

Technical Paper

Quality control planning to prevent excessive scrap production



B. Bettayeb, S.J. Bassetto*, M. Sahnoun

Department of Mathematical and Industrial Engineering, École Polytechnique de Montréal, C.P. 6079, succ. Centre-Ville, Montréal, Quebec, Canada H3C3A7

ARTICLE INFO

Article history:

Received 8 January 2013

Received in revised form

14 November 2013

Accepted 23 January 2014

Available online 25 March 2014

Keywords:

Quality control planning

Risk exposure

Capacity allocation

Scraps products

ABSTRACT

This paper presents a risk-based approach for quality control planning of complex discrete manufacturing processes, to prevent massive scraps to occur. An analytical model is developed to optimize the quality control plan (QCP) subject to inspection capacity limitation and risk exposure objectives. The problem is then formulated as a constrained capacity allocation problem. A dedicated heuristic that solves a simplified instance of an industrial case study, from semiconductor manufacturing, is presented to provide insights into the applicability and the operational use of the approach and its potential gains in terms of risk exposure reduction. The main advancement resulting from this work is the proposal of a model of quality control allocation and an understandable algorithm to prevent the production of excessive amounts of scrap. The industrial illustration shows a decrease in potential losses by a factor of 3.

© 2014 The Society of Manufacturing Engineers. Published by Elsevier Ltd. All rights reserved.

1. Introduction

In mass production industries, like semiconductor manufacturing, the concept of massive scraps is of prime importance. Undetected defects could affect thousands, if not millions, of finished or semi-finished products. Tardy detected defects or failures often lead to product recalls, returns or massive scraps, which are nightmare for industrialists and marketing managers [13]. These catastrophic events are not well publicized as they generate losses due to: re-manufacturing costs, logistics costs, systematic shrinking of their market share, and a severe damage to their image. The consequences for customers concerned can also be catastrophic, ranging from product shortage, injury or death (in case of critical components of health care devices like peacemaker for instance) [11]. Almost each major event, like massive scraps or equipment breakdown, have different origins. However, they share a common characteristic: their causes, even known, have not been detected by the control system and the failure whatever its origins, has affected a lot of products before being detected. In case of massive scraps, quality control plan fails its mission.

This paper acts on the planning of quality controls. It helps in the prevention of massive scraps by planning quality controls regarding production control and control resources constraints. Quality and production control plans are intrinsically linked. However the actual design of these plans are separated. This exposes

industrialists to major losses. Fig. 1 illustrates this risk using one-month real data coming from a semiconductor fab. The risk exposure, referred in this example, is expressed by the number of products processed since the processing date of the last inspected product. It is called “Material-At-Risk” (MAR). The figure draws the temporal evolution of the MAR of two different processing machines. The two machines presented are equivalent as they are qualified for the same operations. Each point in the curve corresponds to one of the following events: (i) if the point represents an increase of the MAR, the event is the processing of a new product (or a lot) by the considered machine; (ii) if the point represents a decrease of the MAR, the event is the quality control of a product (or a lot) that was processed by the considered machine. A control reduces the MAR of a given value depending on production control, that is the added value of the control regarding risk exposure; (iii) else, the event is similar to that in (ii), but the performed control has no added-value.

In the illustration provided in Fig. 1, the first observation is that the risk exposure is not managed equitably between the two machines. Over the monitored period, the maximum value of MAR reached by Machine2 is 500, while Machine1 has a maximum MAR of 200. The mean values of MAR are 222 for Machine2 and 74 for Machine1. Machine2 is more exposed to risks as its MAR is higher than for Machine1. This could be partially justified by the fact that the second machine is more loaded than the first one. However, there are a significant number of controls on Machine1 without any added value (without reducing the MAR). The gap between the maximum values of MAR implies that inspections are not allocated to minimize, or at least to control, the MAR. For example, one might

* Corresponding author. Tel.: +1 5142588565.

E-mail address: samuel-jean.bassetto@polymtl.ca (S.J. Bassetto).

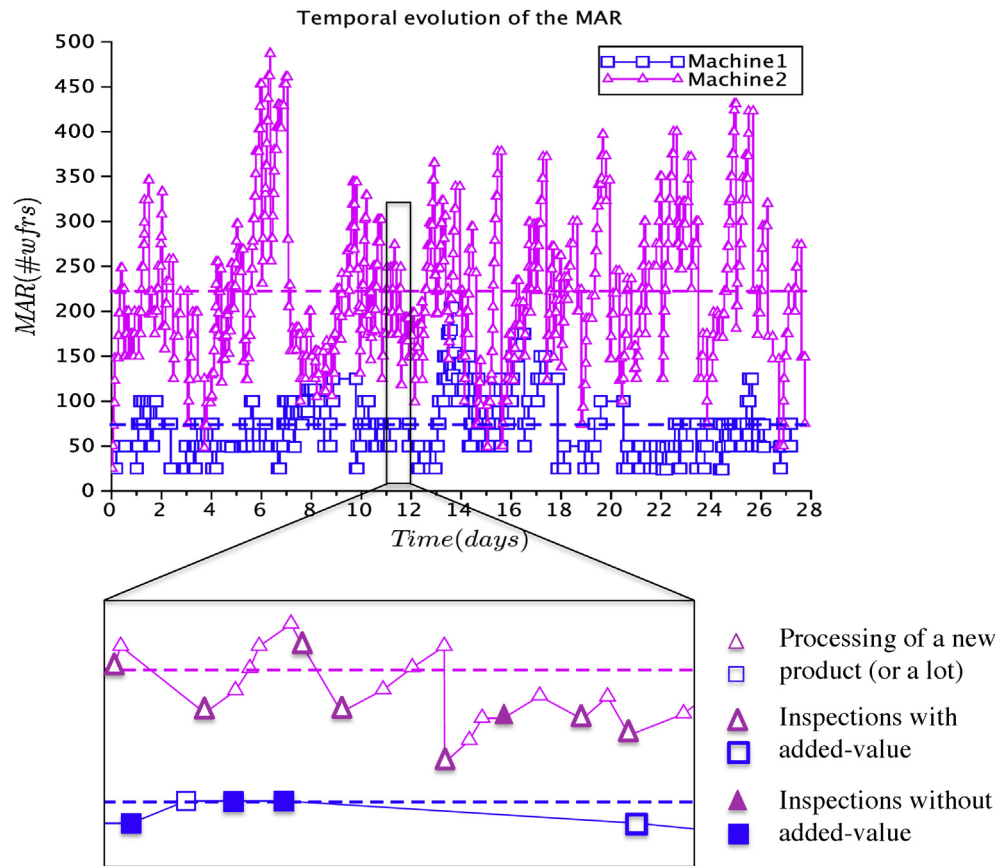


Fig. 1. Illustration of uncontrolled risk exposure.

consider using the control capacity used for controls performed on Machine1 that did not add value, to perform additional controls related to Machine2.

The second observation concerns MAR peaks of a given machine and their significance. Assuming that the control resource is supposed to be error-free, the failure not intermittent and the machines unable to self-repair, each pick corresponds to the maximum number of potentially defective products. A first question arises here: (Q1) If the maximum losses become actual, is the production organization able to face such disturbances? This unveils two others: (Q2) What is the threshold of actual loss that production organization can face and can absorb in reasonable time? and (Q3) How to take into account this threshold, if any, in the quality control plan?

The purpose of this paper is to tackle this issue by providing quality control planning that takes into account both the production plan and a risk exposure insurance level. This article intends to enhance classical quality control model, focused only on the detection speed, by including consideration of massive scraps prevention. There is then a clear inspection allocation problem, constrained by capacity of controls and influenced by Work-In-Progress (WIP) bubbles and evolutions [10].

The remainder of the paper is organized as follows. Section 2, provides a review of the literature related to process control approaches and methodologies, and how quality controls have already been coupled with operations. The main proposal of this paper is described in Section 3. It details definitions, assumptions and formulates a general model of material at risk management. The forth section is an illustration of how the model of MAR can be adapted and used by testing the approach through a real example of an industrial application. It demonstrates the usefulness of such an approach to mastering risk exposure in complex discrete manufacturing systems. Section 5 discusses the limitations and gives

some perspectives of this proposal which constitute the directions for further research. Section 6, which summarizes the aim and the contribution of this research, concludes this paper.

2. Literature review

Quality control crosses various disciplines in an effort to establish appropriate layers of protection [24]. Accordingly, this review focuses on the quality control techniques available to prevent the production of excessive amounts of scraps: risk management, Statistical Process Control (SPC), inspection allocation, and the integration of process control into operational activities.

2.1. Risk management

Almost all semiconductor manufacturers need to provide updated risk analyses about their processes and products, and sometimes their machines, with the objective of assuring their customers of their ability to deliver products on time, and in the quantity and quality required. These analyses concern the operational risks, which have to be determined, evaluated (often using ranking techniques), and mitigated, with follow-up for the best case. These are risks of the occurrence of events that have potentially serious consequences. FMECA (Failure Mode, Effects and Criticality Analysis) is one of the techniques most often applied [30], however there are many others. A general survey of modern risk management methodologies can be found in Ref. [28].

In risk analysis, layers of protection are explicitly mentioned. In FMECA, they are listed in the column labelled "Detectability". However, very few methods link efficiently risk analyses and actual control plan strategies, which would take into account the potential excessive amount of scrap production. From the risk analysis

perspective, controls are often mentioned in a generic phrase, like “Control with SPC” or “Perform maintenance”. To be fully meaningful and operational, these expressions have to be manually reworked by a specialist in the field. For example, the phrase “Perform maintenance” has to be expanded in the maintenance information system to include details such as label, target machine and spare parts, maintenance frequency and operating mode. However, this manual operation often disconnects risk analysis from the control planning process, which means that the assurance perspective of these risk analysis is then not guaranteed. To address this issue, Mili et al. [18] propose an activity model linking risk elicitation, risk evaluation and the risk control plan. This framework has been applied to define maintenance priorities in a photolithography workshop. Every time a failure is detected, the risk analysis is revised and the control plan is modified accordingly. However their method does not guarantee any insurance against substantial losses from excessive scrap production.

The work of Bean [1] presents an interesting development, which is an in-line dynamic inspection plan, based on the probability of deviations due to measurement errors. Bean also introduces the notion of “material at risk” (MAR), as every product processed between two sampled products may be defective. He does not explain what kinds of failure are monitored, but bases his inspection plan only on potential impact, whatever its causes. His work points to a central concept for the prevention of excessive amounts of scrap or of recalls, which is the monitoring and control of MAR in order to avoid as much as possible product being delivered to customers which ultimately proves to be defective.

2.2. Statistical Process Control

Statistical Process Control (SPC) plays a central role in quality assurance. In developing SPC techniques [19], particular effort has been devoted to designing and updating control decision parameters dynamically throughout a production life-cycle. The adaptation of control plans to events as they occur is a major advancement in terms of detecting drifts faster. Varying the sampling interval (sampling frequency), varying the sampling size, or changing the control limits are actions commonly taken whenever data are collected. Investigation of this subject began with the publication of Ref. [23], who demonstrated that a two-level control (each level has its own values of sampling size, frequency and control limits) is the optimal solution to achieve control and detect drifts more quickly, while minimizing the cost of errors. Adaptive process control has been reviewed by Tagaras [26]. Not long ago, De Magalhães et al. [8] presented a very clear overview of these techniques, publishing a key paper in this field. Recent literature focused on the design of control charts for non-normal data distribution [16] or for correlated data [4]. Several studies combine the economic design approach and the adaptive control approach. Prabhu et al. [20] present an integrated statistical and economical design of an adaptive control chart with a variable sampling size and a variable sampling interval. More recently, Torng et al. [29] have investigated a design incorporating a dual sampling control chart, to monitor both the economic and statistical aspects of controls. As the performance of controls is usually measured by the speed of detection, by integrating economic evaluation their work opens the way to addressing the scrap prevention issue. However, we have not found any work in the literature on the limits to designing or adapting quality controls from the insurance perspective.

2.3. Inspection allocation

As control resources are limited, they must be allocated judiciously. Some five decades ago, Lindsay and Bishop [17] developed an economic algorithm for inspection allocation, based on a cost

function per unit produced, taking in account the inspection cost and its location in the process. Since their paper was published, the field of inspection allocation has become an area of intense research. We recommend the surveys of Refs. [22,27] for a complete picture of the field. Among the remarkable works on the subject, several have been of particular importance for the effort to improve production reliability. Villalobos et al. [32] presented a flexible inspection system for serial and multi-stage production systems in the field of Printed Circuit Boards. They provide a dynamic programming algorithm to optimize global goals (like costs), while taking into account some local constraints, like inspection machine availability. Verdusco et al. [31] present an interesting case of information-based inspection allocation. They modeled a cost function taking into account Type I and Type II errors, and the information gain attained by each measurement. Moreover, they formulated the inspection allocation problem as a Knapsack Problem (KP), and proposed a greedy algorithm to solve it. Their simulations reveal that the information-based solution performs better than static inspections, in terms of classification errors. The Works of Rabinowitz and Emmons [21] and Emmons and Rabinowitz [9] present an inspiring non-linear modeling of the inspection allocation problem, in which they introduce the time passed since the last inspection (which they call Z_{ij} in their model) and link the proportion of defect-free items to this time. This conceptualization is strongly related to the concept of MAR. In fact, we have adopted part of their conceptual framework in our work here, which deals with a particular convex loss function that could be modeled by an infinite loss over a defined level. Kogan and Raz [15] also present an interesting model of inspection allocation effort, as they seek to minimize a cost function influenced by inspection costs, as well as multiple failure modes that can affect each production step. They also suggest multiple layers of detection that can be used for every product, at every production stage.

2.4. Integration of quality control with operational activities

In several domains, links between inspection allocation and operations are found. In the food industry, for example, inspection allocation is considered as a traceability issue. This topic has been widely studied. The work of Wang et al. [33], in which an inspection effort is jointly designed to minimize risk and improve operations, is of particular interest to us. They propose an integrated optimization model that computes safety and production parameters to issue an optimal production control plan tracability. In the structural steel industry, an inspection model based on risk modeling has been presented by Straub [25], who investigated the way cracks and failures occur in bridges. He defined a control policy based on a particular level of risk acceptance. This work has been especially inspiring for us, as it is the only work on control allocation based on the risk of failure and a risk acceptance level, rather than on a general cost function. In the flexible manufacturing literature, quantity and quality have been included in an integrated design strategy. As pointed out in the introduction, risks can be strongly influenced by operational management, particularly in terms of the number of products produced between two controls. Hsu and Tapiero [12] were pioneers in terms of proposing a link between operational management and SPC control charts. The papers published on this topic [14] and [5] present academic investigations on how quality and operations control can be linked. Colledani, for example, designed the buffer size of the control machine to meet quality and cycle time expectations. More recently, further developments have been achieved by Colledani and his team [6]. At every stage, the balance between the cost of non-conforming products and the cost of buffers has been knowledgeably investigated.

2.5. Literature conclusion

The literature covers several domains: risk, adaptive SPC control charts, inspection allocation and the integration of quality and production management. The classical process of quality assurance falls short, as a result of failing to take into account its interaction with production control activities. The integration of these activities early in the planning stage, coupled with a risk perspective seems a promising research avenue. Currently, these disciplines seems to be only loosely coupled, since few developments encompass all of them. Moreover, we have found no research conducted on a quality control plan design model aimed at reducing the impact of products scraps or recalls.

3. Problem formulation

The central concept used in our formulation is that of “material at risk” (MAR), which refers to the quantity of products on which no quality control has been performed. These products are potentially defective. With the MAR concept, we can introduce the insurance viewpoint into the design of a quality control plan. We define the resilience limit R_L as the threshold of the number of defective products that can be easily absorbed by the manufacturing system without leading to a major disruption in performance. For instance, such a loss could be absorbed by subcontracting or by increasing productive capacity, temporarily. The manufacturing system is considered to be in job-shop configuration, where process operations can be run in serial or in parallel, or both. A failing machine can produce faulty products even though its successors in the product flow model are “under control”. Incidentally, the MAR can rise even if only one machine is not controlled. Once a machine drifts, all the subsequent products processed by that machine are affected. This is why we consider that the MAR level has to be controlled as close as possible to the source of the defect: the processing machine. For this reason, we consider that every machine has to be accredited before its production level exceeds R_L . This is the main constraint in our formulation, as it prevents the number of scraps or recalls from becoming too large. In a lean manufacturing context, we seek to minimize the number of quality controls, while satisfying the constraint related to the MAR.

3.1. Notations and definitions

More details are introduced in this section, by providing notations and definitions to be employed later in the development of the model. Fig. 2 is used as support for the general context and the problem formulation. The formulation concerns the quality control planning of a manufacturing system during a planning horizon equivalent to a given production plan. The latter enumerates the quantity to produce, denoted by Q_p , of each product $p \in \{1, \dots, P\}$. Each product p has a predefined process flow model which describes successive production operations to be performed in order to obtain a unit of it. Inspection operations are performed or not accordingly to the quality control policy in place. The considered production resources are divided into two categories: processing machines for process operations and inspection machines for inspection operations.

Let Z be the set of processing machines and z the index associated with it. Let H_z be the planning horizon of the processing machine z that is considered. It represents the forecasted load (or planned) of the processing machine z , expressed in number of items to be processed on it, during the period of time necessary to execute the whole production plan. When a product progresses through the various steps of the production process, it passes through several layers of control where it can be inspected, or not, depending on

the actual sampling strategy in place. Let K be the set of all possible layers of control of the production system and k the index of each individual layer of control with $k \in K$. These control layers are placed in the process flow in such a way that they allow the occurrence of a failure mode, denoted by λ , to be detected. Each layer of control is associated with a set of measurement resources with limited capacity. For sake of simplicity, a unique control resource for each control layer k (denoted by IM_k), as presented in Fig. 2. Let Θ be the set of known failure modes of the production system that can be detected by one or more of the control layers in place, and Θ_z the subset of failure modes associated with processing machine z . Let α_λ be the estimated value of the probability of occurrence of failure mode λ .

Let $MAR_z^\lambda(\cdot)$ be the MAR function of machine z that represents the number of items potentially impacted by the failure mode $\lambda \in \Theta_z$, i.e. the number of processed items since the last control of failure mode λ was applied. Let $MAR_z^{\lambda,0}(\cdot)$ be the MAR in a reference situation, and let $MAR_{z,\max}^{\lambda,0} = \max_{t \in [0, H_z]} MAR_z^{\lambda,0}(t)$ be the maximum value

of MAR reached during the horizon H_z considered regarding for the failure mode λ . If the reference situation is the case where no control is planned during the horizon considered, the MAR is characterized by: $MAR_z^{\lambda,0}(t) = \alpha_\lambda t$ $t \in [0, H_z]$ and $MAR_{z,\max}^{\lambda,0} = \alpha_\lambda H_z$.

Now, consider the case where a control plan $x = (n^x, K^x, T^x)$ is defined to manage the MAR for all the machines of the manufacturing system. A control plan is characterized by:

- n_z^x : Number of controls related to processing machine z
- n^x : Total number of controls of control plan x . $n^x = \sum_{z \in Z} n_z^x$
- T^x : Vector of dates of n^x controls, $T^x = (T_1^x, \dots, T_{n^x}^x, \dots, T_{n^x}^x)$, with $T_z^x = (t_{z,1}^x, t_{z,2}^x, \dots, t_{z,i_z}^x, \dots, t_{z,n_z^x}^x)$
- t_{z,i_z}^x : Date of the i_z -th quality control performed on machine z , with $t_{z,i_z}^x \in [1, H_z] \forall i_z \in \{1, \dots, n_z^x\}$ and $t_{z,i_z}^x < t_{z,j_z}^x \forall i_z < j_z$
- K^x : Set of control layers used for the n^x controls, with $K^x = (K_1^x, \dots, K_{n^x}^x, \dots, K_{n^x}^x)$ where $K_z^x = (k_{z,1}^x, k_{z,2}^x, \dots, k_{z,i_z}^x, \dots, k_{z,n_z^x}^x)$, with $k_{z,i_z}^x \in K \forall i_z \in \{1, \dots, n_z^x\}$
- C_k : Capacity of the control layer k , expressed in time units or in cost units
- ct_k : Cost of an unitary control using the layer k
- $MAR_{z,\max}^x$: Maximum value of MAR reached during the horizon H considered in the situation with control plan x in place. It is defined by the maximum value of MAR reached regarding all failure modes. $MAR_{z,\max}^x = \max_{\lambda} MAR_z^{\lambda,x}$

3.2. Assumptions

In order to provide an analytical formulation, the following assumptions are made:

Assumption 1. The probability of occurrence of the failure mode α_λ considered is constant all along the horizon H_z , i.e. α_λ is not affected by control actions.

Assumption 2. Once a failure occurs on a processing machine, it continues to impact the items produced until a control is performed. Once a failure detected, the processing machine concerned is immediately switched off until a corrective action is undertaken.

Assumption 3. The resilience limit is constant ($R_L = Cte$), i.e. the budget allocated to compensate nonconformities ‘crisis’ is immediately replenished each time it is, wholly or partially, used.

These assumptions represent the following scenario: the production continues on until a systematic product control action is applied. The MAR increases regularly with the production. If the threshold R_L is exceeded, a major disruption in production

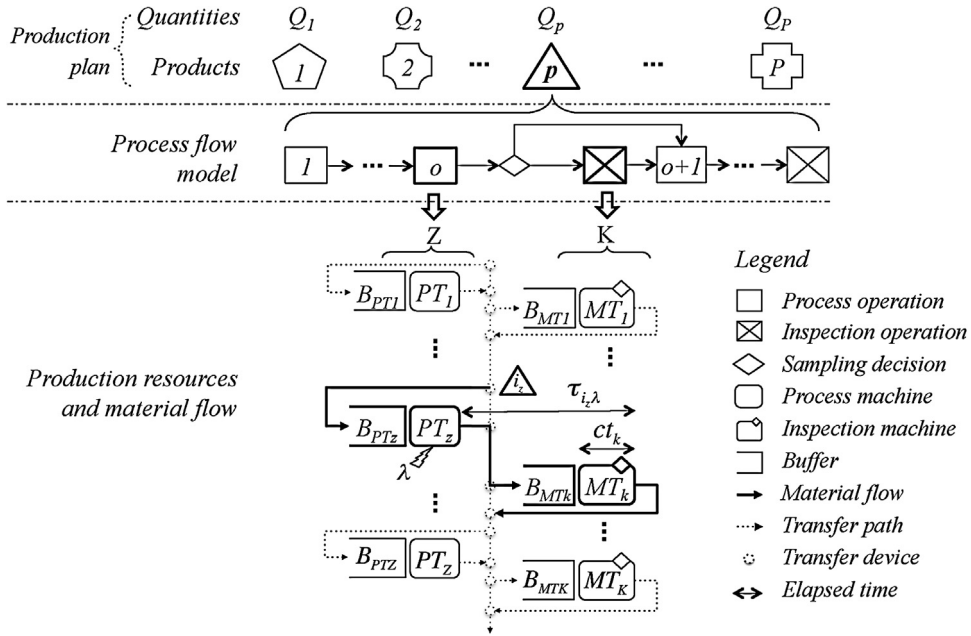


Fig. 2. Problem formalisation.

could occur (leading to customer penalties, for example, and even bankruptcy), while a loss below this threshold could be offset internally or with the help of a partner.

3.3. Model formulation

Assuming that each failure mode $\lambda \in \Theta_z$ can affect the products processed by the processing machine z , the MAR of that machine, relative to quality control plan x and failure mode λ , evolves following a general model as presented in Fig. 3. The x -axis represents the production time expressed in number of processed products and the y -axis represents the value of MAR. During its production horizon H_z , the machine z is controlled by inspecting n_z^x products processed by it. The products sampled for inspection are indexed by $(t_{1z}, t_{2z}, \dots, t_{iz}, \dots, t_{n_z^x})$, where i_z corresponds to the i -th sample taken from the machine z . The effect of quality controls on the evolution of the MAR depends on the parameters α , τ and ρ . The meanings of these parameters are explained below.

The parameter α represents the rate of increase in MAR. It reflects the probability of occurrence of the considered failure mode on the processing machine. α_{i_z} represents the value of this rate between the effective dates of the $(i_z - 1)$ -th and the i_z -th controls. The variation of α from one period to another may be due to the

following reasons: (i) the previous inspection permitted to detect the failure occurrence, which led to a corrective action performed on the machine; (ii) the previous inspection did not detect anything, but the information provided by the inspection permitted to adjust the estimate of the probability of occurrence of the failure mode.

The parameter τ represents the control delay of the MAR. It corresponds to the number of products processed between the instant the inspected product is processed and the instant the value of the MAR is updated. This delay is induced by one or more of the following: transfer time, waiting time, inspection time and data analysis time.

The parameter ρ represents the rate of reduction in MAR. It reflects the accuracy of the inspection resource regarding the monitored failure mode. When a control is performed on a machine, the MAR is reduced by an amount proportional to its value at the time the inspected product was processed. The reduction value of the MAR due to the i_z -th control, denoted by Δ_{i_z} , depends on ρ_{i_z} and is expressed by the following formula:

$$\Delta_{i_z} = \rho_{i_z, \lambda} \times (MAR_z^{\lambda, x}(t_{i_z}) - \alpha_{i_z, \lambda} \tau_{i_z, \lambda}) \quad (1)$$

The evolution of the MAR as modelled in Fig. 3 represents some features to facilitate its use in the design of a quality control plan

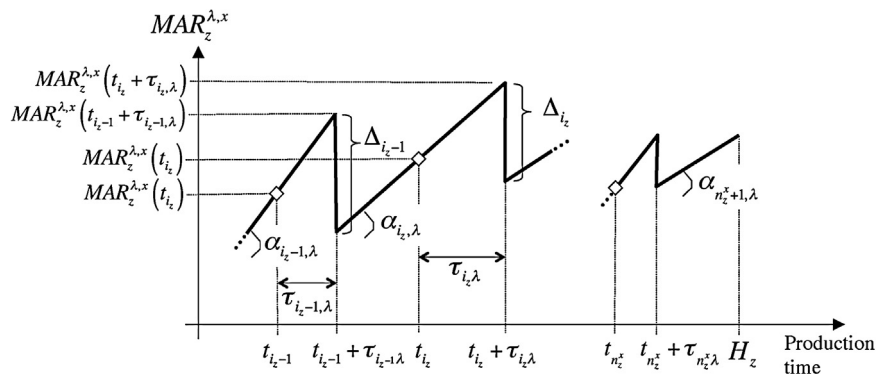


Fig. 3. General model.

based on risk-exposure. As this insurance approach is based on the control of risk-exposure, the peaks values of the MAR are particularly important. These peaks, which correspond to the moments where the quality controls reduce the MAR, are used to characterize the behaviour of the MAR by the following property:

Property 1. For all quality control plan x , for all processing machine z and for all failure mode λ , the peak values of the corresponding MAR model can be expressed by the following recursive expression:

$$\begin{aligned} \text{MAR}_z^{\lambda,x}(t_{iz} + \tau_{iz,\lambda}) &= \text{MAR}_z^{\lambda,x}(t_{iz-1} + \tau_{iz-1,\lambda}) \times (1 - \rho_{iz-1,\lambda}) \\ &\quad + \rho_{iz-1,\lambda} \alpha_{iz-1,\lambda} \tau_{iz-1,\lambda} + \alpha_{iz,\lambda} \\ &\quad \times (t_{iz} + \tau_{iz,\lambda} - t_{iz-1} - \tau_{iz-1,\lambda}) \\ &\quad \forall i_z = 1, \dots, n_z^x + 1 \end{aligned} \quad (2)$$

With: $\rho_{0z,\lambda} = 1$, $\tau_{0z,\lambda} = 0$, $t_{0z} = 0$, $\tau_{n_z^x+1,\lambda} = 0$ and $t_{n_z^x+1} = H_z$.

Proof. See Appendix 1. \square

The Property 1 implies that the value of the peak of MAR of a processing machine z , issued from the current quality control action, depends on: (i) the previous value of the peak of MAR; (ii) the accuracy of the inspection machine used in the previous quality control; (iii) the current probability of occurrence of the failure mode; and (iv) the time delay between the processing machine and the selected inspection machine.

The objective is to find an optimal quality control plan $x^* = (n^*, T^*, K^*)$ where n^* is minimized, while ensuring that the R_L constraint is respected by every machine. If such a control plan exists, the following condition is necessary and sufficient:

Property 2. A QCP aiming to minimize the risk exposure is optimal iff the values of the peaks of the corresponding MAR curve are equal for each of the processing machines, i.e.:

$$t_{1z}^* = \frac{H_z + \sum_{i_z=1}^{n_z^*} (1/\alpha_{i_z+1,\lambda}) (\tau_{i_z,\lambda} \rho_{i_z,\lambda} \alpha_{i_z,\lambda} + \alpha_{i_z+1,\lambda} (\tau_{i_z+1,\lambda} - \tau_{i_z,\lambda})) - \alpha_{1z,\lambda} \tau_{1z,\lambda} \sum_{i_z=1}^{n_z^*} (\rho_{i_z,\lambda} / \alpha_{i_z+1,\lambda})}{(1 + \alpha_{1z,\lambda} \sum_{i_z=1}^{n_z^*} (\rho_{i_z,\lambda} / \alpha_{i_z+1,\lambda}))} \quad (8)$$

Knowing that $t_{0z} = 0$, $\text{MAR}_z^{\lambda,x^*}(t_{0z}) = 0$ and $\text{MAR}_z^{\lambda,x^*}(t_{1z} + \tau_{1z,\lambda}) = \alpha_{1z,\lambda} (t_{1z}^* + \tau_{1z,\lambda})$, the optimal control dates can be expressed by a recursive formula, as described in the following corollary.

Corollary 1. For all processing machine $z \in Z$, for all failure mode $\lambda \in \Theta_z$, the optimal control dates t_{i_z} , $\forall i_z \in \{1, \dots, n_z^* + 1\}$, are expressed by:

$$\begin{aligned} t_{i_z}^* &= t_{1z}^* \left(1 + \alpha_{1z,\lambda} \sum_{j_z=1}^{i_z-1} \frac{\rho_{j_z,\lambda}}{\alpha_{j_z+1,\lambda}} \right) \\ &\quad - \sum_{j_z=1}^{i_z-1} \frac{1}{\alpha_{j_z+1,\lambda}} (\tau_{j_z,\lambda} \rho_{j_z,\lambda} \alpha_{j_z,\lambda} + \alpha_{j_z+1,\lambda} (\tau_{j_z+1,\lambda} - \tau_{j_z,\lambda})) \\ &\quad + \alpha_{1z,\lambda} \tau_{1z,\lambda} \sum_{j_z=1}^{i_z-1} \frac{\rho_{j_z,\lambda}}{\alpha_{j_z+1,\lambda}} \end{aligned} \quad (5)$$

Proof. See Appendix C. \square

As the principle constraint is to stay in an “in-control MAR” situation, the target control plan solution has to verify:

$$\text{MAR}_{z,\max}^{x^*} = \max_{\lambda \in \Theta_z} \max_{i_z \in \{1, \dots, n_z^*\}} \text{MAR}_z^{\lambda,x^*}(t_{i_z}^* + \tau_{i_z,\lambda}) \leq R_L \quad \forall z \in Z \quad (6)$$

Using Eqs. (3) and (5) this constraint can be reformulated as follows:

$$\text{MAR}_z^{\lambda,x^*}(t_{1z}^* + \tau_{1z,\lambda}) = \alpha_{1z,\lambda} (t_{1z}^* + \tau_{1z,\lambda}) \leq R_L \quad \forall z \in Z \quad \forall \lambda \in \Theta_z \quad (7)$$

As Eq. (5) is also valid for $t_{n_z^*+1}^* = H_z$, the expression of t_{1z}^* is given below:

Then, the target control plan $x^* = (n^*, T^*, K^*)$ that minimizes $n = \sum_z n_z$, subject to the R_L constraint, should satisfy the following condition $\forall z \in Z$ and $\forall \lambda \in \Theta_z$:

$$\alpha_{1z,\lambda} \frac{H_z + \sum_{i_z=1}^{n_z^*} (1/\alpha_{i_z+1,\lambda}) (\tau_{i_z,\lambda} \rho_{i_z,\lambda} \alpha_{i_z,\lambda} + \alpha_{i_z+1,\lambda} (\tau_{i_z+1,\lambda} - \tau_{i_z,\lambda})) - \alpha_{1z,\lambda} \tau_{1z,\lambda} \sum_{i_z=1}^{n_z^*} (\rho_{i_z,\lambda} / \alpha_{i_z+1,\lambda})}{(1 + \alpha_{1z,\lambda} \sum_{i_z=1}^{n_z^*} (\rho_{i_z,\lambda} / \alpha_{i_z+1,\lambda}))} + \alpha_{1z,\lambda} \tau_{1z,\lambda} \leq R_L \quad (9)$$

In order to give a standard formulation of the problem, let's define the following decision variables:

$$X_{izk} = \begin{cases} 1 & \text{if the control layer } k \text{ is selected at the } i_z\text{-th control of machine } z; \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Knowing that $\tau_{i_z,\lambda}$, $\rho_{i_z,\lambda}$ and $\alpha_{i_z,\lambda}$ depend on the choice of the control layer used in the i_z -th control of machine z , the latter decision variables are included in the model as follows:

$$\tau_{i_z,\lambda} = \sum_{k \in K} X_{izk} \tau_{k\lambda}$$

$$\rho_{i_z,\lambda} = \sum_{k \in K} X_{izk} \rho_{k\lambda}$$

$$\alpha_{i_z,\lambda} = \alpha_{\lambda} \quad \forall i_z \in \{1, \dots, n_z + 1\}$$

$$\begin{aligned} \text{MAR}_z^{\lambda,x^*}(t_{i_z}^* + \tau_{i_z,\lambda}) \\ = \text{MAR}_z^{\lambda,x^*}(t_{j_z}^* + \tau_{j_z,\lambda}) \quad \forall i_z, j_z \in \{1, \dots, n_z^* + 1\} \quad \forall z \in Z \quad \forall \lambda \in \Theta_z \end{aligned} \quad (3)$$

$$\text{with } \text{MAR}_z^{\lambda,x^*}(t_{n_z^*+1}^*) = \text{MAR}_z^{\lambda,x^*}(H_z)$$

Proof. See Appendix B. \square

Properties 1 and 2 are then used to express the value of the peak of MAR corresponding to the i_z -th quality control, for each failure mode $\lambda \in \Theta_z$, as follows:

$$\begin{aligned} \text{MAR}_z^{\lambda,x^*}(t_{i_z}^* + \tau_{i_z,\lambda}) \\ = \frac{1}{\rho_{i_z-1,\lambda}} (\tau_{i_z-1,\lambda} \rho_{i_z-1,\lambda} \alpha_{i_z-1,\lambda} + \alpha_{i_z,\lambda} (t_{i_z}^* + \tau_{i_z,\lambda} - t_{i_z-1}^* - \tau_{i_z-1,\lambda})) \end{aligned} \quad (4)$$

(Assumption 1) with

$$\sum_{k \in K} X_{izk} = 1 \quad \forall i_z \in \{1, \dots, n_z\}$$

$$\sum_{i_z \in \{1, \dots, n_z\}} X_{izk} \leq n_z \quad \forall k \in K$$

The problem can now be formulated as follows: **(P)**

$$\text{Minimize } \sum_{z \in Z} \sum_{k \in K} \sum_{i_z=1}^{H_z} X_{izk}$$

subject to

$$n_k^{\min} \leq \sum_{z \in Z} \sum_{i_z=1}^{H_z} X_{izk} \leq \sum_{z \in Z} H_z \quad \forall k \in K$$

(C1)

$$\sum_{k \in K} X_{izk} \leq 1 \quad \forall i_z \in \{1, \dots, H_z\} \quad \forall z \in Z$$

(C2)

$$\sum_{i_z=1}^{H_z} \sum_{k \in K} X_{izk} \left(\frac{R_L}{\alpha_{k\lambda}} \rho_{k\lambda} - \tau_{k\lambda} (1 + \rho_{k\lambda}) \right) + X_{i_z+1k} \tau_{k\lambda} - \sum_{k \in K} X_{1zk} \tau_{k\lambda} \geq H_z - \frac{R_L}{\alpha_{k\lambda}} \quad \forall z \in Z \quad \forall \lambda \in \Theta$$

(C3)

$$\sum_{z \in Z} \sum_{i_z=1}^{H_z} X_{izk} c t_k \leq C_k \quad \forall k \in K$$

(C4)

This formulation is lean manufacturing-oriented, as it poses as objective function **(P)** the minimization of the total number of controls. It also takes into account production planning, represented by the H_z values. The insurance perspective toward scraps and recalls is addressed by constraint **(C3)**. Constraint **(C1)** represents the possibility of taking into account the imposed minimum number of controls (n_k^{\min}) for some control layers in the manufacturing system (customer's requirements, standards, etc.). The limited capacity of control layers is also considered in constraint **(C4)**. The resolution of this problem is intended to minimize control activities, while guaranteeing the control of the scrap and recall risks levels.

In order to implement this model in practice, several parameters have to be explicitly identified. Unfortunately, this exercise can not be completed exhaustively in a real industrial case. While the set of control layers K is generally known, the set of failure modes for each processing machine Θ_z is often neither explicitly nor exhaustively identified. Incidentally, the pending parameters $\alpha_{iz\lambda}$, $\tau_{iz\lambda}$ and $\rho_{iz\lambda}$ are uncertain and difficult to evaluate. Otherwise, the general formulation can be easily simplified to give an adapted formulations to specific contexts where one or more parameters could be neglected. Moreover, there could be situations where these parameters are stochastic or dependent to each other, in which case the problem has new different characteristics and the resolution methods could be different.

For these reasons, and in order to illustrate the proposed model with a real case study, the remainder of this paper presents a simplified version of the formulation which is tested in an industrial context.

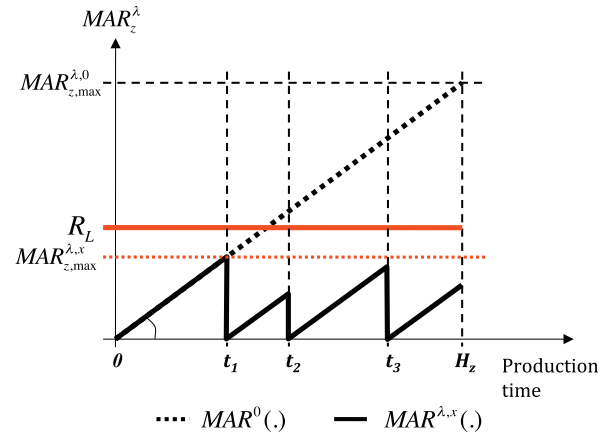


Fig. 4. MAR evolution in the simplified formulation with a given control plan.

4. Illustration

In order to illustrate our proposal, a simplified formulation is provided in this section and then tested in an industrial case study. Although the proposed generic model is tested in its simplest form, the results obtained prove the effectiveness of the risk-based approach for quality control planning.

4.1. Simplified formulation

The simplified formulation is derived in accordance with the need of the industrial case study. In this model, the following additional assumptions are taken into account:

Assumption 4. Only one layer of protection k and one failure mode λ are considered.

Assumption 5. Controls are performed immediately: no time elapses between control decision and control execution, $\tau_{k\lambda} = 0 \quad \forall k \in K; \forall \lambda \in \Theta$.

Assumption 6. Perfect control actions which reset the level of MAR to 0, i.e. $\rho_{k\lambda} = 1 \quad \forall i_z \in \{1, \dots, n_z\}; \forall k \in K; \forall \lambda \in \Theta$.

The MAR evolution for a given control plan can be illustrated as in Fig. 4, where the MAR increases with the production throughput. It is reset to 0 whenever a control is performed. In this figure, four controls have been planned in order not to exceed R_L , three during the production, one at the beginning.

With the previous additional assumptions, the problem is formulated as follows:

$$\text{Minimize } \sum_{z \in Z} \sum_{i_z=1}^{H_z} X_{iz}$$

(P') subject to

$$n^{\min} \leq \sum_{z \in Z} \sum_{i_z=1}^{H_z} X_{iz} \leq \sum_{z \in Z} H_z$$

(C'1)

$$0 \leq X_{iz} \leq 1 \quad \forall i_z \in \{1, \dots, H_z\} \quad \forall z \in Z$$

(C'2)

$$\sum_{i_z=1}^{H_z} X_{iz} \geq \alpha \frac{H_z}{R_L} - 1 \quad \forall z \in Z$$

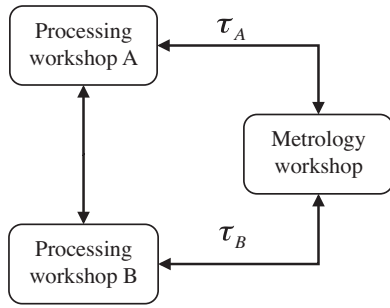


Fig. 5. Simplified flow model of the case study.

(C'3)

$$\sum_{z \in Z} \sum_{i_z=1}^{H_z} X_{i_z} c_t \leq C$$

(C'4)

With these notations and assumptions an analytical result can be provided:

- 1 The equidistant distribution of controls along the considered horizon maximizes the added value of the control plan. The optimal positions of controls are defined by: $t_{i_z}^* = i_z \times (H/n_z + 1) \forall i_z \in \{1, \dots, n_z\}$
- 2 The optimal number of controls of the simplified model is given by the formula $n_z^* = \lceil \alpha \times (H_z/R_L) - 1 \rceil$, where $\lceil x \rceil$ is the first integer greater than or equal to x .

These analytical results can be used to include an insurance viewpoint during the design of a manufacturing system quality control plan. This simplified version generates an NP-Hard problem by analogy with the 0–1 Knapsack problem [2].

4.2. Illustrative application in the semiconductor industry

We had the opportunity to illustrate our model with an industrial case study concerning an application in a semiconductor manufacturing company. The case focuses on two key workshops (named A and B in the remainder). The illustration consists of a before/after comparison of how MAR is controlled. Although the quality control techniques used were identified as among the best in their class in the field, major scrap production remained a reality. This convinced the managers to complement their quality techniques with tools that control the MAR level throughout the plant.

The shop floor considered is presented in Fig. 5. The case study involved 3 production instances, each of 1 month's duration. In each instance, six product families are produced in several quantities. For each instance, the production plan corresponds to about 600,000 wafer operations per month which represents about 65,000 wafers. In this case study the number of qualified machines in each workshop and the processing times of the process and measurement steps are supposed constant over the planning period. Delays between processing machines and measurement machines were constant for all the workshops throughout the experiment. Delays express the number of wafers processed during the time a lot moves from the processing machine to the measurement machine. The characteristics of the case study are given below:

$NPM_A = 23$: Number of Process machines in A

$NPM_B = 13$: Number of Process machines in B

$\tau_A = 177W/frs$: Delay between A machines and Metrology workshop

Table 1

Measurement case study: before/after results.

	Before	After	MAR variation
<i>Mean_zMAR_{max,z}</i>			
Instance1	383	224	−41.5%
Instance2	450	252	−44%
Instance3	436	251	−42.4%
<i>Mean_zMAR_z</i>			
Instance1	137	46	−66.4%
Instance2	168	56	−66.7%
Instance3	142	46	−67.6%
<i># OverRiskyLots</i>			
Instance1	692	28	−95%
Instance2	2598	41	−98.4%
Instance3	7329	1054	−85.6%

$\tau_B = 116W/frs$: Delay between B machines and Metrology workshop

$NMT = 12$: Number of machines in the Metrology workshop

$NP = 6$: Number of products (technologies or product families)

$NOP = 313$: Number of operations to be controlled by the Measurement machines

The objective of this case study was to test, on a scale-1 case, the usefulness of our strategy for allocating quality controls dynamically and managing the MAR level. The target control plan is used as input to the dynamic sampler of the measurement in the fab. This input, called the Warning Limit, is one of the characteristics used in the real-time decision-making of the sampler process: to sample, to measure, or to skip a lot [7]. The Warning Limit of a processing machine corresponds to the expected maximum value of its MAR minus the delay: $MAR_{max,z}^* - \tau_z$ and serves as an “alarm” as the machine goes over its WL_z . An item should be sampled for control so that the MAR remains as close as possible to its forecasted (optimal) value MAR_{max}^* . The latter is computed using a 2-stages heuristic algorithm that solves the simplified model including the capacity constraint [3]. Comparison is performed of the evolution of the MAR before and after the use of a sampler based on our approach to Warning Limit computation. The results have been compared with 3 indicators:

- *Mean_zMAR_{max,z}*: This indicator is used to monitor the maximum MAR for each machine. It is an estimate of the maximum risk exposure of the production system. Note that for each machine z the value of $MAR_{max,z}$ is not an estimated value but the real value achieved by the machine.
- *Mean_zMAR_z*: This indicator computes the mean of the MAR over the production horizon and over all the processing machines. It is the mean of the risk exposure of the manufacturing system.
- *# OverRiskyLots*: The number of lots where the machine was above MAR_{max}^* . Each of these lots warns the production manager that the manufacturing system is facing a major risk of disruption.

Table 1 presents the results for the three instances cited above. They show the advantages of using a risk-based control planning strategy for the real-time management of MAR. This case study reveals that the manufacturing system was exposed to a mean MAR of 149 wafers per month. That value is divided by three when the Warning Limit computation using our approach is used as input to the sampler. This means that the expected monthly potential loss has been divided by three, a result that is of prime importance. This industrial case confirms that it is possible to compensate for a massive loss. Instead of potentially losing 149 products (on average), the number is reduced to 49. This example demonstrates the ability of our technique to control potential losses.

Material At Risk

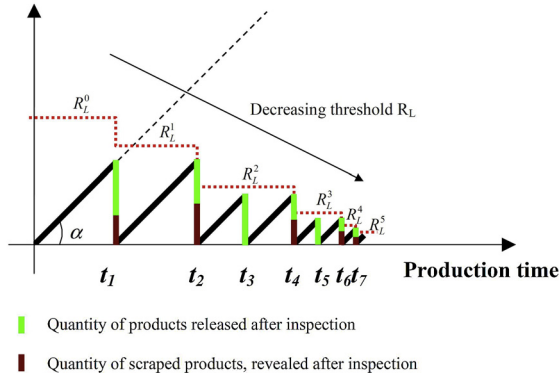


Fig. 6. A decreasing risk exposure threshold.

Material At Risk

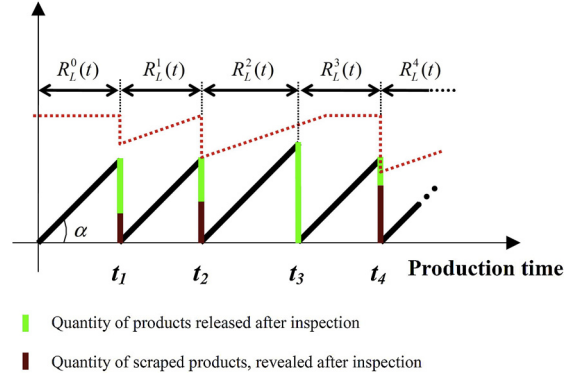


Fig. 7. A refundable risk exposure threshold.

The difference between the values of $Mean_zMAR_{max,z}$ and $Mean_zMAR_z$ for the three instances reveals that some processing machines present some MAR peaks. These peaks become less significant with our strategy, because it reduces the $Mean_zMAR_{max,z}$ and $Mean_zMAR_z$ by 42% and 66% respectively, on average. Another advantage of using the risk-based control planning strategy is the substantial decrease in the number of over-risky lots. This makes it possible to decrease the number of alarms which are always a source of conflict, and the stress that can exist between the various departments of the manufacturing system.

5. Discussion and model enhancement

This section presents several possible improvements to the model that would surpass weaknesses related to some restrictive assumptions.

5.1. Threshold risk exposure sensitivity

A possible enhancement is the sensitivity of R_L to failures. Assumption 3 enables an infinite capacity of the production system or the refunding system. For small amounts of loss this assumption can be valid but not for a large one. An initial rough estimation of the exposure limit R_L as a fraction of the capacity (1/10 for instance) can lead to an iterative revision of the control plan. If at most $\rho_{k\lambda} \times MAR_z^x(t_{iz})$ products can be rejected each time a control (t_{iz}) is applied, then these products have to be then reordered, leading to a capacity consumption and then a decrease in R_L according to the equation $R_L(t_{i+1}) = f(Capacity(t_i), \rho_{k\lambda})$.

Fig. 6 illustrates a simple case, in which controls are acceptance tests, the risk exposure threshold is sensitive to rejects (i.e. the more rejects there are, the less insurance is available), and exposure to loss is non refundable. This situation can be easily encountered in industry, where R_L is a budgeted number of products that can be scrapped over a given period. During this period, the budget cannot be revised. This leads to a direct increase in the number of controls, and in the worst case scenario, to the need for a control for every product produced.

A refunding mechanism acting on the threshold R_L can also be envisioned. Every time a product is produced and sold, a fraction of the funding is devoted to R_L , which becomes dependent on the production time, enhancing the previous equation to $R_L(t_{i+1}) = f(Capacity(t_i), \rho_{k\lambda}, t_i)$. This mechanism (illustrated in Fig. 7) leads to another process control dynamic, as it can offset the decreasing risk exposure threshold presented in Fig. 6.

5.2. The model's stochastic approach

The proposed model could be turned into a completely stochastic one. As explained earlier with α_λ , in the general model, every failure mode has a probability of occurrence. The deterministic view adopted reflects an average tendency. Fig. 8 presents the simplest case, where every product manufactured has a probability α of being rejected by an acceptance test. The rejected (scrapped) products among the n produced follow a binomial law. At the t th product, there is a non null probability of scrapping every product produced up to that point, and a non null probability of scrapping none of them. There is, then, a non null probability of encountering a major loss, which is represented by the probability of being over-exposed to risk ($MAR > R_L$). The goal of a stochastic model should be to find a strategy for setting control times and allowances that make it possible to remain below the threshold of risk exposure within a pre-defined confidence interval. As capacity is finite and there is a physical measurement, the outcomes of corrective actions are also variable. $\tau_{\lambda k}$ is stochastic. The ATS (Average Time to Signal) and the ARL (Average Run Length) are well-known examples of the intrinsic characteristics of control strategies (control charts). Every time a control is applied, information is provided about processes and production states. However, owing to the ARL_1 and ATS_1 concepts [19], a control may not retrieve the actual characteristics of products, and it can falsely release a fraction of non conforming products or scraps, or falsely investigate good products. This directly impacts the value of $\rho_{k\lambda}$, which ranges between 0 and 1, with a given confidence interval. This suggests a promising research avenue, which would involve building a full stochastic model that mixes quality and product recall prevention.

Material At Risk

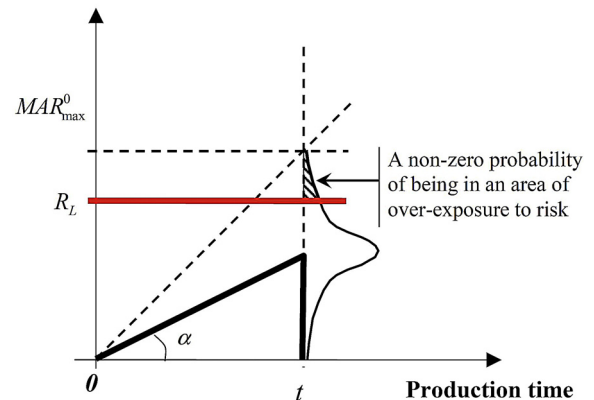


Fig. 8. Stochastic approach.

6. Conclusion

The prevention of recalls and of excessive amounts of scrap is proposed through the careful allocation of quality controls. This paper presents an exposure-based quality control planning approach, which provides an analytical model that has to be solved to plan quality control under capacity and insurance constraints. The resulting control plan ensures that the risk remains below a threshold limit. If a quantity of Material At Risk greater than this value is released onto the market, and if these products are found to be defective, large amounts of product could be scrapped. The paper presents an industrial illustration that employs, in a simplified case, analytical results provided by the model. With this process and quality control in place, teams can determine the level of MAR produced. This model and the proposed algorithm, which can be applied in any job shop manufacturing system, will become a key element in the design of a policy to prevent excessive scrap production.

Authors contributions

BB: main author, design of the algorithm, test and write and revise the paper. MS: developed the simulator for the test. SB: write the paper and perform major revision of the paper, analysis of results and associated discussions, direct the research,

Acknowledgments

The authors are grateful to STMicroelectronics, and especially Philippe Vialletelle, for providing data and his support for their research. This article has been started in the context of the European IMPROVE project (Implementing Manufacturing science solutions to increase equipment productivity and fab performance). IMPROVE was a JTI Project supported by the ENIAC Joint Undertaking contract 12005. It has been finalized with the support of discovery grants funded by NSERC, under the number RGPIN418674-12 CRSNG NIP392730.

Appendix A. Proof of property 1

For better clarity, the machine index z and the failure mode index λ will be removed from this proof.

$$MAR_i = MAR(t_i + \tau_i) = \alpha_i \cdot (t_i + \tau_i - t_{i-1} - \tau_{i-1}) + MAR_{i-1} - \Delta_{i-1}$$

with

$$\Delta_{i-1} = \rho_{i-1} \cdot (MAR_{i-1} - \alpha_{i-1} \tau_{i-1})$$

The MAR reduction Δ_{i-1} , which is effective after the delay τ_{i-1} of the control date t_{i-1} , is proportional to ρ_{i-1} and the value of the MAR at time t_{i-1} ($MAR(t_{i-1}) = MAR_{i-1} - \alpha_{i-1} \tau_{i-1}$). Then, a recursive form of MAR is defined as:

$$MAR_i = MAR_{i-1}(1 - \rho_{i-1}) + \rho_{i-1} \alpha_{i-1} \tau_{i-1} + \alpha_i(t_i + \tau_i - t_{i-1} - \tau_{i-1}) \quad \forall i \in \{1, \dots, n+1\} \quad (A.1)$$

With: $\rho_0 = 1$, $\tau_0 = 0$, $t_0 = 0$ and $t_{n+1} = H$. \square

Appendix B. Proof of property 3

For better clarity, the machine index z and the failure mode index λ will be removed from this proof. Let us consider a control plan $x = (n^x, T^{n^x}, K^{n^x})$. Two cases are possible:

Case 1. x verifies the condition

$$MAR_i^x = MAR_{\max}^x \quad \forall i \in \{1, \dots, n+1\}$$

Let us demonstrate that: $MAR_{\max}^{x'} > MAR_{\max}^x \quad \forall x' \neq x$ with $n^{x'} = n^x$ and $K^{x'} = K^x$. Suppose that x' is a control plan such that $t_i^{x'} = t_i^x - \Delta t_i \quad \forall i \in \{1, \dots, n\}$

$$\begin{aligned} MAR_1^{x'} &= MAR_1^x - \alpha_1 \Delta t_1 = MAR_1^x - \Delta MAR_1 \\ MAR_2^{x'} &= \alpha_2(t_2^{x'} + \tau_2 - t_1^{x'} - \tau_1) + MAR_1^{x'}(1 - \rho_1) + \alpha_1 \rho_1 \tau_1 \\ &= \alpha_2(t_2^x + \tau_2 - t_1^x - \tau_1) - \alpha_2(\Delta t_2 - \Delta t_1) + MAR_1^x(1 - \rho_1) \\ &\quad + \alpha_1 \rho_1 \tau_1 - \alpha_1 \Delta t_1(1 - \rho_1) \\ &= MAR_2^x - (\alpha_2(\Delta t_2 - \Delta t_1) + \alpha_1 \Delta t_1(1 - \rho_1)) \\ &= MAR_2^x - (\alpha_2(\Delta t_2 - \Delta t_1) + \Delta MAR_1(1 - \rho_1)) \end{aligned}$$

Using Eq. (A.1), the following recursive form can be easily demonstrated by recurrence:

$$\begin{aligned} MAR_i^{x'} &= MAR_i^x - (\alpha_i(\Delta t_i - \Delta t_{i-1}) + \sum_{j=1}^{i-1} \alpha_j(\Delta t_j - \Delta t_{j-1}) \prod_{k=j}^{i-1} (1 - \rho_k)) \\ &= MAR_i^x - \Delta MAR_i \quad \forall i \in \{1, \dots, n+1\} \end{aligned}$$

with

$$\Delta MAR_i = (1 - \rho_{i-1}) \Delta MAR_{i-1} + \alpha_i(\Delta t_i - \Delta t_{i-1}) \quad \forall i \in \{1, \dots, n+1\}$$

with $\Delta t_{n+1} = 0$ and $\Delta t_0 = 0$ because $t_{n+1} = H$ and $t_0 = 0$ always. x' dominates x regarding the first variant of the control plan's added value iff $\Delta MAR_i^x > 0 \quad \forall i \in \{1, \dots, n+1\}$

$$\begin{aligned} MAR_{\max}^{x'} < MAR_{\max}^x &\Leftrightarrow \Delta MAR_i^x > 0 \quad \forall i \in \{1, \dots, n+1\} \\ &\Leftrightarrow \Delta t_1 > 0 \\ &\Leftrightarrow \Delta t_i - \Delta t_{i-1} \geq 0 \quad \forall i \in \{2, \dots, n+1\} \end{aligned}$$

Knowing that $\Delta MAR_{n+1} = (1 - \rho_n) \Delta MAR_n - \alpha_i \Delta t_n$

$$\begin{aligned} MAR_{\max}^{x'} < MAR_{\max}^x &\Leftrightarrow \Delta t_n \leq 0 \\ &\Rightarrow \Delta t_{n-1} \leq t_n \leq 0 \quad \text{because } \Delta t_n - \Delta t_{n-1} \geq 0 \\ &\Rightarrow \Delta t_{i-1} \leq t_i \leq 0 \quad \forall i \in \{1, \dots, n\} \\ &\Rightarrow \Delta t_1 \leq 0 \quad \Rightarrow \text{ABSURD because } \Delta t_1 > 0 \end{aligned}$$

Case 2. x does not verify the condition

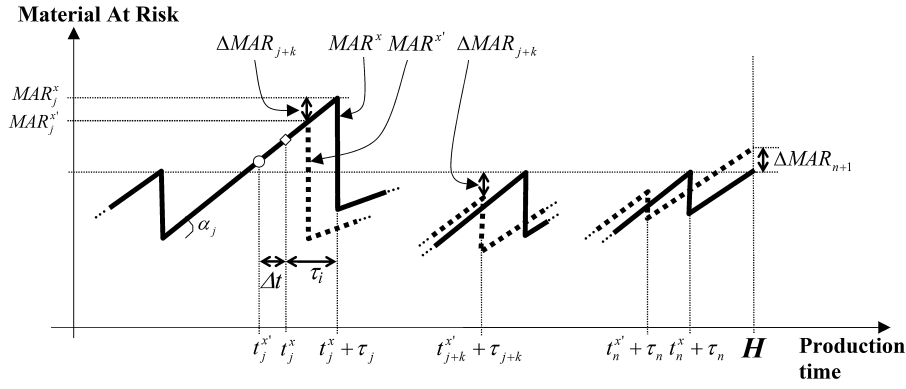
$$MAR_i^x = MAR_{\max}^x \quad \forall i \in \{1, \dots, n+1\}$$

Let us demonstrate that x cannot be optimal (is dominated), because $\exists x' = (n^{x'}, T^{x'}, K^{x'} = K^{n^x})$ such that $MAR_{\max}^{x'} < MAR_{\max}^x$.

In this case (see Fig. 9), there exists at least $j \in \{1, \dots, n\}$ with $MAR_{\max}^x = MAR_j^x$ and $MAR_i^x = cte < MAR_j^x \quad \forall i \in \{a, \dots, n+1\} - \{j\}$

Suppose $x' = (n^{x'} = n^x, T^{x'}, K^{x'} = K^x)$ is another control plan with:

$$\begin{aligned} t_i^{x'} &= t_i^x \quad \forall i \in \{1, \dots, j-1\} \\ &= t_i^x - \Delta t \quad \forall i \in \{j, \dots, n\} \end{aligned}$$

Fig. 9. MAR^x and $MAR^{x'}$.

Then,

$$\begin{aligned} MAR_i^{x'} &= MAR_i^x \quad \forall i \in \{1, \dots, j-1\} \\ MAR_j^{x'} &= MAR_j^x - \alpha_j \Delta t \\ &= MAR_j^x - \Delta MAR_j \end{aligned}$$

$$\begin{aligned} MAR_{j+k}^{x'} &= MAR_{j+k}^x - \alpha_j \Delta t \prod_{k'=j}^{j+k-1} (1 - \rho_{k'}) \quad \forall k \in \{1, \dots, n-j\} \\ &= MAR_{j+k}^x - \Delta MAR_{j+k} \end{aligned}$$

$$\begin{aligned} MAR_{n+1}^{x'} &= MAR_{n+1}^x + \alpha_{n+1} \Delta t - \alpha_j \Delta t \prod_{k=j}^n (1 - \rho_k) \\ &= MAR_{n+1}^x + \Delta MAR_{n+1} \end{aligned}$$

As $\Delta MAR_j > 0$; $\Delta MAR_{j+k} > 0 \quad \forall k \in \{1, \dots, n-j\}$ and $\Delta MAR_{n+1} > 0$ then

$$\begin{aligned} MAR_{\max}^{x'} &= \max \{MAR_{j-1}^x; MAR_j^{x'}; MAR_{n+1}^{x'}\} \\ &= \max \{MAR_{j-1}^x; MAR_j^x - \Delta MAR_j; MAR_{n+1}^x + \Delta MAR_{n+1}\} \end{aligned}$$

Knowing that $\alpha_i = \alpha_\lambda \quad \forall i \in \{1, \dots, n+1\}$, if $\Delta t = \frac{MAR_j^x - MAR_{j-1}^x}{2\alpha_\lambda}$ then

$$\begin{aligned} \Delta MAR_j &= \frac{MAR_j^x - MAR_{j-1}^x}{2} \\ \Delta MAR_{n+1} &= \frac{MAR_j^x - MAR_{j-1}^x}{2} \cdot (1 - \prod_{k=j}^n (1 - \rho_k)) \end{aligned}$$

Then

$$\begin{aligned} MAR_{\max}^{x'} &= \max \{MAR_{j-1}^x; MAR_j^x - \frac{MAR_j^x - MAR_{j-1}^x}{2}; \\ &MAR_{n+1}^x + \frac{MAR_j^x - MAR_{j-1}^x}{2} (1 - \prod_{k=j}^n (1 - \rho_k))\} \\ &= \max \{MAR_{j-1}^x; \frac{MAR_j^x + MAR_{j-1}^x}{2}; \frac{MAR_j^x + MAR_{j-1}^x}{2} \\ &- \frac{MAR_j^x - MAR_{j-1}^x}{2} \prod_{k=j}^n (1 - \rho_k)\} \\ &= \max \{MAR_{j-1}^x; \frac{MAR_j^x + MAR_{j-1}^x}{2}\} \\ &= \frac{MAR_j^x + MAR_{j-1}^x}{2} \end{aligned}$$

Knowing that $MAR_j^x > MAR_{j-1}^x$,

then $MAR_{\max}^{x'} < MAR_{\max}^x$

So, a condition of Eq.(3) is a necessary and sufficient optimality condition for the objective of minimizing MAR_{\max}^x . \square

Appendix C. Proof by recurrence of Eq. (5)

For better clarity, the machine index z and the λ index will be removed from this proof.

The equation is verified for $i=1$ because $t_0^* = 0$ and $\rho_0 = 1$

Suppose that this formula is verified for i

$$\begin{aligned} t_i^* &= t_1^* \left(1 + \alpha_1 \sum_{j=1}^{i-1} \frac{\rho_j}{\alpha_{j+1}} \right) - \sum_{j=1}^{i-1} \frac{1}{\alpha_{j+1}} (\tau_j \rho_j \alpha_j + \alpha_{j+1} (\tau_{j+1} - \tau_j)) \\ &+ \alpha_1 \tau_1 \sum_{j=1}^{i-1} \frac{\rho_j}{\alpha_{j+1}} \end{aligned}$$

Let us demonstrate that the formula is still valid for $j=i+1$ Using Eq. (4)

$$MAR_j^* = \frac{1}{\rho_{j-1}} (\tau_{j-1} \rho_{j-1} \alpha_{j-1} + \alpha_j \cdot (t_j^* + \tau_j - t_{j-1}^* - \tau_{j-1})) \quad (A)$$

and knowing that

$MAR_j^* = MAR_{j-1}^* = MAR_i^* = MAR_1^* = \alpha_1 t_1^* + \alpha_1 \tau_1$ and $t_{j-1}^* = t_i^*$ Eq. (A) becomes

$$\begin{aligned} \alpha_1 t_1^* + \alpha_1 \tau_1 &= \frac{1}{\rho_{j-1}} (\tau_{j-1} \rho_{j-1} \alpha_{j-1} + \alpha_j \cdot (\tau_j - \tau_{j-1})) \\ &+ \frac{\alpha_j}{\rho_{j-1}} \cdot t_j^* - \frac{\alpha_j}{\rho_{j-1}} \cdot t_i^* \end{aligned}$$

Then, the expression of t_j^* is obtained as follows:

$$\begin{aligned}
 t_j^* &= \frac{\rho_{j-1}\alpha_1}{\alpha_j} \cdot t_1^* + t_1^*(1 + \alpha_1 \sum_{i=1}^{j-1} \frac{\rho_i}{\alpha_{i+1}}) \\
 &\quad - \sum_{i=1}^{j-1} \frac{1}{\alpha_{i+1}} (\tau_i \rho_i \alpha_i + \alpha_{i+1} (\tau_{i+1} - \tau_i)) \\
 &\quad + \alpha_1 \tau_1 \sum_{i=1}^{j-1} \frac{\rho_i}{\alpha_{i+1}} + \frac{\rho_{j-1}}{\alpha_j} \alpha_1 \tau_1 \\
 &\quad - \frac{1}{\alpha_j} (\tau_{j-1} \rho_{j-1} \alpha_{j-1} + \alpha_j \cdot (\tau_j - \tau_{j-1})) \\
 &= t_1^* (\frac{\rho_{j-1}\alpha_1}{\alpha_j} + 1 + \alpha_1 \sum_{i=1}^{j-1} \frac{\rho_i}{\alpha_{i+1}}) \\
 &\quad - \sum_{i=1}^{j-1} \frac{1}{\alpha_{i+1}} (\tau_i \rho_i \alpha_i + \alpha_{i+1} (\tau_{i+1} - \tau_i)) \\
 &\quad - \frac{1}{\alpha_j} (\tau_{j-1} \rho_{j-1} \alpha_{j-1} + \alpha_j \cdot (\tau_j - \tau_{j-1})) + \alpha_1 \tau_1 (\sum_{i=1}^{j-1} \frac{\rho_i}{\alpha_{i+1}} + \frac{\rho_{j-1}}{\alpha_j}) \\
 &= t_1^* (1 + \alpha_1 \sum_{i=1}^{j-1} \frac{\rho_i}{\alpha_{i+1}}) - \sum_{i=1}^{j-1} \frac{1}{\alpha_{i+1}} (\tau_i \rho_i \alpha_i + \alpha_{i+1} (\tau_{i+1} - \tau_i)) \\
 &\quad + \alpha_1 \tau_1 \sum_{i=1}^{j-1} \frac{\rho_i}{\alpha_{i+1}}.
 \end{aligned}$$

□

References

- [1] Bean WJ. Variation reduction in a wafer fabrication line through inspection optimization. Massachusetts Institute of Technology; 1997 <http://hdl.handle.net/1721.1/43446>
- [2] Bettayeb B, Bassetto S, Vialletelle P, Tollenaere M. Quality and exposure control in semiconductor manufacturing part i: modelling. *International Journal of Production Research* 2011; <http://dx.doi.org/10.1080/00207543.2011.630042>.
- [3] Bettayeb B, Vialletelle P, Bassetto S, Tollenaere M. Optimized design of control plans based on risk exposure and resources capabilities. In: *International symposium on semiconductor manufacturing (ISSM2010)*. 2010, October. p. 18–20.
- [4] Chen YK, Hsieh KL, Chang CC. Economic design of the VSSI control charts for correlated data. *International Journal of Production Economics* 2007;107(2):528–39.
- [5] Colledani M. Integrated analysis of quality and production logistics performance in asynchronous manufacturing lines. In: *17th IFAC world congress (IFAC'08)*. 2008. p. 8368–74.
- [6] Colledani M, Tolio T. Performance evaluation of production systems monitored by statistical process control and off-line inspections. *International Journal of Production Economics* 2009;120(2):348–67.
- [7] Dautère-Pérès S, Rouveyrol J-L, Yugma C, Vialletelle P. A smart sampling algorithm to minimize risk dynamically. In: *Advanced semiconductor manufacturing conference*. 2010. p. 307–10.
- [8] De Magalhães M, Costa A, Moura Neto F. A hierarchy of adaptive control charts. *International Journal of Production Economics* 2009;119(2):271–83.
- [9] Emmons H, Rabinowitz G. Inspection allocation for multistage deteriorating production systems. *IIE Transactions* 2002;34:1031–41.
- [10] Hassoun M, Rabinowitz G, Lachs S. Identification and Cost Estimation of WIP Bubbles in a Fab. *IEEE Transactions on Semiconductor Manufacturing* 2008;21(2):217–23.
- [11] Hora M, Bapuji H, Roth AV. Safety hazard and time to recall: the role of recall strategy, product defect type, and supply chain player in the U.S. toy industry. *Journal of Operation Management* 2011;7–8:766–77.
- [12] Hsu LF, Tapiero CS. Quality control of the M/G/1 queue. *European Journal of Operational Research* 1989;42(1):88–100.
- [13] Jacobs RM. Product recall, a Vendor/Vendee nightmare. *Microelectronics and Reliability* 1996;36(1):101–3.
- [14] Kim J, Gershwin SB. Integrated quality and quantity modeling of a production line. *OR Spectrum* 2005;27:287–314.
- [15] Kogan K, Raz T. Optimal allocation of inspection effort over a finite planning horizon. *IIE Transactions* 2002;34(6):515–27.
- [16] Lin Y, Chou C. On the design of variable sample size and sampling intervals charts under non-normality. *International Journal of Production Economics* 2005;96(2):249–61.
- [17] Lindsay GF, Bishop AB. Allocation of screening inspection effort – a dynamic-programming approach. *Management Science* 1964;10(2):342–52.
- [18] Mili A, Bassetto S, Siadat A, Tollenaere M. Dynamic risk management unveil productivity improvements. *Journal of Loss Prevention in the Process Industries* 2009;22(1):25–34.
- [19] Montgomery D. *Introduction to statistical quality control*. 6th ed. New York: John Wiley & Sons; 2009.
- [20] Prabhu SS, Montgomery DC, Runger GC. Economic-statistical design of an adaptive X chart. *International Journal of Production Economics* 1997;49(1):1–15.
- [21] Rabinowitz G, Emmons H. Optimal and heuristic inspection schedules for multistage production systems. *IIE Transactions* 1997;29:1063–71.
- [22] Raz T. A survey of models for allocating inspection effort in multistage production systems. *Journal of Quality Technology* 1986;18(4):239–47.
- [23] Reynolds Jr MR, Amin RW, Arnold JC, Nachlas JA. X charts with variable sampling intervals. *Technometrics* 1988;30(2):181–92.
- [24] Schippers WAJ. An integrated approach to process control. *International Journal of Production Economics* 2001;69(1):93–105.
- [25] Straub D. *Generic Approaches to risk based inspection planning for steel structures*. Zurich: NETH15529, Swiss Federal Institute of Technology; 2004 (PhD Dissertation).
- [26] Tagaras G. A survey of recent developments in the design of adaptive control charts. *Journal of Quality Technology* 1998;30(3):212–31.
- [27] Tang K, Tang J. Design of screening procedures: a review. *Journal of Quality Technology* 1994;26(3):209–47.
- [28] Tixier J, Dusserre G, Salvi O, Gaston D. Review of 62 risk analysis methodologies of industrial plants. *Journal of Loss Prevention in the Process Industries* 2002;15(4):291–303.
- [29] Torng C, Lee P, Liao N. An economic-statistical design of double sampling control chart. *International Journal of Production Economics* 2009;120(2):495–500.
- [30] U.S. Department of Defense. *Procedures for performing a failure mode, effects, and criticality analysis*; 1980.
- [31] Verduzco A, Villalobos JR, Vega B. Information-based inspection allocation for real-time inspection systems. *Journal of Manufacturing Systems* 2001;20(1):13–22.
- [32] Villalobos JR, Foster JW, Disney RL. flexible inspection system for serial multistage production systems. *IIE Transactions* 1993;25(3):16–26.
- [33] Wang X, Li D, O'brien C, Li Y. A production planning model to reduce risk and improve operations management. *International Journal of Production Economics* 2010;124(2):463–74.