Integrated quality and production logistics modelling for the design of selective and adaptive assembly systems

Marcello Colledani [a,*], Dariush Ebrahimi [a], Tullio Tolio (1)[a,b]

[a] Politecnico di Milano, Department of Mechanical Engineering, Via la Masa 1, 20156 Milan, Italy
[b] ITIA-CNR, Institute of Industrial Technologies and Automation, Via Bassini 15, 20133 Milan, Italy

Keywords:
Assembly
Performance
Electric drives

1. Introduction, motivation and objectives

The current global manufacturing context is posing serious challenges on manufactures due to high customer expectations on product quality and delivery reliability, while companies have to continuously reduce costs. To cope with these challenges, production processes and systems need to be designed and operated close to their technological limits, by simultaneously paying attention to quality and productivity requirements, in accordance to the "zero-defect" manufacturing paradigm.

With respect to this background, the concept of selective assembly represents a formal approach to obtain high precision assemblies from relatively low precision manufactured parts. Therefore, this approach allows either to overcome technological limitation in very precise assemblies (e.g. micro-assembly) or to reduce the cost of the components for a given required precision of the final assembly. Selective assembly consists in performing in-line component inspection and in partitioning the components into quality classes, depending on the specific outcome of the measurement process. Then, only matching components extracted from compliant quality classes are assembled. This strategy increases the assembly yield, but at the same time it increases the operational complexity of the system. As a matter of fact, the selective assembly strategy translates a product quality issue into a system logistics issue. In the past, this strategy was applied mostly to traditional sectors, such as mechanical components production. One classical example is piston and cylinder assembly where the tolerance on the clearance between the two components is narrower than the dimensional variability of the two components. Due to the increasing pressure on high precision manufacturing and thanks to the development of advanced and fast measurement technologies supporting on-line applications and 100% inspection [1], selective assembly systems have attracted increasing interest in the last five years, especially in fast growing sectors such as micro-production [2], renewable energy equipment production, in the e-mobility sector (e.g. electrical drives, batteries assembly [3]), and in the automotive body assembly [4]. In the literature, the performance of selective assembly systems has been addressed with quality-oriented approaches by mainly focusing on the effect of the component sorting policy on the assembled product quality [5,6]. Other studies investigate the effects of process adaptation on the performance of selective assembly systems [7]. Process adaptation means that the nominal value of the key quality characteristics of the component produced with the more capable production process can be adjusted according to pre-defined states of the system, in order to increase the component matching. Recently, simulation has been used for predicting the impact of specific adaptation policies on system performance [8]. While all these works are important to shed light on the selective assembly process they neglect important production logistics features of the system, such as finite capacity buffers and unreliable machines which, in turn, create significant phenomena such as the arising of deadlock states in the system. As a matter of fact, a methodology to support the design of selective and adaptive assembly systems that integrates quality and production logistics implications has never been proposed, reducing the potential benefits of these systems in industry. For example, important questions like "What is the impact of the number of quality classes and of the process adaptation on the throughput of conforming assembled products?" remain unsolved, resulting in sub-performing system configurations. To overcome these limitations, in this paper an integrated model of selective and adaptive assembly systems and a new method for the prediction of the throughput of conforming products, the system yield and the WIP (Work in Progress) in these systems are developed for the first time. This approach allows deriving new insights on the implementation of selective assembly systems in real settings.

---

[*] Corresponding author.

2. System description

The considered selective assembly system model is represented in Fig. 1. Although this system model can be integrated into models of longer process-chains, in this paper we will focus on the selective assembly cell. Specifically, we consider a selective assembly system where two sub-components, namely x and y, are assembled. Extensions to higher number of sub-components can be easily managed within the same framework. The sub-components are respectively processed by machines M_x and M_y (blue squares). After the process, components x and y are respectively inspected in-line by machines I_x and I_y (red squares) and sorted into buffers (yellow circles) by splitting stations S_x and S_y (white switches), according to the measured key quality characteristic value. A total number of F buffers or bins are defined for both x and y, namely $B_{x,i}$ and $B_{y,i}$, with $i = 1, \ldots, F$. The capacity of each buffer is respectively $N_{x,i}$ and $N_{y,i}$. A subcomponent of type x is then assembled with one matching subcomponent of type y by the assembly station M_a (light blue rectangle). The main performance measures of interest are:

- Average total production rate of the system, E^{Tot}.
- Average effective production rate, E^{Eff}, of conforming parts.
- System yield, Y^{system}, that is the fraction of conforming assembled products produced by the system (E^{Eff}/E^{Tot}).
- Yield of class i, Y_i, that is the fraction of conforming parts obtained by coupling components from the quality class $i = 1, \ldots, F$.
- *WIP*, that is the total average inventory of the system.

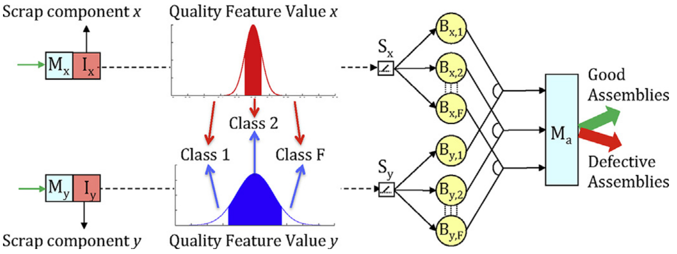


Fig. 1. Representation of the analyzed selective assembly system.

3. Single stage models

3.1. Component processing, inspection and sorting stations

Without loss of generality, we focus on component x. Stage M_x is unreliable. It fails with probability $p_x = 1/MTTF_x$ (Mean Time to Failure) and is repaired with probability $r_x = 1/MTTR_x$ (Mean Time to Repair). The in-line inspection station I_x is considered to be perfectly accurate. Both I_x and S_x are considered to be reliable. From a material flow point of view, the splitting station S_x sorts the incoming material flow into F output flows, based on the product quality feature measurement. The splitting fractions are not known a priory but depend on the component quality feature distribution and on the specific sorting policy.

Component quality. M_x produces parts with key quality characteristic values distributed according to a probability density function f_x and the cumulative distribution function F_x. The mean of the distribution is μ_x and the standard deviation is σ_x. Specification limits are imposed by design on x. They are defined as Upper Specification Limit (USL_x) and Lower Specification Limit (LSL_x). A component is considered to be defective and is scrapped by I_x if its quality characteristic is outside these limits. The fraction of defective component is γ_x.

Sorting policy. Each buffer $B_{x,i}$ contains components with key quality feature x included between a lower limit $l_{x,i}$ and an upper limit $L_{x,i}$. The quality classes are contiguous, i.e. they respect the following properties: $L_{x,i} = l_{x,i+1}, \forall i = 1, \ldots, F - 1$; $L_{x,F} = USL_x$; and $l_{x,1} = LSL_x$.

The probability of sorting a part into $B_{x,i}$ after inspection is:

$$\alpha_{x,i} = \int_{l_{x,i}}^{L_{x,i}} F_x(s)ds \tag{1}$$

3.2. Assembly station

The assembly station M_a assembles one component x with one component y to meet a desired key quality feature value of the assembled product. It is unreliable with parameters p_a and r_a.

Assembled product quality. The key quality characteristics z of the assembled product is expressed in form of a function $z = g(x, y)$ of the components' key quality feature values x and y. Therefore, the probability density function f_z and the cumulative distribution function F_z can be obtained from the cumulative distribution functions of x and y. Lower (LSL_z) and Upper (USL_z) Specification Limits are imposed by design on the key quality feature z of the assembled product (i.e. gap tolerance), depending on the specifications imposed on the final product. According to these tolerances, every assembly flow has a specific fraction of non-conforming parts associated, namely $\gamma_{a,i}$, with $i = 1, \ldots, F$. If the number of classes is sufficiently high, $\gamma_{a,i} = 0$ is theoretically possible. However, practitioners tend to limit this number.

Assembly policy. M_a can only assemble parts in $B_{x,i}$ with compliant parts in $B_{y,i}$, with $i = 1 \ldots, F$. The assembly machine selects the quality classes to match according to a probabilistic rule. If all the upstream buffers are not empty, then the assembly machine selects buffer i with fixed probability $\alpha_{a,i}$. If one or more upstream buffer is empty, the selection probability is scaled according to the available components. If all the upstream buffers are empty, the assembly machine cannot process parts and is starved.

4. Deadlock states

Since buffers have finite capacity, a deadlock state is observed if all the following conditions hold:

- Machine M_x cannot deposit component x of quality class i in buffer $B_{x,i}$ since it is full;
- Machine M_y cannot deposit component y of quality class j in buffer $B_{y,j}$ since it is full;
- The assembly machine M_a cannot assemble parts from any couple of upstream buffers.

More formally, if the buffer levels are described by the vector $\mathbf{n} = (n_{x,1}, n_{x,2}, \ldots, n_{x,F}, n_{y,1}, n_{y,2}, \ldots, n_{y,F})$, a deadlock state is every state undergoing the following condition:

$$(n_{x,i} = N_{x,i} \wedge n_{y,i} = 0) \vee (n_{x,i} = 0 \wedge n_{y,i} = N_{y,i}) \quad \forall i = 1, \ldots, F \tag{2}$$

Deadlock avoidance policies have been introduced for these systems. Specifically, two strategies are possible:

- *Ignore*: allocate sufficient space to accommodate all the parts.
- *Discard*: scrap incoming components that cannot be accommodated in the selected full buffer.

In presence of machines' failures, the ignore policy entails unlimited buffers. Since unlimited buffers cannot be introduced in real systems, only the "*discard*" policy is considered in this paper.

5. Process adaptation

To decrease part scrapping due to the discard policy, a shift in the mean of the higher capability component, let say x, can be implemented, for each specific flow-dependent blocking condition. We assume the shift requires negligible set-up times (only process target adjustments). A shift τ modifies the process mean level $\mu_x(\tau) = \mu_x + \delta(\tau)$. Therefore, it affects the probability of sorting the component x in the downstream bins as follows:

$$\alpha_{x(\tau),i} = \int_{l_{x,i}}^{L_{x,i}} F_{x(\tau)}(s)ds \quad \forall i = 1, \ldots, F \tag{3}$$

where $F_{x(\tau)}$ is the cumulative distribution function with shifted mean $\mu_{x(\tau)}$. The fraction of components x in quality class i produced under the target process mean $\mu_{x(\tau)}$ is denoted as $\theta_{x(\tau),i}$. Since it involves a change in the quality feature distribution of one component, the process adaptation modifies the assembly yield. Under the process shift τ, the adjusted fraction of non-conforming assembled products for the flow i is $\gamma_{a,i}(\tau)$.

6. Analytical method for system performance evaluation

The system performance is estimated with an analytical method based on a two-level decomposition approach, inspired by the method proposed in [9] and further developed in [10]. One level of analysis, the *Machine Level Decomposition* (MLD), is based on the evaluation of the probability of all the states of each single machine, also taking into account the influence of its neighbouring buffers. With this approach, it is possible to estimate the portion of the working time each machine dedicates to each specific activity. Starting from this analysis, in the second level of decomposition, the *Buffer Level Decomposition* (BLD), the goal is to approximate the flow of material through each buffer of the original line. The BLD decomposes the system into a set of $2F$ two-machine one-buffer sub-systems, F for component x, $ss_x(i)$, and F for y, $ss_y(i)$, i.e. one for each buffer in the system. The performance of each sub-system in terms of production rate, $E_x(i)$ and $E_y(i)$, and average inventory, $WIP_x(i)$ and $WIP_y(i)$, is evaluated with the exact method developed in [11]. By studying alternately the BLD and the MLD and by using the results obtained in one level as input for the other level, it is possible to evaluate the performance of the original system, once convergence conditions are met. The system performance can be calculated as:

$$Y_i = (1-\gamma_{a,i}) \left. 1-\sum_{\tau} \vartheta_{x(\tau),i} \right) + \sum_{\tau} \vartheta_{x(\tau),i}(1-\gamma_{a,i}(\tau))$$

$$E^{Tot} = \sum_{i=1}^{F} E_x(i), \quad E^{Eff} = \sum_{i=1}^{F} E_x(i)\cdot Y_i, \quad WIP = \sum_{i=1}^{F} WIP_x(i) + WIP_y(i) \tag{4}$$

The accuracy of this method was validated against simulation on a wide set of 80 test systems with randomly generated data. The maximum error on the E^{Eff} is 2.8%, while it is 3.89% on the *WIP*.

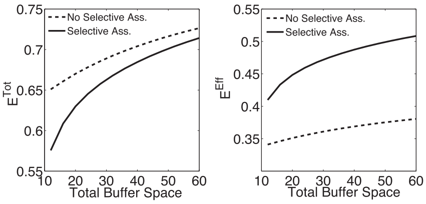
7. Numerical results and system behaviour

In the first experiment, selective assembly is compared with the case where selective assembly is not adopted, i.e. $F = 1$, under variation of the total buffer space. The parameters of the tested system, based on a real cylinder and piston assembly case [7], are reported in Table 1. The quality characteristics of the assembled product is the gap between components x and y, i.e. $z = x - y$. No process adaptation policy is applied. Since the reliability parameters are identical, the total buffer space is equally distributed between the buffers in the system. The results are reported in Fig. 2. The following considerations hold:

- Selective assembly negatively affects the total production rate of the system, E^{tot}, since non-matching components are discarded to avoid deadlocks (Fig. 2a).
- Selective assembly improves the system yield due to the pre-sorting of the components into classes. As a consequence, it increases the effective throughput E^{Eff} (Fig. 2b).

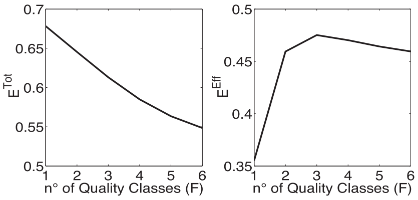
Table 1
Data of the experimental analysis [7].

M_x	M_y	M_a
$x \sim N(\mu_x, \sigma_x^2)$	$y \sim N(\mu_y, \sigma_y^2)$	$z = x - y; z \sim N(\mu_x - \mu_y, \sigma_x^2 + \sigma_y^2)$
$\mu_x = 4$; $\sigma_x = 0.116$	$\mu_y = 3.3$; $\sigma_y = 0.05$	$\mu_z = 0.7$; $\sigma_z = 0.126$
$LSL_x = 3.5$; $USL_x = 4.5$	$LSL_y = 2.8$; $USL_y = 3.8$	$LSL_z = 0.61$; $USL_z = 0.79$
$\alpha_1^x = 0.5$; $\alpha_2^x = 0.5$	$\alpha_1^y = 0.5$; $\alpha_2^y = 0.5$	$\alpha_1^a = 0.5$; $\alpha_2^a = 0.5$
$\gamma_x \approx 0$	$\gamma_y \approx 0$	$\gamma \approx 0.4761$ if $F = 1$
		$\gamma_1 = \gamma_2 \approx 0.2881$ if $F = 2$
$p_x = 0.01$; $r_x = 0.05$	$p_y = 0.01$; $r_y = 0.05$	$p_a = 0.01$; $r_a = 0.05$



(a) (b)

Fig. 2. Comparison between selective and traditional assembly.



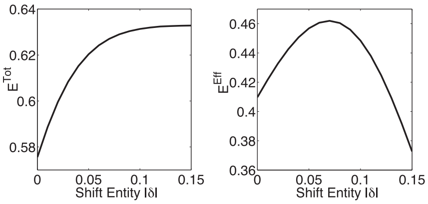
(a) (b)

Fig. 3. Effect of the number of quality classes.

In the second experiment, the impact of the number of quality classes is investigated. The same data as the previous experiment are considered. The equal probability binning policy is assumed, i.e. the quality class limits are set in order to uniformly fill the quality classes. A fixed total buffer capacity of 24 is considered, equally spread in the available buffers. Results are reported in Fig. 3. The following considerations hold:

- As the number of quality classes increases, the total throughput decreases (Fig. 3a). Indeed, for each quality class the buffer size is smaller and more parts are discarded to avoid deadlocks.
- However, since the limits of each quality class get closer, the control of the quality feature of the assembled product improves with F, thus the system yield increases.
- As a result, the effective throughput curve is concave (Fig. 3b), with a maximum for a specific number of classes, i.e. $F = 3$.

In the third experiment, the impact of process adaptation on the system effective throughput is analyzed. The buffer sizes are all set to 3. If $B_{y,1}$ is full, the target mean of the more capable process M_y is shifted to $\mu_y(\tau) = \mu_y + \delta$, while, if $B_{y,2}$ is full, the target mean of M_y is shifted to $\mu_y(\tau) = \mu_y - \delta$. Results are reported in Fig. 4. The following considerations hold:



(a) (b)

Fig. 4. Effect of the process adaptation on the system performance.

- The shift has a double effect on the performance of the system. On the one hand, it increases the total production rate of the system, since it decreases the probability of producing a part in quality class i when the buffer $B_{y,i}$ is full (Fig. 4a). On the other hand, by increasing the shift, the yield increases up to a certain level, and then it decreases.
- As a result of these two effects, the effective production rate is a concave function that is maximized for a certain value of the shift entity $\delta = 0.07$ (Fig. 4b). For this level, the effective throughput of the system (0.464) is higher then the effective throughput without process adaptation (0.41), i.e. +13%.
- The yield is maximized for $\delta = 0.056$. For this value of the shift, the resulting effective throughput of the system is 0.455. This means that traditional quality-oriented methods that neglect the interactions between quality and production logistics, only provide sub-performing configurations of the system (−2%).

8. Real case study

The proposed approach has been applied to the rotor assembly line of the electric drive assembly plant in Bosch [12]. Cars are with about 12 million manufactured units (EU27) one of the most important products in Europe. Due to the threatening lack of petrol and efforts towards sustainable development, the shift from the combustion engine towards electrical drives in cars is on going. For example, in Germany the ambitious goal of 1 million sold electrical vehicles by 2020 has been stated. The current manufacturing processes for producing electrical drives must be improved to support the achievement of this goal.

The line under analysis is composed of seven main stages. Stage M_1 loads the laminated steel stacks, which are rings of 22.6 cm diameter, on the pallet; stage M_2 assembles the magnets on the stacks' external ring. Stage M_3 is the stack magnetization process where the total magnetic flux intensity is also measured. Stage M_4 is a heating station that prepares the magnetized stacks for the rotor assembly process at M_5. At this stage, the stacks are axially assembled on the central shaft of the rotor, currently under a first-in-first-out loading policy. Stages M_6 and M_7 respectively perform balancing and marking operations. After assembling the rotor and the stator, the completed motor undergoes the EOL (End of Line) inspection. Motor characteristics as well as specifications such as torque, and speed are tested. However, the company's strategic goal is to move the inspection and quality control processes upstream in the line. The main objective is to improve the quality of the rotor assembly process, by exploring the possibility of implementing selective assembly, thus saving downstream production capacity currently wasted in processing rotors that are already defective.

The rotor assembly line under the current configuration, where selective assembly is not implemented, has been modelled. The operational and down times were provided by the company. The performance of the system was evaluated and the results were validated through comparison with historical production data. Table 2 shows that the model well captures the dynamics of the system in the current configuration (CI is the confidence interval).

Table 2
Performance of the assembly line without selective assembly (the time unit is omitted for confidentiality reasons).

Perf. measure	Model	Production data			Err. (%)
		95% LCI	Mean	95% UCI	
E^{Tot}	0.536	0.534	0.538	0.542	0.37
E^{Eff}	0.508	0.507	0.511	0.515	0.59
Y^{System}	0.948	0.947	0.95	0.953	0.21
WIP	18.6	17.8	18.2	18.6	2.19
Scrap rate	0.052	0.047	0.05	0.053	4

The approach proposed in this paper has been applied to analyze the potential impact of selective assembly on the overall quality and production logistics performance of the rotor assembly line. Specifically, the company idea was to exploit the availability of the measurements of the total magnetic flux intensity at stack level, performed at M_3, to cluster the stacks and assemble only compliant stacks at M_5. With this approach, weak magnetic field stacks can be coupled with strong magnetic field stacks resulting in a reduced variability of the magnetic field intensity at rotor level. Adaptation policies are not applicable since the company currently sees the magnetization stage as a black-box process. Following the company requirements, the total magnetic flux intensity measured over a large set of stacks has been used to characterize the probability density function of the component key quality characteristics. The sum of the total fluxes has been considered as the quality characteristics of the assembled product. Four different reconfigurations of the line have been tested, with increasing number of quality classes varying from 2 to 8. The total buffer size after stage M_3 has been kept identical to the current configuration (40 stacks). Results are reported in Table 3. The effective throughput is maximized with 2 quality classes. In this configuration, only the 0.3% of rotors is defective against the 5% of the current configuration (a reduction of defective rotors of 94%). Moreover, the total fraction of scrapped parts, including both discarded stacks and scrapped defective rotors, is globally reduced of a 40%. The effective throughput is increased (+2.5%) and the WIP of the system is reduced (−20%) with respect to the current configuration ($F = 1$). Since recovering value from scrapped stacks is easier than from scrapped rotors, this solution entails a large potential economic benefit, moving towards zero defect manufacturing solutions. Therefore, an overall benefit on both the quality and the production logistics performance of the system has been achieved.

Table 3
Rotor assembly line performance as a function of the number of quality classes (the time unit is omitted for confidentiality reasons).

F	E^{tot}	Y^{System}	E^{Eff}	WIP	Stack scrap rate	Rotor scrap rate	Total scrap rate
1	0.536	0.948	0.508	18.6	0	0.052	0.052
2	0.521	0.997	0.519	15	0.028	0.003	0.031
4	0.512	0.9996	0.512	17.4	0.045	0	0.045
6	0.498	0.9998	0.498	14.6	0.07	0	0.07
8	0.487	0.9999	0.487	15.4	0.091	0	0.091

9. Conclusions

This paper proposes an integrated modelling framework for supporting the design of selective and adaptive assembly systems. Results prove that, if properly designed, selective assembly can significantly reduce the output fraction of defective assemblies. Traditional quality-driven selective assembly designs that neglect the finite storage capacity fail in this goal as sub-performing solutions are provided. Future research will investigate new approaches to jointly optimize buffers, classes, process shifts and tolerances. Given the industrial benefits of this approach, in view of its full implementation in high-tech industry, in-line inspection technologies and smart manufacturing solutions development should be boosted together with appropriate system design and management policies.

Acknowledgements

The authors would like to thank Eng. Jan Aichele from Bosch GmbH and Eng. Daniel Coupek from ISW for the support in this research.

References

[1] Maropoulos P-G, Ceglarek D (2010) Design Verification and Validation in Product Lifecycle. *CIRP Annals – Manufacturing Technology* 59(2):740–759.
[2] Löchte C, Kayasa J, Herrmann C, Raatz A (2012) Methods for Implementing Compensation Strategies in Micro Production Systems Supported by a Simulation Approach. *Precision Assembly Technologies and Systems* 371:118–125.
[3] Yang S, Wang H, Hu S-J, Lin Y (2013) Modeling Assembly Systems with Repetitive Operations. *CIRP Annals – Manufacturing Technology* 62/1:5–8.
[4] Ceglarek D, Huang W (2007) Mode-based Decomposition of Part Form Error by Discrete-Cosine Transform with Implementation to Assembly and Stamping System with Compliant Parts. *CIRP Annals – Manufacturing Technology* 51(1):21–26.
[5] Arai T (1992) A Simulation System on Assembly Accuracy. *CIRP Annals – Manufacturing Technology* 41(1):37–40.
[6] Mease D-A, Sudjinato V, Nair N (2004) Selective Assembly in Manufacturing: Statistical Issues and Optimal Binning Strategies. *Technometrics* 46(2):165–175.
[7] Matsuura S, Shinozaki N (2011) Optimal Process Design in Selective Assembly when Components with Smaller Variance are Manufactured at Three Shifted Means. *International Journal of Production Research* 49(3):869–882.
[8] Kayasa J, Herrmann C (2012) A Simulation Based Evaluation of Selective and Adaptive Production Systems (SAPS) Supported by Quality Strategies. *Procedia CIRP* 3:14–19.
[9] Colledani M, Tolio T (2005) A Decomposition Method to Support the Configuration/Reconfiguration of Production Systems. *CIRP Annals – Manufacturing Technology* 54(1):441–444.
[10] Colledani M (2013) Integrated Quality and Production Logistics Analysis of Selective Assembly Systems. *11th A.I.Te.M. Congress*, Italy.
[11] Tolio T, Gershwin S-B, Matta A (2002) Analysis of Two-Machine Lines with Multiple Failure Modes. *IIE Transaction* 34(1):51–62.
[12] EU, FP7, FoF.NMP. 2011-5 (2011) *MuProD-Innovative Proactive Quality Control System for In-process Multi-stage Defect Reduction – Grant Agreement n° 285075.*