30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021)
15-18 June 2021, Athens, Greece.

Throughput Sensitivity Analysis in Closed Loop Manufacturing Systems for Hairpin Stator Production

Alberto Loffredo[a,*], Mengyi Zhang[a], Ziwei Lin[a,b], 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 Minhang, 200240 Shanghai, China*

Abstract

With the development of the electrical vehicle (EV) field, multiple closed loop manufacturing systems (CLMS) handling will have a predominant role in near future industry, since they are used to produce key elements of EVs. Nevertheless, literature for multiple CLMS is not wide and developed. This work is developed from an industrial case of multiple CLMS in EV industry. Through the system analysis, a phenomenon called near flatness (NF) arouse our interest. It identifies the capability of a CLMS to keep a nearly constant throughput when the number of items circulating in the CLMS varies. Further analysis about the relationship between NF and system parameters is carried out. Numerical results, obtained through discrete event simulation, show that NF phenomenon is related to the efficiency in isolation of the machines and the balancing in terms of system processing time. Finally, practical guidelines for multiple CLMS handling, related to NF, are provided.

1. Introduction

1.1. Motivation

In recent years, Multiple Closed Loop Manufacturing Systems (or multiple CLMS) are becoming increasingly important in automotive industry. A crucial factor for this importance growth is the electrification of power trains for vehicles, aimed to reduce fuel cost and the ecologic impact of mobility [15]. The electric motor is composed by two principal components: the rotor and the stator. Specifically, multiple CLMS are utilized in the production of stators. As example, a new technology of stator called "hairpin stator" is increasingly used as an alternative to the traditional technology based on wound-coils [3, 7]. It assures lower size for equal power and a superior production rate compared to traditional technology [5, 6]. The production

process to realize a hairpin stator involves a repetition of operations aimed at inserting more stacked devices, called "layers", on top of each stator. Using an internal closed loop, so that the system assumes a multiple CLMS architecture, the same machines can be visited a fixed number of times sequentially by each part, obtaining the repetition of operations without additional investments for buying new machines. In this way, there is a double advantage: cost saving and space saving. CLMS are composed by more than one "sub-loop" [8], going to build the structure of the principal system. Literature for multiple CLMS is not so deeply developed and a complete study on their handling is not present yet.

1.2. Problem

This work originates from a study conducted on a real manufacturing system producing hairpin stators. This production line is characterized by a multiple CLMS architecture, where stators are carried on pallets, whose number remains constant. The system architecture consists of two main single loops: an external loop (the "main loop") and an internal loop (Figure 1). Each pallet, which follows the main loop, travels through all the sta-

* Corresponding author
E-mail address: alberto.loffredo@polimi.it (Alberto Loffredo).

tions from the first loading station to the last unloading one and finally is carried again to the first station. In the main loop, each pallet circulates only one time, to perform operations related to the hairpin devices and to the coating of the stator itself (more details in Section 2.1). Moreover, each stator repeats for a fixed number of times the operations performed in the internal loop and, consequently, it repeats for a fixed number of times the internal loop path (the red one in Figure 1), before getting out of it. The operations performed in the internal loop concern the insertion and the welding of the layers, semicircular fixtures stacked on top of each other on each stator, completing the assembly.

With the actual design, the system may be subject to deadlock condition because of the presence of the loops. To avoid this problem, a proper control policy to manage the load and release of pallets must be applied to the internal loop, in order to limit and control the "internal loop population". A simple control policy can be the following: imposing that the internal loop population must be lower than or equal to a specific threshold value, guaranteeing the movement of items in the internal loop. At the same time, the threshold value might not guarantee the best performance of the system: few pallets in the loop might bring to high starvation of the loop machines and a decrease of the performance, while an excessive number of pallets might lead to the high blocking of the loop machines and, again, to a decrease of the performance. In conclusion, the threshold value must be optimized, since the performance might be sensitive to the internal loop population (or "population") imposed by the threshold limit.

The trend of the throughput according to the population can be graphically represented. This graphical representation is defined in this work as Throughput-over-Population curve, or T-o-P curve (more details in Section 2.2). It is of great interest to understand how the T-o-P curve, i.e. the throughput sensitivity to the internal loop population, is affected by the different design parameters of the machines. Indeed, understanding how to manage the throughput-population relation means understanding how to manage the system flexibility in terms of population choice. The overall analysis is referred only to the internal loop population because in the studied CLMS the slowest machines, i.e. the layers-welding stations, are positioned inside the internal loop and consequently the system performance is mainly affected by the internal loop.

The main goal of this work is to offer new prospects, intended as practical guidelines, on multiple CLMS handling and optimization in terms of system sensitivity to the internal loop and potential implications on the system performance.

1.3. Theoretical Background

Literature for multiple CLMS is not wide and developed, but some studies have been performed to understand their behavior and find ways to evaluate their performances. In particular, detailed studies were developed on specific class of production lines in automotive sector: through analytical models [13] and discrete event simulation [1, 2]. For what concerns more general cases, in literature [10, 11] the effect of different parameters

(pallet number, failure rates, machining times and buffer capacities) was evaluated through simulation on a two-loops semiconductor assembly production line; but, since it was referred to the semiconductor field, it concerned about a multi-product situation. Approximate analytical method for single loops was extended to multiple loops [8, 14], even though limitations in the model were still present. However, their method was not used to derive some insights on the system handling.

A very important phenomenon concerning CLMS studied in literature is the *near flatness* [4]. It is possible to talk about "near flatness" situation, or NF, in a single CLMS where there is a range of population values able to guarantee the highest (or near-highest) throughput. According to the width of this range, it is defined a high (Figure 2 - (a)) or a low (Figure 2 - (b)) NF phenomenon. The authors consider deterministic and identical processing time for each machine in each case, unreliable machines with same failure parameters, only the buffer capacities change. They found out that in symmetrical loops (i.e. same capacity for each buffer) NF was low. On the other hand, loops which were very asymmetrical (i.e. different capacity for each buffer), exhibit high NF. The degree of NF seemed to increase with the degree of asymmetry and unbalancing in the loop. In conclusion, they observed that with buffer capacities really distant and different among them, the so-called NF increased. High NF leads to a more flexible choice in the optimal population value. On the contrary, with buffer capacity-values very close among them, the population should be picked carefully. Nevertheless, their work is the only one regarding the NF present in literature and the authors focused only on single CLMS. Moreover, a quantitative definition of the phenomenon was not provided in their work. Finally, the only studied leading factor to NF was the balancing in terms of capacity of the buffers.

The focus of this work is on the phenomenon of the NF for *multiple CLMS*. For this reason, analysis are performed to evaluate the possible presence of NF phenomenon for the multiple CLMS case. Moreover, a quantitative definition of the NF phenomenon is presented. This is obtained through a quantitative indicator able to describe the NF of the T-o-P curve. Starting from this, other possible factors leading to this phenomenon are investigated in this work: efficiency in isolation of the unreliable machines [9] and processing time balancing.

In detail, the work is organized in 5 sections. In Section 2, the industrial system analysed in this work is presented, along with the description of the T-o-P curve and of the numerical indicators describing the NF phenomenon. In Section 3, multiple CLMS are analysed focusing on the effect of specific machine parameters on the system T-o-P curve. The experimental campaign with numerical cases, to support theoretical analysis, is provided. Furthermore, the confirmation of the results on real case of industrial multiple CLMS is reported. Starting from the analysis performed in Section 3, different practical guidelines for multiple CLMS handling are presented in Section 4. Conclusions and further developments are discussed in Section 5, that closes the work.
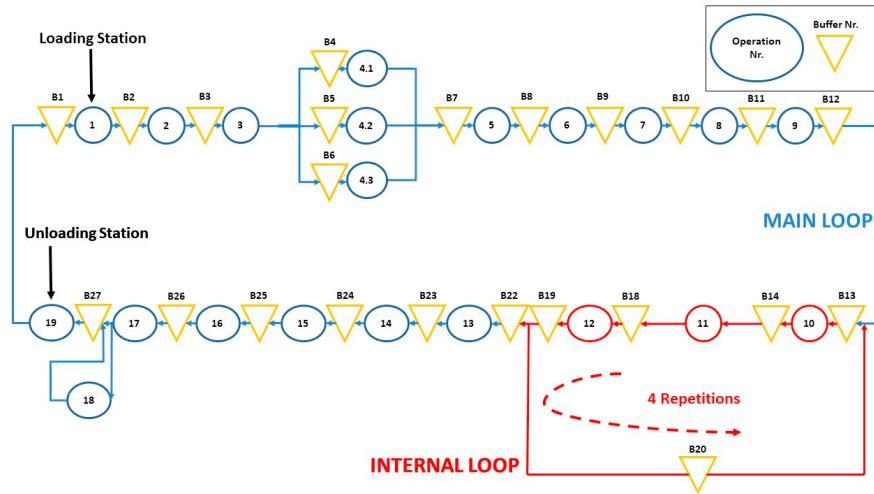
Fig. 1. Industrial system scheme : "main loop" and "internal loop" are visible.
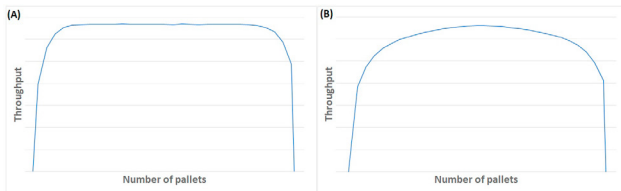

Fig. 2. Two examples of T-o-P curve with high NF (a) and low NF (b).

2. System Description and Near Flatness Indicators

2.1. System Description

The manufacturing system from which this work originates is a hairpin stator production line, characterized by multiple CLMS architecture (Figure 1). In the production system workstations are connected by a conveyor. Stators flow through the line on pallets. Once a complete piece is unloaded from the last station, the empty pallets are conveyed again towards the first loading station. The production processes consists of 18 automatic operations and 1 manual operation (Op. 17). All the machines are unreliable. Each machine can work one piece at a certain time. First Come First Serve (FCFS) and Blocking After Service (BAS) rules are applied. Operations from 1 to 9 regard the loading of the stator cores on pallets, the insertion of insulation paper in the stator slots, the hairpin insertion (performed in parallel by three machines during operation number 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. Each layer is inserted, crimped and finally welded respectively in Op. 10, Op. 11 and Op. 12. Operations from 13 to 16 regard the blowing operation, removing the remaining dust on the stator, and the resining operation. Here, the stators are covered with insulating

gel and resin. Then a final visual check is performed (Op. 17) by an operator. Parts rejected from the tests are sent to a repairing station (Op. 18). Complete stators with no defects, instead, are unloaded by a robot (Op. 19). Finally, empty pallets are sent again to Op. 1.

2.2. The T-o-P Curve

The trend of the multiple CLMS throughput according to the internal population value can be graphically represented (T-o-P curve), as in Figure 3. Starting from a limit value of zero pallets in the central loop, the throughput increases as the central loop population limit increases, since the starvation probability of the central loop machines is reduced. Then, increasing again the value of the central loop population limit, the throughput reaches an optimal value. A further increase of the loop population limit leads to a slight growth in terms of central loop machines blocking. However, when the buffer downstream the central loop is not full, there is no control on the release of parts from the loop. In this situation, there is always the possibility of releasing parts to the station downstream the central loop. As a result, the average number of pallets in the loop is lower than the limit imposed by the policy. For this reason, the central loop does not get to extremely high blocking probability, which is the situation of single CLMS. This leads to a lower reduction in the throughput than the expected one. Therefore, the T-o-P might be not symmetric. The first part of the T-o-P curve has an increasing trend, while the second half of the curve might show a less evident curvature.

2.3. NF and System Parameters

In literature, a quantitative definition of the NF phenomenon is not provided. Thus, a numerical indicator of the NF, based on the T-o-P curve, is proposed. The indicator must describe the flexibility of a CLMS, since the NF phenomenon is related
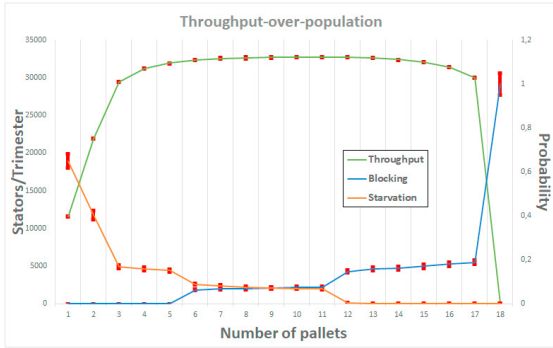
Fig. 3. Example of T-o-P curve.

Fig. 4. The framework of indicators to describe the T-o-P curves.

to it. A CLMS is considered flexible in this work when it is able to guarantee the optimal (or near optimal) throughput for a wide range of population values, and when it is less sensitive, in terms of throughput, to population changes.

The first parameter introduced is related to the width of the near flat region, where the throughput can be considered nearly constant and optimal (or near-optimal), according to the optimal throughput *indifference zone*. Indeed, as visible from Figure 3, it is possible to define an indifference zone in terms of optimal throughput. In this zone, the value of the throughput is considered "optimal" even if it is slightly lower than the maximum. In this work, values greater or equal to 95% of the maximum throughput are considered an "optimal" situation. The population limit value (or values) able to lead to optimal (or near-optimal) throughput is defined as "optimal limit of population value" (or values). The range of optimal population values corresponds to the "length" of the near flat part of the T-o-P curve (Figure 4). Let us define this measure as *near flat distance*, or *d*, able to quantify the optimal population range where a variation does not lead to a relevant decrease of the system performance. Moreover, it gives a measure of the system flexibility in terms of population choice. The higher the *d*, the more values of population level leading to an optimal throughput can be identified. This can be normalized on the total length of the curve d_{tot}. d_{tot} corresponds to the total buffer capacity of the loop plus the number of machines in the loop. This sum represents the total amount of pallets that can be present in the loop. In this way, a general parameter not depending on the system in consideration is obtained. This is called d_r. In fact, the maximum amount of pallets that can circulate in the system depends on the system design. As a result, with this formulation, different systems can be compared (Equation 1).

$$d_r = \frac{d}{d_{tot}} \tag{1}$$

Nevertheless, just this measure is not enough to describe the NF of the T-o-P curve. Indeed, in addition, the flexibility can be seen as "low sensitivity" of the system in respect to the population. This means that a system is flexible when is less sensitive, in terms of throughput, to population changes. The example shown in Figure 5 explains this further definition of flexibility.
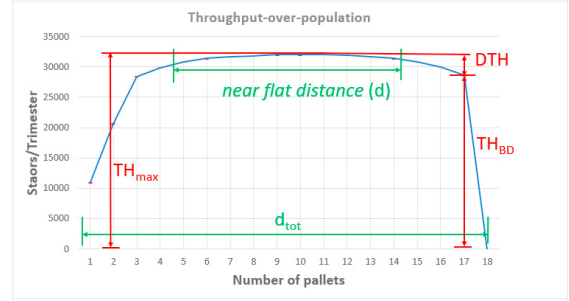
The two curves have the same *d*. However, one appears to have a higher NF than the other. This means that "Curve 1" is more sensitive to changes in population value. In "Curve 1", when the population changes significantly, the potential lost production in respect to the maximum possible throughput achievable is higher. For this reason, another indicator to describe the NF of the system (and, consequently, its flexibility) is required. Hence, another important characteristic of the T-o-P curve is the curvature through the entire range. To describe this curvature, an important indicator is the difference between the maximum throughput (TH_{max}) and its last value before the deadlock of the system due to total blockage of the loop (TH_{BD}). The difference between the two is denoted by DTH. These indicators are visible in Figure 4. Moreover, DTH can be normalized on the maximum throughput (Equation 2). The normalized value is indicated with the symbol DTH_r. If a variation of the population is executed, this indicator measures the lost production in respect to the maximum possible throughput achievable.
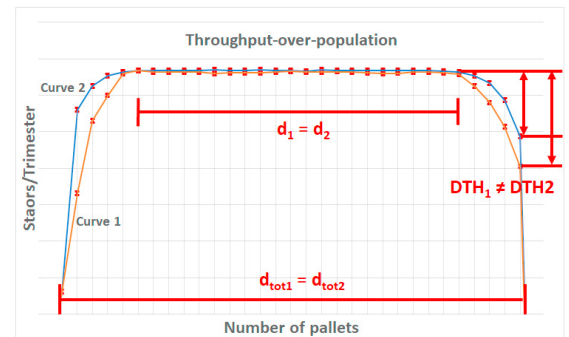

Fig. 5. Comparison about two T-o-P curves with same *near flat distance*.

$$DTH_r = \frac{DTH}{TH_{max}} = \frac{TH_{max} - TH_{BD}}{TH_{max}} \tag{2}$$

Finally, the two indicators can be coupled considering a unique coefficient that completes the framework. The aim is to obtain a single indicator able to estimate the NF. This indicator is denoted by γ from now on, defined as *NF indicator*. It is

equal to the ratio between the two previous indicators, according to Equation 3.

$$\gamma = \frac{DTH_r}{d_r} \qquad (3)$$

γ can be considered an estimator of the NF of the T-o-P curve. In fact, its value is lower in case either DTH_r is low or d_r is high, which indicates a near flat behaviour of the curve. This implicates that a low value of γ is associated with a wide "optimal population range" and/or with a low cost in terms of throughput, in case a variation of the population is applied. This means that the system throughput has a low sensitivity on the population choice, that corresponds to a high NF situation. On the other hand, γ increases as DTH_r increases or d_r decreases: the more γ increases, the less near-flatness is exhibited by the T-o-P curve. This means that the system throughput has a high sensitivity about the population choice, that corresponds to a low NF situation.

3. Near Flatness Analysis for Multiple CLMS

The dependency of system throughput on loop population is studied in this section. In this section, the analysis are performed investigating the relationship of the NF in respect to specific parameters in multiple CLMS. The focus is moved on the following parameters:

- **Efficiency in isolation** of the unreliable machines: it is observed whether and how a variation in the efficiency of the internal loop-machines may lead to generating NF in a multiple CLMS, with [9]:

$$\eta = Eff_{InIsolation} = \frac{MTTF}{MTTR + MTTF} = \frac{MTTF}{MTBF} \qquad (4)$$

- **Processing time balancing**: it is analysed whether and how an unbalancing in the internal loop machine processing times in a multiple CLMS may lead to NF situations.

Furthermore, the results of the analysis for the abovementioned parameters are reported (Section 3.3 and Section 3.4). The analysis are carried out through discrete event simulation, under the assumptions presented in Section 3.2. In this case, they are firstly performed on a "simplified system" (described in Section 3.1), to demonstrate the existence, or not, of the relations among NF and the factors studied and then the same analysis are executed on the real industrial case (presented in Section 2.1).

3.1. Simplified System Model

The simplified system model (Figure 6) used to perform the analysis is derived from a reduced version of the real system (Figure 1). The main loop structure is simplified. Just one machine upstream and one downstream the central loop are kept. The section of the line upstream the central loop and also the

one downstream the central loop are not totally removed from the model. On the contrary, in their place, two buffers of infinite capacity (B1 and B9) are positioned to simulate the two sections. All the other buffers have finite capacity. Furthermore, the time spent by each pallet in these sections is considered. In the simulation model, it is considered as transportation time in the two conveyors. The two time lengths are estimated by data fitting. The deterministic processing times of the machines out of the loop are fixed at 50 seconds for each considered case while the processing times of the machines inside the loop is 50 seconds, but this has to be multiply by 4 for a total of 200 seconds. Indeed each operation is repeated 4 times. The slowest machines are positioned inside the central loop, as in the real system.
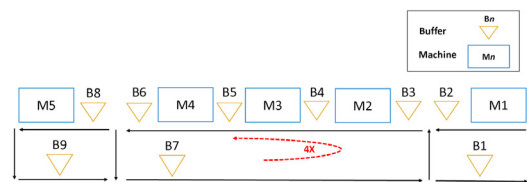

Fig. 6. Simplified System scheme, modelling the real production system.

3.2. Simulation Models Assumptions

In both simulation models (simplified system and real case) the processing times analysed are all deterministic, which is an appropriate assumption for automated lines, widely applied in industry. Blocking After Service (BAS) policy is applied. Transportation time is already included in the processing times. Machines are unreliable. The "Time To Failure" (*TTF*) of the assigned failures follows a Weibull distribution with shape factor equal to 1.5 and variable scale factor. The shape factor is selected higher than 1, because in this case the failure rate increases with time, as in an aging process: machines that are more likely to fail as time goes on [12]. The aging process is considered a realistic assumption for an industrial system, so it is considered also in the simulation models. The same value of shape factor is utilized also for the "Time To Repair" (*TTR*) of the assigned failures, again with variable scale factor. *MTTR* and *MTTF* correspond to the two expected values of the Weibull distributions used for the *TTR* and *TTF*. Given the presence of stochastic parameters as *TTR* and *TTF*, both simulation models are stochastic.

Processing times of the machines and buffer capacities can be varied depending on the considered experiment, and the same holds for the failure data (*MTTR*, *MTTF* and efficiency). All the experiments are performed with a simulation length of 90 days (about three months of production time). This length is obtained considering 7.5 hours per shift, and three shifts per day. The same warm-up period is imposed for each experiment, equal to 3 days of production period. The number of replication for each single experiment is set fixed and equal to 10. The main indicator computed in each replication is γ. It must be noticed

that γ is stochastic. Indeed, it derives from experimental results obtained through computer simulation on a stochastic system. Given its stochasticity, its average value is extracted with a confidence level of the 95% on its confidence interval.

3.3. Analysis on Efficiency of the Machines

This analysis concerns the influence of the efficiency of the central loop machines on the NF phenomenon. This is first performed on the simplified system. The *MTTRs* of the central loop machines are kept constant. The value of the efficiency in isolation is the same for all the central loop machines. Thus, *MTTF* and *MTBF* are computed according to Equation 4. All the other assumptions remain unchanged from the ones reported in Section 3.1. Starting from a value of efficiency in isolation equal 0.80, this is improved gradually to 0.95 with a step of 0.05, for a total of 4 cases. In Figure 7 the 4 resulting T-o-P curves are reported. Each T-o-P curve is associated with one different value of efficiency in isolation. The T-o-P curves and the corresponding values of the NF indicator γ are displayed: the higher the efficiency, the lower the γ and the higher the NF. The reason behind this behaviour might be the influence of the number of pallets on the starvation and blocking propagation. When a machine fails, if the population is small, the starvation propagates easily in the system, from the failed machine to the other devices. In addition, if we decrease the machines efficiency and we keep constant the *MTTR*, the frequency of failures increases. Increasing failures frequency leads to increasing number of occasions where a machine is failed during the production time. If the population is small, increasing number of times where the machines are failed leads to increasing starvation propagation during the production time and consequent decreasing throughput. For this reason, for low values of population, the decreasing of the throughput due to starvation of the machines is amplified for low values of efficiency. Hence, to slow down the starvation propagation (and get to the optimal throughput), a higher number of pallets is needed with low efficiency. Thus, for low values of efficiency, the first part of the T-o-P curve has increasing curvature. On the other hand, considering the second part of the T-o-P curve, a similar situation can be noticed for the blocking propagation. When a machine fails, if the population is high, the blocking propagates easily in the system, from the failed machine to the others. Thus, if the population is high, increasing number of times where the machines are failed leads to increasing blocking propagation during the production time and consequent decreasing throughput. For this reason, for high values of population, the decreasing of the throughput due to blocking of the machines is amplified for low values of efficiency. Consequently, to slow down the blocking propagation (and get to the optimal throughput), a lower number of pallets is needed with low efficiency. Therefore, for low values of efficiency, the second part of the T-o-P curve has increasing curvature. In conclusion, for a multiple CLMS with this specific architecture, increasing central loop machines efficiency leads to increasing NF in the T-o-P curve.

The same analysis are repeated on the real case. To vary the efficiency in isolation of the central loop devices, their real fail-

Fig. 7. T-o-P curves for different efficiency in isolation values (simpl. system).

ure data are removed and substituted. Thus, the three central loop machines have an identical imposed efficiency. To each of the three, the same failure distribution is assigned, as explained in Section 3.2. Two different values of efficiency for the central loop machines selected are 0.99 and 0.95. Consequently, identical failure data are computed and assigned to each machine. All the other parameters are the same and remain unchanged. The results of this analysis are shown in Figure 8. The case with higher efficiency in isolation (0.99) has lower γ value, and, consequently, higher NF. The γ values confirm the consistency of this analysis with the ones of the simplified system case: increasing the efficiency of the central loop machines leads to lower γ values and to increasing NF.
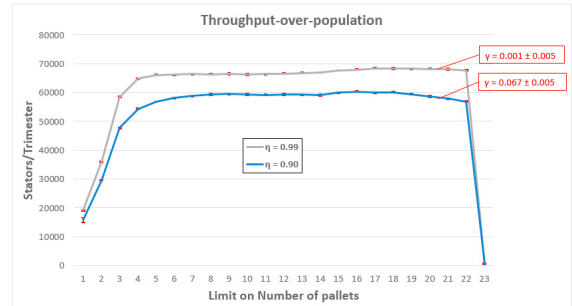

Fig. 8. T-o-P curves for different efficiency in isolation values (real system).

3.4. Analysis on Processing Time Balancing

In this section, the influence of the central loop processing time balancing on the NF phenomenon is studied. This is first performed on the simplified system described. Starting from a perfectly balanced central loop, i.e. identical processing time, *MTTF*, *MTTR* for all the three machines, two machines are gradually improved in terms of processing time reduction. The station M3 is kept as the slowest machine, with a processing time of 50 seconds per layer (a total of 200 seconds for the 4 layers). The initial processing time of all the machines is 50 seconds (a total of 200 seconds for the 4 layers), then M2 and M4 are improved with a step of 2.5 seconds until a processing time of 27.5 seconds for layer is reached (110 seconds for the 4 layers). The three central loop machines have the same reliabil-

ity: the efficiency is kept at 0.9, the *MTBF* at 3600 seconds. All the other parameters and assumptions remain unchanged.

Line balancing is measured using a coefficient named Balancing Index, whose definition is derived from Equation 5 (PT stands for Processing Time in the equation). It has unitary value for perfectly balanced system and decreases as line is unbalanced. In Figure 9, two T-o-P curves of the configurations with balancing index equal to 0.55 and 1 are displayed. The case with higher balancing index (equal to 1) has higher γ value, and, consequently, lower NF. On the other hand, in an unbalanced line, the throughput mainly depends on the performance of the slowest machine in isolation.
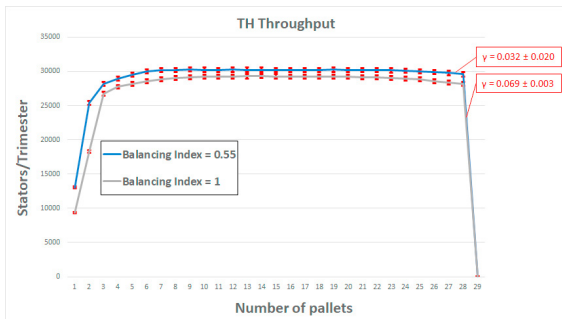

Fig. 9. T-o-P curves for different processing times balancing (simpl. system).

$$Bal_{Index} = 1 - \frac{PT_{SlowestMachine} - PT_{OtherMachines}}{PT_{SlowestMachine}} \qquad (5)$$

In a very unbalanced line with a strongly slowest machine, indeed, the upstream buffer of the slowest machine is basically always full, while its downstream buffer is basically always empty. Thus, it is possible to say that the slowest machine is less starved and less blocked than in a more balanced case. For this reason, the influence of the number of pallets on the starvation and blocking propagation to the slowest machine is decreased. Indeed, in this case, blocking and/or starvation propagation from a failed machine to the slowest one is not easily achievable. In conclusion, the throughput mainly depends on the performance of the slowest machine in isolation. This effect is lowered once a more balanced situation in terms of processing times is reached. The more the balancing grows, the more easily the blocking and the starvation can propagate in the system. Thus, to reduce the starvation propagation (and get to the optimal throughput), an increasing number of pallets is needed in the first part of the T-o-P curve. Similarly, a decreasing number of pallets is needed in the second part of the T-o-P curve to slow down the blocking propagation (and get to the optimal throughput). For this reason, the range of optimal values (where an optimal throughput is achieved) is reduced when the line balancing grows. Thus, the loop shows a stronger sensitivity to the loop population and the NF phenomenon is less relevant in more balanced cases.

The same analysis is repeated on the industrial case. The unbalancing is inserted by lowering of about the 10-20% the processing times of the two fastest machines in the internal loop

(Op. 11 and Op. 12). All the other parameters (failure data, buffer allocation, conveyor length, etc.) are the same of the original configuration and remain unchanged. This configuration is addressed as "unbalanced case" and is compared to the actual situation, the "base case". The results of this analysis are presented in Figure 10.
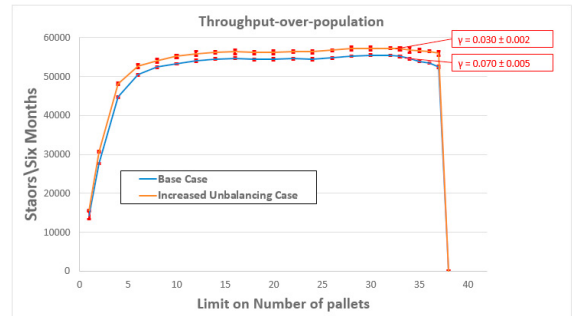

Fig. 10. T-o-P curves for different processing times balancing (real system).

The results are consistent with the ones reported for the simplified system case. In the "unbalanced case", there is a lower value of the NF indicator γ than in the "base case". This indicates increasing NF. As a conclusion, the impact of the processing time unbalancing (of the central loop machines) on the NF is confirmed also in a real production system. For a multiple CLMS with this architecture, the more unbalanced the central loop is in terms of processing times, the stronger the NF is.

4. Practical Guidelines

In this section, practical guidelines for multiple CLMS handling are presented. They are originated from the experimental results obtained in Section 3.3 and Section 3.4. Known this, it is possible to define ways to handle eventual dependencies, to better control and manage multiple CLMS, since understanding how to manage the NF phenomenon means understanding how to manage the system flexibility in terms of population choice.

These guidelines are referred to multiple CLMS with deterministic processing times, finite buffers and unreliable machines. From now on, the term "machines" is referred just to the central loop devices. Further assumptions, specific for each guideline, are reported in the respective guideline. The practical guidelines presented are the following:

1. When designing a new system, if all the machines have the same failure parameters and high efficiency in isolation, additional effort for finding the exact optimal population value may be avoided: the population choice is not so relevant to obtain an optimal (or near optimal) system performance. On the other hand, if their efficiency in isolation is low, the population choice becomes relevant for the system performance and for this reason additional effort for finding the exact optimal population value is required.

2. If all the machines have the same failure parameters, increasing the efficiency of all the devices leads to higher

throughput and to a system less sensitive to the population choice. Thus, there is a double advantage: higher performance are obtained and flexibility, in terms of population choice, is enhanced.

3. When designing a new system, if a strongly slowest machine is present in the system, additional effort for finding the exact optimal population value may be avoided.: the population choice is not so relevant to obtain an optimal (or near optimal) system performance. On the other hand, it is balanced in terms of processing time, the population choice becomes relevant for the system performance and for this reason additional effort for finding the exact optimal population value is required.

4. When a strongly slowest machine is present in the system, lowering its processing time improves the throughput. However, this also leads to further analysis to identify the exact optimal population value, since population choice becomes more relevant to obtain an optimal (or near-optimal) throughput.

5. When a strongly slowest machine is present in the system, if the processing time of the other machines is lowered, the system becomes less sensitive to the population value. Thus, the flexibility, in terms of population choice, is enhanced.

5. Conclusions and Further Developments

In this work, a phenomenon called *Near Flatness* affecting the T-o-P curve in multiple CLMS, is investigated. A quantitative definition of the NF of the T-o-P curve is introduced by means of an indicator γ. In addition, a deep analysis to evaluate the NF dependency on system parameters is performed. Efficiency in isolation of the unreliable machines [9] and processing time balancing of the machines are discovered to be leading factors to NF. Finally, starting from this analysis, some guidelines for multiple CLMS handling to better control and manage these systems are presented.

To extend this study, a more in-depth analysis of the NF phenomenon can be carried out. An interesting study could be performed repeating the same experiments but with different assumptions and/or different parameters as potential leading factor to NF, to obtain new practical guidelines. Different guidelines can be also extracted utilizing a new and different definition of the NF indicator γ. Another challenging topic might be the identification of an analytical correlation between the change in T-o-P curve shape and the change in the efficiency values or the degree of unbalancing. Finally, significant results might be extracted repeating the same studies on multiple CLMS but with different architecture.

Acknowledgements

References

[1] Ferreira, L.P., Gómez, E.A., Lourido, G.C., Quintas, J.D., Tjahjono, B., 2012. Analysis and optimisation of a network of closed-loop automobile assembly line using simulation. The International Journal of Advanced Manufacturing Technology 59, 351–366.

[2] Ferreira, L.P., Gomez, E.A., Lourido, G.P., Salgado, M., Quintas, J.D., 2011. Analysis on the influence of the number of pallets circulating on an automobile closed-loops assembly line. manufacturing systems 14, 15.

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

[4] Gershwin, S.B., Werner, L.M., 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, 3085–3111.

[5] Ishigami, T., Tanaka, Y., Homma, H., 2014. Development of motor stator with rectangular-wire lap winding and an automatic process for its production. Electrical Engineering in Japan 187, 51–59.

[6] Ishigami, T., Tanaka, Y., Homma, H., 2015. Motor stator with thick rectangular wire lap winding for hevs. IEEE Transactions on Industry Applications 51, 2917–2923.

[7] Kim, S.H., 2017. Electric motor control: Dc, ac, and bldc motors. Elsevier.

[8] Levantesi, R., 2001. Analysis of multiple loop assembly/disassembly networks.

[9] Li, J., Meerkov, S.M., 2008. Production systems engineering. Springer Science & Business Media.

[10] Li, N., Zheng, L., Li, Q.L., 2007. Semiconductor system with multiple closed-loops constrains, in: Automation Science and Engineering, 2007. CASE 2007. IEEE International Conference on, IEEE. pp. 484–488.

[11] Li, N., Zheng, L., Li, Q.L., 2009. Performance analysis of two-loop closed production systems. Computers & Operations Research 36, 119–134.

[12] Pham, H., Lai, C.D., 2007. On recent generalizations of the weibull distribution. IEEE transactions on reliability 56, 454–458.

[13] Resano Lazaro, A., Luis Perez, C., 2009. Dynamic analysis of an automobile assembly line considering starving and blocking. Robotics and Computer-Integrated Manufacturing 25, 271–279.

[14] Zhang, Z., et al., 2006. Analysis and design of manufacturing systems with multiple-loop structures. Ph.D. thesis. Massachusetts Institute of Technology.

[15] Zhou, Y., Wang, M., Hao, H., Johnson, L., Wang, H., 2015. Plug-in electric vehicle market penetration and incentives: a global review. Mitigation and Adaptation Strategies for Global Change 20, 777–795.