,240,300)	Triang(1800,12600,27000)
6	Triang(120,180,240)	Triang(7500,9000,10500)
8	Triang(100,120,150)	Triang(3000,3600,4200)
13	Triang(120,180,300)	Triang(2700,5400,27000)
Others (All)	Triang(20,30,45)	Triang(54000,67500,81000)