mathematics

Article

On the Dynamics in Decoupling Buffers in Mass Manufacturing Lines: A Stochastic Approach

Gilberto Pérez-Lechuga [1,*], Francisco Venegas-Martínez [2], Marco A. Montufar-Benítez [1] and Jaime Mora-Vargas [3]

[1] Instituto de Ciencias Básicas e Ingeniería-AAIA, Universidad Autónoma del Estado de Hidalgo, Pachuca de Soto 42184, Mexico; montufar@uaeh.edu.mx
[2] Escuela Superior de Economía, Instituto Politécnico Nacional, Ciudad de Mexico 11350, Mexico; fvenegas1111@yahoo.com.mx
[3] Instituto Tecnológico y de Estudios Superiores de Monterrey, Escuela de Ingeniería y Ciencias, Av Lago de Guadalupe KM 3.5, Margarita Maza de Juárez, Cd López Mateos 52926, Mexico; jmora@tec.mx
* Correspondence: glechuga@uaeh.edu.mx; Tel.: +52-7712367599

Abstract: This paper analyzes the flow of the contents of interleaved buffers with continuously operating machines in a mass production line. Under this framework, the products to be manufactured advance from station to station to receive a physical–chemical transformation that adds value as they progress in the process. The existence of decoupling buffers between operations (between two consecutive workstations) is a common practice in order to alleviate the pressure that is ahead due to the lack of synchronization between consecutive operations, which causes leisure and/or bottlenecks in the system. In this proposal, we analyze the dynamics of a mass manufacturing line with intermediate decoupling buffers. To do that, we use a regenerative stochastic process approach to build a model where the products stored in each buffer are taken all at once by the consecutive machine. In a second approach, we use a homogeneous birth–death process with constant input–output and assume that the products are taken one by one by the consecutive machine. Finally, we use a non-homogeneous birth–death process to analyze the dynamics of a system whose inputs and outputs depend on time. These proposals are accompanied by numerical examples that illustrate its practical utility.

Keywords: stochastic manufacturing; flexible manufacturing systems; decoupling buffers; birth–death processes

MSC: 90B30; 60J27; 60J28

Citation: Pérez-Lechuga, G.; Venegas-Martínez, F.; Montufar-Benítez, M.A.; Mora-Vargas, J. On the Dynamics in Decoupling Buffers in Mass Manufacturing Lines: A Stochastic Approach. *Mathematics* **2022**, *10*, 1686. https://doi.org/10.3390/math10101686

Academic Editors: Aldina Correia, Eliana Costa e Silva and Ana Isabel Borges

Received: 6 April 2022
Accepted: 11 May 2022
Published: 14 May 2022

1. Introduction

Manufacturing is a process carried out in industry in order to change either the consistency, dimension, shape, firmness or beauty of raw material. They are activities that transform raw materials into finished products. Such processes involve automated systems, computers, robots, programmable logic controllers, or autonomous vehicles. As usual, demand is the trigger mechanism of a factory system. The demand states the quantity to be manufactured, the manufacturing time and even the production sequence to be carried out. Other no less important factors that also contribute to establishing the initial conditions of the manufacturing process are, for example, the materials, the equipment, the machinery, the physical–chemical transformation, and the distribution of the plant (layout), among others.

The amount of product that is produced can be determined by the standardization that is made of it (to manufacture it in large volumes), the specific requirements for select clients, the difficulty of the design, the production costs, and the capacity of the plant. These elements determine the type of process to be selected. Therefore, manufacturing is a process where value is added to the product that moves along a production line. As a consequence, the finished product has a higher value than the sum of the value of its components (raw material). There are several types of manufacturing processes, but all of them involve some combination of human labor and machinery. Some of them include the following:

1. Continuous manufacturing, the flow of materials is constant without pause and without any type of transition between operations.
2. Batch production, a standardized method whereby a group of identical products are produced simultaneously (instead of one at a time). It is the Master Production Schedule (MPS) that dictates the size of the lot and the frequency of its production.
3. Production by process, characterized by manufacturing a variety of products in different quantities. Here, the product must visit the machines that will carry out the operations, not necessarily in an ordered sequence due to the complication of moving production equipment.

The modeling of this class of dynamic systems is of great importance for understanding their behavior and, even better, to estimate key parameter values such as the expected quantity of manufactured products, the leisure and the bottlenecks. Therefore, we are interested in studying the behavior of mass production systems with decoupling buffers between machines. In this case, it is common to observe the flow of materials at high speeds (for example, from one hundred to one thousand units per minute). In this type of manufacturing line, there are often temporary storage systems to regulate the feeding of products between machine and machine in order to avoid over-saturation of the line. These systems are called buffer decoupling systems.

Buffers are basic tools for the regular transfer of materials between consecutive stations. Some of the most common uses are: (a) regulate planned stoppages for the replacement of materials, (b) carry out unplanned stoppages due to surrounding line equipment, and (c) have storage time due to the product process. In our proposal, we are interested in the analysis of the interaction between the decoupling and feeding operations of subsequent machines to a buffer in order to characterize the expected production of the line during a given time interval $[0, T]$. To do this, the manufacturing entities are divided into two important elements, the machines and the buffers.

Two cases are analyzed in this proposal. In the first one, we use a format called take all, in which the totality of materials contained in any buffer of the line is fed to the subsequent machine of the process. Here, we used a regenerative stochastic process to model the dynamics of the system. In the second approach, called one-by-one, we assume that the products contained in the buffers are taken individually to feed the subsequent machine. Hence, we use a birth–death process to model the dynamics presented. In both situations, quantitative indicators of the model are obtained, which are illustrated with their respective numerical examples.

2. The Model Background

Fortunately, the literature in this topic is recent and abundant. For example, in [1], an extensive analysis of this type of distributions with diverse variations of the topology of the design are elaborated. An interesting review of the application of mathematical models in the food and beverage industry can be found in [2]. In the proposal developed in [3], the authors create a multi-objective linear integer mathematical model to balance assembly lines. The focus of their work is on optimizing the volume of production, taking care of cycle times and leisure time at work stations.

Regarding the use of buffers in manufacturing lines, an interesting proposal is found in [4]. Here, a detailed analysis of the throughput rate with content-limited buffers is presented. A mathematical model for the productivity rate is built from the technological parameters, the capacity of the buffers and the number of stations and sections with different failure rates and cycle times. The buffers associated with the Work In Process (WIP) are considered in the proposal of [5]. In this case, a model is developed to obtain the optimal WIP stored in the manufacturing buffers at minimum cost by calculating their optimal capacity during the manufacturing and interoperation process. The approach is carried out through mathematical models, which are solved by means of a series of algorithms. The final proposal is validated via simulation, and the results are applied to a car manufacturing company.

A proposal very close to ours can be seen in [6]. The authors analyze the performance of a flexible manufacturing system through the geometric reliability approach. They use Markovian analysis for the study of two machines and find estimators for the distribution, mean and standard deviation, of the process completion time. An advanced model for the study of robotic flexible manufacturing systems (FMS) that include buffers is found in [7]. The authors develop and apply an analytical model for performance in flexible production networks based on autonomous mobile robots (AMR) using a circular loop between workstations and interoperable buffers. In their paper, it is shown that it is possible to avoid line congestion by using multiple crossings and analyzing both the flow and the loading/unloading phases.

An analysis of the impact of Technology 4.0 on FMS can be found in [8]. The importance of Industry 4.0 systems and their autonomous management capacity is shown in this research. Production is based on a scheme based on planning strategies of tolerances (tolerance scheduling problem) to determine the feasible changes in the manufacturing line. Finally, the resequence of the production process in operations associated with late customization through the use of intermediate buffers is also analyzed.

It is important to mention the work developed by [9] on the allocatiion and size of buffers. The authors propose a model for the optimal allocation of buffers in production lines by means of a hybrid algorithm that incorporates the method of nested partitions (NP) and Tabu Search (TS). Through numerical applications, the methods used are promising for the problem of buffer allocation in a large production line. Equally interesting is the approach found in [10]. The research focuses on the problem of buffer size and the positioning of inspection stations on production lines using mixed-integer nonlinear optimization. The technique is illustrated by an instance incorporating n machines and n fixed-size buffers (storage) in series. Under accept–reject conditions, the optimal buffer sizes and the number and positions of inspection stations are found at minimum cost.

Another interesting approach to the problems stated above is using the method of experimental designs. In [11], a performance study of the main factors of the manufacturing systems of a serial manufacturing line is developed. The model includes the reliability and cycle time of workstations, the length of a manufacturing line, and the capacity and location of internal buffers. The analysis is also supported by a simulation model of discrete systems. The combination of both approaches produces good results. Similarly, in [12], a two-stage production system with an intermediate buffer is considered. In the first stage, two similar machines are incorporated in parallel, and in the second, only one machine is considered. The model is analyzed from a simulation perspective to quantify the effect of spare capacity and machine repair rates on production line efficiency.

On the other side, in [13], a buffer is incorporated to study the increase in efficiency in a transfer line where several workstations are linked by a conveyor. Advanced computational techniques have been incorporated to better understand the dynamics that exist in a storage system such as the one described here. The search for learning automata is based in game theory to solve the optimal buffer allocation problem on production lines as in [14]. Here, the proposal is based on the application of the Automata Learning Theory (ALT) to find the optimal buffer sizes in a production line.

In [15], the authors propose a time-based parametric model to calculate a cluster size for a given buffer type. Moreover, they propose an optimal buffer pool policy applicable to an online configuration with deterministic processing time. Finally, in [16], an interesting proposal is developed based on the theory of Markovian processes applied to repairable modular machines. The authors propose a series of composition operators and an improved Universal Generation Function (UGF) vector technique to build the system model. Subsequently, they use a Genetic Algorithm (GA) to solve the optimization mathematical model associated with the buffer capacity. The problem is applied to a production line for engine heads.

3. Setting Up the Mathematical Model of the Process

Normally, a flexible manufacturing line is considered as an adaptable system; however, nothing is further from the truth, since there are a large number of factors associated with it that occur in the form of random events, which makes everything random. Throughout this section, some indicators for both approaches are mentioned.

3.1. The Deterministic Approach

In any development of a representative model, there are some assumptions to achieve a mathematical approximation to the reality of the archetype. In a manual manufacturing line, the operations are fundamentally machining, and the operating times at each station are evaluated by estimating standard times. In an automated line, there are also no exact standard operating times. Usually, these are represented through the mean, variance and other moments that define a probability distribution (pdf). However, it is customary to approximate the average operation times of the work stations by means of punctual values using the arithmetic mean as an unbiased estimator of the sampling times in repetitive operations and use these data as information to obtain other indicators such as:

1. The cycle time of the manufacturing line, T_c, that is defined as the processing time of the slowest station plus the transfer time. Therefore, if τ_i represents the operating time of each station i and τ_{c_i} is the transfer time from machine i to the next machine, then for a total of n machines, the cycle time is given by

$$T_c = \max\{\tau_i\} + \tau_{c_i}.$$

2. The actual time of the production cycle, T_r, that includes the times assigned to unavoidable random stoppages expressed as an average value satisfies

$$T_r = T_c + F_d\, T_o$$

where F_d represents the stoppage frequency in (stops/cycle) and T_o is the downtime per line stop. The above measures lead to simple expressions for the ideal P_i and actual P_r production rates, that is $P_r \approx T_c^{-1}$ and $P_i \approx T_r^{-1}$.

3.2. The Stochastic Approach

Deterministic modeling of the operation of a manufacturing line is an almost impossible task. There are many random quantities involved in this operation, for example: repair times of damaged equipment, installation preparation times (set-up times), mechanical, electrical and electronic failures, different operator skills, etc. Our proposal is focused on a more realistic stochastic modeling of any manufacturing line, manual or automatic. Under this framework, an important indicator is the probability that a manufacturing line is operating satisfactorily at any moment, namely the reliability function. Therefore, the reliability at time t, $R(t)$, of a manufacturing line is a metric such that $R(t) \in (0,1)$; rarely does a system reach the value 1.

Consider a mass production line with intermediate buffers, as shown in Figure 1.

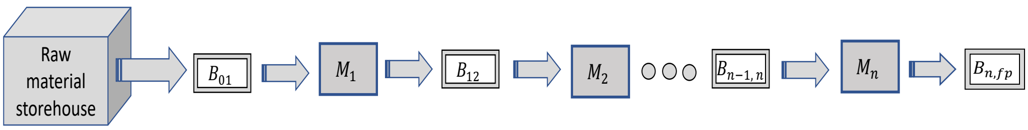

Figure 1. Sequencing of operations with decoupling buffers.

The raw material warehouse is a place where all the products that will be used in the process are stored until they are required. A machine is any mechanical, electrical or electronic device designed and used to carry out some operation or process on a material for a specific product.

A buffer B_{ij} is considered as an intermediate workstation (out of a total of n) between machines i and j, and it is used as temporary storage for the product. Typically, these facilities are located between two consecutive machines and serve to cushion the advance of the product between two stations. Therefore, the overall system can be analyzed as a queuing model embedded in a processing system. The view of a buffer as a queue of finite length is not entirely new. There are several works that have focused with great accuracy on this scheme; see for example [17,18].

In our proposal, we will assume that the buffers satisfy the conditions of the conservation of flux law given by

$$u_t + \psi(u)_x = 0, \tag{1}$$

where $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is the flux and $u : [0, +\infty) \times \mathbb{R} \to \mathbb{R}^n$ is the conserved quantity and where, for any interval $[a, b]$, it is satisfied that [19]

$$\frac{d}{dt} \int_a^b u(t, x)dx = -\int_a^b f(u(t, x))dx = f(u(t, a)) - f(u(t, b)) = \varsigma(t). \tag{2}$$

where $\varsigma(t)$ is a function only of t. Thereby, in the first model, we take up some ideas from [20] to develop our own model associated with a manufacturing line with these characteristics.

Next, we will obtain the dynamics of the products inside the buffers. To do this, we will consider two cases. In the first one, we will suppose that the total content of each buffer is taken by the consecutive machine in a unique way. That is, all the content of the buffer is taken to be processed (takes all), which presupposes that the machine has the capacity to receive the entire content of the buffer in a single emission. In addition, we will use the average times to failure and repair as the elements that trigger the availability of the equipment, which constitute the central axis of our analysis.

In a second model, we will suppose that each piece of the buffer is taken one by one and the consecutive machine processes them in a unitary way (one by one). Here, the triggers of the dynamics of the system are the rate of arrival (α) of the products to the machine M_i, as well as the rate of abandonment (β) to be incorporated downstream in the manufacturing line. As a result, the main indicators contained in the system buffers will be obtained. The main results are as follows:

3.3. Takes All Format

Under this framework, the amount of content in the buffers may vary due to different processing times and random shutdowns due to equipment failure, different operator skills, and random events affecting the system. In order to set up the mathematical model of this system we consider the following aspects:

1. All the raw material for the process comes from a single place called the raw material storehouse. This is a warehouse with infinite capacity.
2. Each machine i has a mean time to failure given by

$$\int_0^\infty x\, dF_X^i(x) = \lambda_i{}^{-1}, \ i = 1, \ldots, n. \tag{3}$$

3. Similarly, the mean time to repair of machine i is given by

$$\int_0^\infty y\, dF_Y^i(y) = \mu_i{}^{-1}, \ i = 1, \ldots, n. \tag{4}$$

4. From (3) and (4), it follows that the unavailability of equipment i is given by

$$U_i = \lambda_i \mu_i{}^{-1}, \ i = 1, \ldots, n. \tag{5}$$

5. Each machine M_i that is part of the system has a maximum productivity rate (measured in products manufactured per unit of time) given by ξ_i. In this proposal, we will assume that if a machine is operating, it always does so at its maximum capacity.

6. The totality of the production moves only in the main manufacturing line, and there is no lost or spilled material.

To begin this analysis, suppose that the quantity of product $\phi(t)$ existing in each buffer $B_{ij}, i < j, j = 1, 2, \ldots, n$, with respect to time t is driven by the stochastic differential equation

$$\frac{d}{dt}\,\phi(t) = \xi_i\,\delta_i(t) - \xi_j\,\delta_j(t) = \zeta(t),\ t \geq 0, \tag{6}$$

here, $\phi(t) \in \left(\frac{-h}{2}, \frac{h}{2}\right)$, and $\phi(0) = 0$. In addition, assume that

$$\delta_i(t) \in \{0, 1\}. \tag{7}$$

In this approach, each buffer B_{ij} has a capacity $h_{ij} \leq \max\{\xi_i, \xi_j\}$, such that the products that arrive at it are formed to be served in a strict First Input, First Output (FIFO) order. When the content of a buffer is a negative quantity, it is said to be starved.

Equations (6) and (7) define a regenerative stochastic process with state space $\{0\}, \{1\}$, where $\delta_i(t) = 1$ means that the corresponding machine is working, and zero means that it is not. Let $\pi(\vartheta, t)$ be the probability of being in the ϑ-state. If we assume that $F_X(x)$ and $F_Y(y)$ are exponentially distributed under steady-state conditions, as in [21,22], then

$$\pi(0,t) = \lim_{t\to\infty}\left[\pi(0,0)\,(1-\lambda-\mu)^t + \frac{\lambda}{\mu+\lambda}\left[1-(1-\lambda-\mu)^t\right]\right] = \frac{\lambda}{\mu+\lambda}, \tag{8}$$

and

$$\pi(1,t) = \lim_{t\to\infty}\left[\pi(1,0)\,(1-\lambda-\mu)^t + \frac{\mu}{\mu+\lambda}\left[1-(1-\lambda-\mu)^t\right]\right] = \frac{\mu}{\mu+\lambda}, \tag{9}$$

thus

$$P(\delta_i(t) = 1) = \pi(1,t),\ P(\delta_i(t) = 0) = \pi(0,t). \tag{10}$$

The above assumption leads to the noise $\zeta(t)$, which is a Markov process with state space Ω defined as

$$\Omega = \{\xi_i - \xi_j,\ -\xi_j,\ \xi_i,\ 0\}, \tag{11}$$

where

$$P(\xi_i - \xi_j) = \pi_i(1)\,\pi_j(1) = \frac{\mu_i\,\mu_j}{(\mu_i+\lambda_i)(\mu_j+\lambda_j)} = \frac{\mu_i\,\mu_j}{\varrho}, \tag{12}$$

$$P(-\xi_j) = \pi_i(0)\,\pi_j(1) = \frac{\lambda_i\,\mu_j}{(\mu_i+\lambda_i)(\mu_j+\lambda_j)} = \frac{\lambda_i\,\mu_j}{\varrho}, \tag{13}$$

$$P(\xi_i) = \pi_i(1)\,\pi_j(0) = \frac{\mu_i\,\lambda_j}{(\mu_i+\lambda_i)(\mu_j+\lambda_j)} = \frac{\mu_i\,\lambda_j}{\varrho}, \tag{14}$$

$$P(0) = \pi_i(0)\,\pi_j(0) = \frac{\lambda_i\,\lambda_j}{(\mu_i+\lambda_i)(\mu_j+\lambda_j)} = \frac{\lambda_i\,\lambda_j}{\varrho}. \tag{15}$$

Hence, there is an initial probability vector Π_0^* given by

$$\Pi_0^* = \varrho^{-1}\left[\mu_i\,\mu_j,\ \lambda_i\mu_j,\ \mu_i\,\lambda_j,\ \lambda_i\,\lambda_j\right], \tag{16}$$

where Π^* means transpose of Π. Similarly, we have that the infinitesimal transition probabilities $\hat{Q}$ is defined as

$$\hat{Q} = \begin{pmatrix} -(\lambda_i+\lambda_j) & \mu_i & \mu_j & 0 \\ \lambda_i & -(\mu_i+\lambda_j) & 0 & \mu_j \\ \lambda_j & 0 & -(\mu_j+\lambda_i) & \mu_i \\ 0 & \lambda_j & \lambda_i & -(\mu_i+\mu_j) \end{pmatrix}. \tag{17}$$

An estimator for the expected value of the buffer contents for all t is given by

$$\mu_{\phi(t)} = E[\phi(t)] = \varrho^{-1}\left[(\xi_i - \xi_j)(\mu_i\mu_j) - \xi_j(\lambda_i\mu_j) + \xi_i(\mu_i\lambda_j)\right], \tag{18}$$

and the variance satisfies

$$\sigma^2_{\phi(t)} = \varrho^{-1}\left[(\xi_i - \xi_j)^2(\mu_i\mu_j) + \xi_j^2(\lambda_i\mu_j) + \xi_i^2(\mu_i\lambda_j)\right] - E^2[\phi(t)]. \tag{19}$$

Derived from the above, we have that the total expected production φ_i of each machine i during the planning horizon T can be obtained by adding the partial productivities generated during the k production cycles, when process (6) is in state 1, that is (see Figure 2),

$$\varphi_i = \sum_{l=1}^{k}\left(\frac{\mu_l}{\mu_l + \lambda_l}\right)h_l, \quad i = 1, \ldots, n \tag{20}$$

The expected total production of the line is given by

$$M_{total} = \sum_{i=1}^{n}\varphi_i = \sum_{i=1}^{n}\sum_{l=1}^{k}\left(\frac{\mu_l}{\mu_l + \lambda_l}\right)h_l. \tag{21}$$

The number of average empty slots S_e in each B_{ij} buffer is given by

$$S_e = \begin{cases} h_i - E[\phi(t)], & \text{if } E[\phi(t)] \geq 0, \\ h_i, & \text{otherwise.} \end{cases} \tag{22}$$

This model constitutes an important manufacturing case where a machine can be fed in a single way. That is, to start its operation, it considers taking all the material that precedes it in its operation. A numerical example for this situation will be shown later.

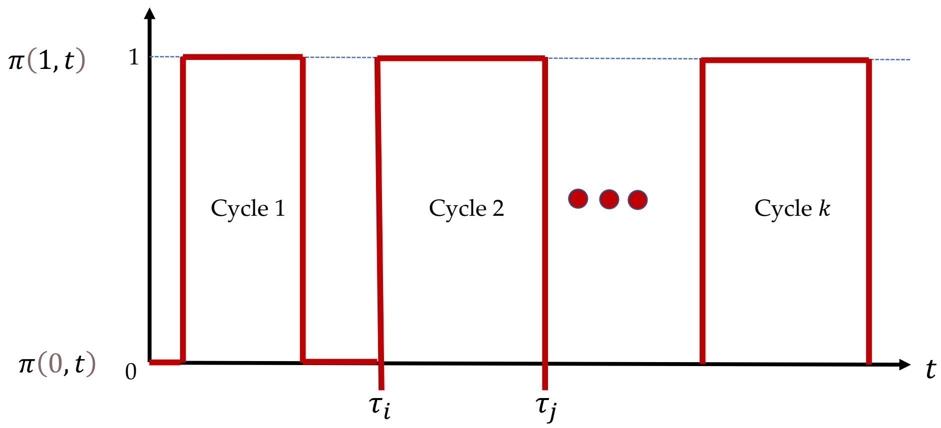

Figure 2. Alternative process cycles $\frac{d}{dt}\,\phi(t)$.

3.4. One by One Format

In this alternative model, we consider that the elements of the buffer content are dislodged from it one by one by the subsequent machine. In addition, each machine is always available when it is requested by the production line.

3.4.1. The Birth-and-Death Model

Let α_i and β_i be the number of pieces that arrive and leave machine i per unit of time, respectively, and let $\mathcal{S}$ be an integer interval (finite or infinite). Then, the process of arrivals and departures of products to a certain buffer can be seen as a Markov chain with subspace $\mathbb{Z}$ such that $\alpha \to [0, \infty)$ and $\beta \to [0, \infty)$. Again, we will suppose that $h \leq \max(\alpha, \beta)$.

Let $X(t) = X_t$ be the number of pieces in buffer B_{ij} at time t, and let $P_x(t) = P\{X(t) = x\}$, $x = 0, 1, \ldots$ be the probability associated with the length of the queue at time t.

We will also assume the following:

1. The probability of transition $x \to x + 1$ in the interval $(t + \Delta t)$ is $\alpha_x\Delta t + o(\Delta t)$;
2. The probability of transition $x \to x - 1$ in the interval $(t, t + \Delta)$ is $\beta_x\Delta t + o(\Delta t)$;
3. The probability of transition to another neighboring state is $o(\Delta t)$;

4. The probability of no change is $1-(\alpha_x+\beta_x)\Delta t+o(\Delta)$.

Therefore, the system can be seen as a Markov process in continuous time of the type birth-and-death, i.e.,

$$\begin{aligned}P_x(t+\Delta t)=\alpha_{x-1}\,P_{x-1}(t)\,\Delta t+[1-(\alpha_x+\beta_x)\,\Delta t]P_x(t)+\\ \beta_{x+1}P_{x+1}(t)\Delta t+o(\Delta t).\end{aligned} \tag{23}$$

A reasonable assumption in this model is that $\alpha>\beta$, i.e., no backorders are allowed. The above result is associated with the following Kolmogorov equation

$$\frac{dP_x(t)}{dt}=\alpha_{x-1}\,P_{x-1}(t)-(\alpha_x+\beta_x)\,P_x(t)+\beta_{x+1}\,P_{x+1}(t),\ x=1,2,\ldots$$

and, in particular for $x=0$, we have

$$\frac{dP_0(t)}{dt}=\beta_1P_1(t).$$

Furthermore, if it is satisfied that in the initial state, the system is in the state $x=x_0$, $0<x_0<\infty$, then the initial conditions are $P_x(0)=1$ for $x=x_0$, and $P_x(0)=0$ otherwise. Without loss of generality, we will assume that $\lambda_x=\lambda$ and $\mu_x=\mu$. Using the generating function of the probabilities, $F(s,t)=\sum_{x=0}^{\infty}P_x(t)s^x,|\ x\ |\leq 1$, we have as in [23]

$$\frac{\partial F(s,t)}{\partial t}=\left[\alpha_s^2-(\alpha+\beta)s+\beta\right]\frac{\partial F(s,t)}{\partial s},$$

whose general solution is given by

$$F(s,t)=f\left(\frac{\beta-\alpha s}{1-s}\right)e^{-(\alpha-\beta)t}.$$

Note that when $X(0)=x_0=1$, then $F(s,0)=s$. Therefore,

$$s=f\left(\frac{\beta-\alpha s}{1-s}\right),$$

and

$$f(\xi)=\frac{\beta-\xi s}{\alpha-\xi},$$

where f is an arbitrary function. Hence,

$$F(s,t)=\frac{\beta\left(1-e^{(\alpha-\beta)t}\right)-\left(\alpha-\beta e^{(\alpha-\beta)t}\right)s}{\beta-\alpha e^{(\alpha-\beta)t}-\alpha\left(1-e^{(\alpha-\beta)t}\right)s}. \tag{24}$$

The coefficients of s^x are given by (using the expansion method in power series)

$$P_x(t)=\begin{cases}[1-\gamma(t)][1-\kappa(t)][\kappa(t)]^{x-1},\ x=1,2,\ldots\\ 0,\ x=0.\end{cases} \tag{25}$$

where

$$\gamma(t)=\frac{\beta\left(e^{(\alpha-\beta)t}-1\right)}{\alpha e^{(\alpha-\beta)t}-\beta}, \tag{26}$$

$$\kappa(t)=\frac{\alpha\left(e^{(\alpha-\beta)t}-1\right)}{\alpha e^{(\alpha-\beta)t}-\beta}, \tag{27}$$

$$E\{X_t\}=e^{(\alpha-\beta)t}, \tag{28}$$

and

$$Var\{X_t\}=\sigma_{X_t}^2=\left[\frac{\alpha+\beta}{\alpha-\beta}\right]\left[e^{(\alpha-\beta)t}\right]\left[e^{(\alpha-\beta)t}-1\right]. \tag{29}$$

An estimator for the transition matrix Q is given by

$$Q(x) = \begin{cases} p(x) = \frac{\alpha(x)}{\alpha(x)+\beta(x)}, & \text{if } x \to x+1 \\ q(x) = \frac{\beta(x)}{\alpha(x)+\beta(x)}, & \text{if } x \to x-1 \\ r(x) = 0, & \text{if } x \to x \end{cases} \tag{30}$$

If $p(x) = q(x) = 0$, then $r = 1$. That is, the state x is absorbing. Therefore, the generating matrix $\hat{Q}$ is given by

$$\hat{Q}(x) = \begin{cases} -[\alpha(x)+\beta(x)], & \text{if } x \to x \\ \alpha(x), & \text{if } x \to x+1 \\ \beta(x), & \text{if } x \to x-1 \end{cases} \tag{31}$$

Two important measures in the above context are the bottleneck and the leisure generated in the system. Leisure is defined as the time when one machine or a set of machines are free from work or other duties. In this topology, leisure in the B_{ij} buffer occurs when the number of α products coming from machine i is less than the number of β products that will access machine j. Similarly, a bottleneck is a point of congestion in a production system. This occurs in an assembly when workloads arrive too quickly for the production process to handle. In our case, a bottleneck occurs in buffer B_{ij} when the number of α products coming from machine i is greater than the number of β products that will access machine j. Formally, the buffer B_{ij} can be in any of the following states

$$\text{Buffer status}\,(B_{ij}) = \begin{cases} \text{Leisure } (\mathcal{L}), & \alpha < \beta \\ \text{Bottleneck } (\mathcal{B}), & \alpha > \beta \\ \text{Steady flow}, & \alpha = \beta \end{cases} \tag{32}$$

From (28), we have

$$\lim_{t\to\infty} E\{X_t\} = \begin{cases} 0, & \text{for } \alpha < \beta \\ 1, & \text{for } \alpha = \beta \\ \infty, & \text{for } \alpha > \beta \end{cases} \tag{33}$$

During the product movement in the buffers, it may happen that they become empty or become oversaturated due to the large amount of product accumulated in them. In a birth–death process, these phenomena are known as extinction and population explosion. Formally, in this analysis, any of the following events may occur:

1. If $X_s = 0$ for some $s \in \mathcal{S}$, and hence $X_t = 0,\ \forall\ t \in [s,\infty)$, then the phenomenon of extinction of the population is verified. The probability of emptying the buffer at time t (extinction probability) is obtained from Equations (25)–(27) in the following way:

$$\lim_{t\to\infty} P_0(t) = \lim_{t\to\infty} \left[\frac{\beta\left(e^{(\alpha-\beta)t} - 1\right)}{\alpha e^{(\alpha-\beta)t} - \beta} \right] = \begin{cases} 1, & \text{for } \alpha < \beta \\ (\beta/\alpha), & \text{for } \alpha > \beta \end{cases} \tag{34}$$

2. Similarly, if $X_t \to \infty$, as $t \to \infty$, $(T = \infty)$, then the phenomenon of population explosion is verified. The probability of explosion, p_e, is given by

$$p_e = 1 - \lim_{t\to\infty} P_0(t) \tag{35}$$

Finally, we are interested in the absorption probability function, i.e., $h(x) = P(X_t = 0$ for some $t \in [0,\infty) \mid X_0 = 0, T < \infty)$. This is given by

$$h(x) = \frac{1}{q} \sum_{i=0}^{x-1} \frac{\beta(1)\,\beta(2)\ldots\beta(i)}{\alpha(1)\,\alpha(2)\ldots\alpha(i)} \tag{36}$$

where

$$q = \sum_{i=0}^{\infty} \frac{\beta(1)\,\beta(2)\ldots\beta(i)}{\alpha(1)\,\alpha(2)\ldots\alpha(i)} \tag{37}$$

It is interesting to note the similarity between Equations (34) and (36). One more important result is the following:

$$\lim_{t\to\infty} E\{X_t\} \to x_0, \text{ if } \alpha = \beta$$

The application of this model assumes the existence of constant rates of arrival and departure of the system. In order to illustrate its application, a numerical example will be carried out later.

3.4.2. The Non-Homogeneous Birth-and-Death Model

In practice, the α_i arrival and the β_i departure rates of products on the machine i are not the same at any instant of time. These change according to the variable speeds of the equipment, the demand requirements and the Master Production Plan (MPS). In this case, these rates should be expressed as a function of the time in which they will be required. Formally, we now define $\alpha_i(t)$ and $\beta_i(t)$ as the number of pieces arriving and leaving the machine M_i, respectively, depending on their observation time. Again, let $X(t) = X_t$ be the number of pieces in buffer B_{ij} at time t. Now, we will assume that the rates are arbitrary functions of time of the state variable. The adjustments to the previous model are based in [24] and reported in [25]. In this case, we have

$$\gamma(t) = 1 - \frac{e^{-\omega(t)}}{w(t)}, \tag{38}$$

$$\kappa(t) = 1 - \frac{1}{w(t)}, \tag{39}$$

where

$$\omega(t) = \int_0^t [\beta(\tau)] - \alpha(\tau)]d\tau, \tag{40}$$

and

$$w(t) = e^{-\omega(t)}\left[1 + \int_0^t \beta(\tau)\, e^{\omega(\tau)} d\tau\right]. \tag{41}$$

The mean and variance, respectively, are now given by

$$E\{X_t\} = e^{-\omega(t)}, \tag{42}$$

$$Var\{X_t\} = \sigma^2_{X_t} = e^{-2\omega(\tau)} \int_0^t [\beta(\tau)] + \alpha(\tau)]\, e^{\omega(\tau)} d\tau. \tag{43}$$

Under these conditions, the probability of extinction of pieces in each buffer is given by

$$P_0(t) = \frac{\int_0^t \beta(\tau)\, e^{\omega(\tau)} d\tau}{1 + \int_0^t \beta(\tau)\, e^{\omega(\tau)} d\tau}. \tag{44}$$

The intensity function and the transition probabilities are, respectively, given by

$$q_i = (\alpha + \beta)i, \;\; i = 1, 2 \ldots \tag{45}$$

and

$$Q_{ij} = \begin{cases} \frac{\alpha}{\alpha+\beta}, & \text{for } j = i+1,\, i = 1, 2, \ldots \\ \frac{\beta}{\alpha+\beta}, & \text{for } j = i-1, \\ 0, & \text{for } \mid i - j \mid > 1 \end{cases} \tag{46}$$

The infinitesimal matrix of transition probabilities satisfies

$$\mathcal{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \beta & -(\alpha+\beta) & \alpha & 0 & 0 & \cdots & 0 & 0 \\ 0 & 2\beta & -2(\alpha+\beta) & 2\alpha & 0 & \cdots & 0 & 0 \\ 0 & 0 & 3\beta & -3(\alpha+\beta) & 3\alpha & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \end{bmatrix} \tag{47}$$

The corresponding Kolmogorov equations are given by

$$\frac{dP_{ij}}{dt} = \alpha_{j-i}P_{i,j-1}(t) - (\alpha_j + \beta_j)P_{ij}(t) + \beta_{j+1}P_{i,j+1}(t), \tag{48}$$

$$\frac{dP_{ij}}{dt} = \alpha_i P_{i+1,j}(t) - (\alpha_i + \beta_i)P_{ij}(t) + \beta_i P_{i-1,j}(t), \tag{49}$$

and using the generating function, we have that

$$(F(s,t))^i = \left[\frac{\gamma(t) + [1-\gamma(t) - \kappa(t)s]}{1-\kappa(t)s}\right]^i, \tag{50}$$

where

$$[\gamma(t) + (1-\gamma(t)-\kappa(t))s]^i = \sum_{n=0}^{\infty} \binom{i}{n} [\gamma(t)]^i \, [1-\gamma(t)-\kappa(t)]^i \, s^i,$$

and

$$[1-\kappa(t)]^{-i} = \sum_{n=0}^{\infty} \binom{i+n-1}{n} [\kappa(t)]^i \, s^i, \quad | \, \kappa(t)s \, | < 1.$$

Therefore,

$$P_{ij}(t) = \sum_{n=0}^{i} \binom{i}{n} \binom{i+j-n-1}{i-1} [\gamma(t)]^{i-n} [\kappa(t)]^{j-n} [1-\gamma(t)-\kappa(t)]^n, \;\; i \geq j. \tag{51}$$

The above equation is useful for defining a set of Ω states of the transition matrix and calculating other properties of the system such as absorption first passage time. For the recurrence time, an important result in [26] is the following. If $H_{ij}(t)$ and $H_{ii}(t)$ are the first-passage time and recurrence time distribution of the process $X(t)$, then

$$\Upsilon = \int_0^{\infty} dH_{ii}(t),$$

defines the probability that the system starts at i and ends at i during a finite period of time.

An interesting extension to birth–death processes is when there exists a jump rate expressed as an asymptotic polynomial dependence on the position of the process; see, for example, [27]. The authors obtain an approximate asymptotic exponential distribution for the probability of excursions of a rescaled process contained within a neighborhood of a given non-negative continuous function.

4. Numerical Examples

Below are three numerical examples associated with the cases analyzed in this document. First, an instance associated with the model of the 'takes all formats' type is shown. The second refers to the one-by-one format model. In this case, we show two examples associated; the first one uses the homogeneous birth–death model and the second one uses the non-homogeneous case. Each subsection corresponds to the previously developed models in the same order of presentation. The results obtained are the following.

4.1. The Case for the Format Takes All

In order to show a numerical example with this format, we assumed a sequence of six machines with five intermediate buffers. The operation dynamics of this model is as follows:

1. Enter the known values of h_i, for all B_{ij} $i = 1, \ldots, n$;
2. Enter the known values of ξ_i, λ and μ for all M_i $i = 1, \ldots, n$;
3. For all M_i, $i = 1, \ldots, n$, obtain the magnitudes:

$$\mu_{\phi(t)},\ \sigma^2_{\phi(t)},\ S_e,\ U,\ \varphi;$$

4. Obtain the states $\mathcal{S}$ of the process: leisure ($\mathcal{L}$) or bottleneck ($\mathcal{B}$).

The presentation is made in three blocks in Table 1. The first block shows the numerical values assigned to the instance. The second block shows the states of Ω. Finally, in the third block, we show the numerical values obtained using our proposal.

For the block Θ_1, the following values are proposed (typical quantities in a metal-working manufacturing system). The unit of measurement of time is hours; therefore, $\lambda = 1/1200$ means one failure every 1200 h, and $\mu = 1/25$ is interpreted as one repair every 25 h. The values of ξ were arbitrarily proposed and represent the maximum number of pieces processed per hour on each machine M_i, $i = 1, \ldots, 6$. Here, $\theta_{ij} = \xi_i - \xi_j$, and $M_{tot} = 37$. Finally, h_{ij} represents the capacity of the buffer B_{ij} in pieces processed per hour.

Table 1. Parametric values and results obtained from the takes all format.

Θ_1	M_1	B_{12}	M_2	B_{23}	M_3	B_{34}	M_4	B_{45}	M_5	B_{56}	M_6
h		6		8		8		7		8	
ξ	6		5		8		7		4		8
λ	$\frac{1}{1200}$		$\frac{1}{1150}$		$\frac{1}{1000}$		$\frac{1}{1060}$		$\frac{1}{1200}$		$\frac{1}{1100}$
μ	$\frac{1}{25}$		$\frac{1}{32}$		$\frac{1}{27}$		$\frac{1}{29}$		$\frac{1}{28}$		$\frac{1}{34}$
Θ_2	M_1	B_{12}	M_2	B_{23}	M_3	B_{34}	M_4	B_{45}	M_5	B_{56}	M_6
θ_{ij}		1		-3		1		3		-4	
$-\xi_j$		-5		-8		-7		-4		-8	
ξ_i		6		5		8		7		4	
0		0		0		0		0		0	
$\mathcal{S}$		$\mathcal{B}$		$\mathcal{L}$		$\mathcal{L}$		$\mathcal{L}$		$\mathcal{B}$	
Θ_3	M_1	B_{12}	M_2	B_{23}	M_3	B_{34}	M_4	B_{45}	M_5	B_{56}	M_6
$\mu_{\phi(t)}$		1.017		-2.901		0.096		2.910		-3.851	
$\sigma_{\phi(t)}$		1.162		1.597		1.679		1.278		1.489	
S_e		4.982		8.000		7.038		4.089		8.000	
U	0.020		0.027		0.027		0.027		0.023		0.030
φ	5.882		4.864		7.789		6.8136		3.988		7.760

The Θ_2 block represents the state space associated with the Markovian process $\phi(t)$ (Equation (11)) once the values of the block Θ_1 have been replaced. The value of $\mathcal{S}$ defines the buffer positions as leisure ($\mathcal{L}$) or bottleneck ($\mathcal{B}$).

In the Θ_3 block, the row assigned to $\mu_{\phi(t)}$ represents the average value of the contents of the system buffers, which is the expected value of the Markovian process $\phi(t)$ (Equation (18)). It is important to note that buffers B_{12}, B_{34} and B_{45} contain a positive value. For example, the case of buffer B_{12} is interpreted as follows. Machines M_1 and M_2 have respective processing capacities of six pieces/h and five pieces/h, respectively. The buffer B_{12} will have an average content of 1017 pieces/h. However, the processing capacity of M_3 is eight pieces/h, that is, a processing capacity greater than M_2; therefore, there will be leisure in the buffer B_{23}, since it remains starved for material that does not arrive. This situation is reflected in its average content—that is, -2.92 pieces/h. The rest of the analysis for this row is done with the same logic.

The row associated with $\sigma_{\phi(t)}$ represents the standard deviation of the $\phi(t)$ process (Equation (19)). It is important to highlight that these quantities are very similar, which defines little variability of the process. In the case of the S_e variable, it defines the number of empty slots in each B_{ij} buffer. Special cases are buffers B_{23} and B_{56} (Equation (22)). Similarly, the variable U (Equation (5)) represents the unavailability of the machine M_i, i.e., a measure of the probability that such a machine is offline and φ is the expected production in the system.

Another interesting measure of the activity of each machine is given by Equations (12)–(14), that is, the state probabilities. Table 2 shows these results.

Table 2. Probability distribution for the process $\phi(t)$ in each buffer.

State (i)	M_1	B_{12}	M_2	B_{23}	M_3	B_{34}	M_4	B_{45}	M_5	B_{56}	M_6
$\xi_i - \xi_j$		0.9531		0.9473		0.9478		0.9512		0.9479	
$-\xi_j$		0.0199		0.0264		0.0256		0.0260		0.0221	
ξ_i		0.0265		0.0256		0.0259		0.0222		0.0293	
0		0.0006		0.0007		0.0007		0.0006		0.0007	

The components of the matrix represent the probability of the state s_i, $i = 1, \dots, 4$ in the B_{ij} buffer. For example, for buffer B_{23}, we have $P(\xi_i - \xi_j) = 0.9473$, and the rest is the same. The values found reveal a high activity of the state $\xi_i - \xi_j$. Finally, from the values given in this instance, we obtain the initial vector of probabilities (Equation (16))

$$\Pi_0^{B_{12}^*} = [0.9531, 0.0199, 0.0265, 0.0006],\ \Pi_0^{B_{23}^*} = [0.9473, 0.0264, 0.0256, 0.0007]$$

$$\Pi_0^{B_{34}^*} = [0.9478, 0.0256, 0.0259, 0.0007],\ \Pi_0^{B_{45}^*} = [0.9512, 0.0260, 0.0222, 0.0006]$$

$$Pi_0^{B_{56}^*} = [0.9479, 0.0221, 0.0293, 0.0007].$$

4.2. The Case One by One

For the analysis of these instances, we proceed as before, separating the corresponding subsections in the order in which they were previously exposed.

4.2.1. The Case of the Homogeneous Birth–Death Model

In order to illustrate the second model numerically, we define a punctual estimator of the content of the buffers (independent of t), and the following statistic is proposed

$$\bar{y} = \frac{1}{T} \sum_{t=1}^{T} E\{X(t)\}. \tag{52}$$

Similarly, for the standard deviation of the content of the buffer B_{ij}, we propose the statistic $\bar{z}$ defined as

$$\bar{z} = \frac{1}{T} \sum_{t=1}^{T} \sigma_{X_t}. \tag{53}$$

For the exploration of this instance, we use two methods. In the first, the expected value of $P_X(t)$ is proposed as the content of the buffer B_{ij}. In the second, the dynamic input–output, Work In Process (WIP) and Work In Buffers (WIB) are used. For this case, we define the Work In Process (WIP) as the raw materials, labor, and overhead costs incurred for products that are at various stages of the production process. Similarly, the Work In Buffers (WIB) refers to the amount of material stored in buffer B_{ij} at time t of the planning horizon.

The dynamics of operation in the first approach is as follows

1. Enter the known values of h_i, for all B_{ij}, $i = 1, \dots, n$;
2. Enter the known values of ξ_i, α and β for all M_i $i = 1, \dots, n$;
3. For all t and for all M_i, $i = 1, \dots, n$, obtain the magnitudes:
   (a) $\beta\left(e^{(\alpha-\beta)t} - 1\right)$, $\alpha\left(e^{(\alpha-\beta)t} - 1\right)$, $\alpha e^{(\alpha-\beta)t} - \beta$
   (b) $\kappa(t)$, $\gamma(t)$, $P_x(t)$, $E\{X(t)\}$, σ_{X_t};

4. For each buffer B_{ij}, estimate the approximate values of $\bar{y}$ and $\bar{z}$, $P_0(t)$ and $\mathcal{S}$;
5. Assign to the buffer $B_{n,fp}$ the expected value of the production of the system.

Thus, for a $T = 100$, we have the numerical values for this case in Table 3.

Table 3. Parametric values and results obtained from the one by one format.

Θ_1	M_1	B_{12}	M_2	B_{23}	M_3	B_{34}	M_4	B_{45}	M_5	B_{56}	M_6	$B_{6,fp}$
h		6.80		6.75		6.72		6.71		6.70		6.65
ξ	10		10		10		10		10		10	
α	6.80		6.70		6.66		6.59		6.57		6.56	
β	6.70		6.68		6.60		6.58		6.56		6.50	
$\bar{y}$		55.81		3.22		69.10		1.72		1.72		69.10
$\bar{z}$		642.56		68.30		1019.42		39.43		39.43		1011.70
P_0		0.9853		0.9970		0.9910		0.9985		0.9985		0.9909
$\mathcal{S}$		$\mathcal{B}$		$\mathcal{L}$		$\mathcal{B}$		$\mathcal{L}$		$\mathcal{L}$		

From the values proposed for α and β, it is observed that in the row of states $\mathcal{S}$, there are two bottlenecks ($\mathcal{B}$) in buffers B_{12} and B_{34}. Similarly, buffers B_{23}, B_{45} and B_{56} have respective leisures ($\mathcal{L}$). The content of the buffer $B_{6,fp} = 69.10$ represents the amount of finished product (the expected total production) when the system is in steady-state conditions. Figure 3 shows graphically the probability distribution $P_X(t)$ for $t = 30$ h of operation in machine 1 for $x = 1, 2, 3$.

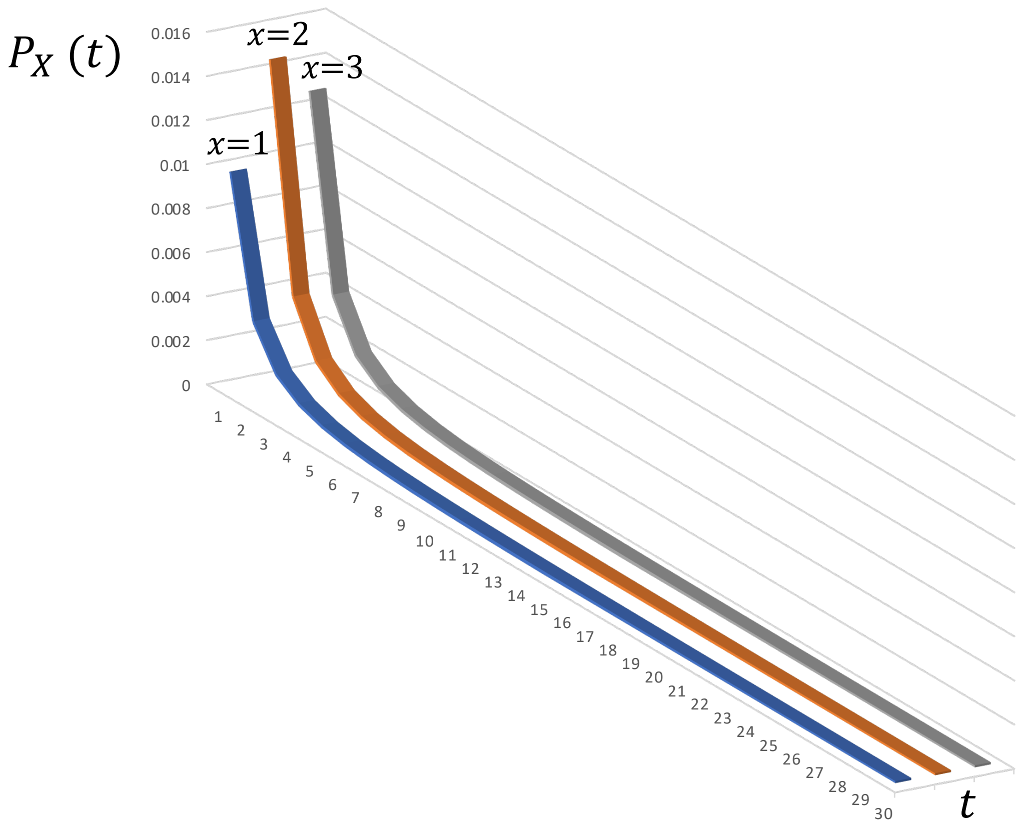

Figure 3. Probability distribution $P_X(t)$ for $t = 30$ h of operation in machine 1 for $x = 1, 2, 3$.

The dynamics of operation for the second approach is as follows:

1. Enter the known values of h_i, for all B_{ij} $i = 1, \ldots, n$;
2. Enter the known values of ξ_i, α and β for all M_i $i = 1, \ldots, n$;

3. At $t = 0$, the raw material arrives at the machine M_1 at a rate given by the function α_1;
4. The number of products leaving machine one during the first hour is given by the function β_1;
5. The output rate of products (s_1) from machine 1 is obtained from the quantity

$$s_1(t) = \min\{\xi_1, \beta_1\}; \tag{54}$$

6. The WIP of the first machine is defined as the difference between what is supported and what is delivered as finished product.

$$WIP_1 = \alpha_1 - s_1;$$

7. Similarly, the WIB_1 is given by

$$WIB_1 = \min\{s_1, h_1\};$$

8. The WIB_1 and the quantity ξ_2 constitutes the input to M_2 in the following form

$$\alpha_2 = \min\{WIB_1, \xi_2\}.$$

Therefore, the process is repeated updating the respective indices.

Table 4 shows a concentration of the results found with the parameters described.

Table 4. Average values for the homogeneous birth–death process for $T = 100$ h.

Θ	M_1	B_{12}	M_2	B_{23}	M_3	B_{34}	M_4	B_{45}	M_5	B_{56}	M_6	$B_{6,fp}$
h		6.80		6.75		6.72		6.71		6.70		6.65
ξ	10		10		10		10		10		10	
α	12.000		6.161		6.121		6.096		6.080		6.080	
β	16.000		16.000		16.000		16.000		16.000		16.000	
s	6.700		6.700		6.700		6.700		6.700		6.700	
WIP	3.451		−2.387		−2.427		−2.451		−2.467		−2.467	
WIB		6.161		6.121		6.096		6.080		6.080		6.040
$\mathcal{S}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$

The expected final quantity of production of this model is given by

$$M_{tot} = \sum_{i=1}^{6} WIB_i = 36.578$$

and $s_i = \min\{\alpha_t, \beta_t\}, i = 1, \ldots, n$, stands for the amount emitted from machine i at time t.

4.2.2. The Case of the Non-Homogeneous Birth–Death Model

In order to illustrate the non-homogeneous birth–death model, consider the following $\alpha(t)$ and $\beta(t)$ functions (see Figure 4).

$$\alpha(t) = \begin{cases} \frac{5}{7}\,t + \frac{9}{7}, & 0 \le t \le 15 \\ -0.08\,t^2 + 3\,t - 17, & 15 < t \le 30 \end{cases} \tag{55}$$

$$\beta(t) = \begin{cases} t + 1, & 0 \le t \le 15 \\ -0.25\,t + 8, & 15 < t \le 30 \end{cases} \tag{56}$$

In this case, we will assume that the arrival rate to the system (machine 1) is given by α_1. Here:

$$WIP[P_i(t)] = \beta_i - \alpha_i,$$

$$WIB[B_{ij}(t)] = \min\{\beta_i, \xi_i\}.$$

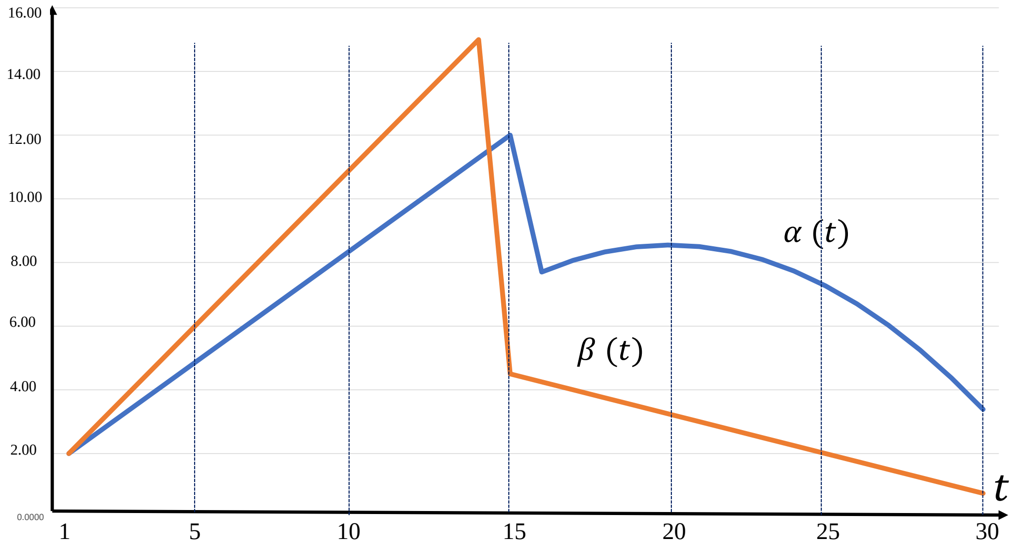

Figure 4. Functions $\alpha(t)$ and $\beta(t)$ used in the non-homogeneous birth–death model.

In this model, machine 1 is the access door to the system. Thus, the dynamics of the process is described as follows.

1. At $t = 0$, the raw material arrives at the machine M_1 at a rate given by the function $\alpha_1(t)$.
2. The number of products leaving machine one during the first hour is given by the function $\beta_1(t)$.
3. The output rate of products (s_1) from machine 1 is obtained from the quantity

$$s_1(t) = \min\{\xi_1, \beta_1(t)\}.$$

4. The WIP of the first machine is defined as the difference between what is supported and what is delivered as finished product.

$$WIP[M_1] = \alpha_1(t) - s_1(t)$$

5. Similarly, the $WIB[B_1]$ is given by

$$WIB[B_1] = \min\{s_1, h_1\}.$$

6. The WIB_1 and the quantity ξ_2 constitutes the input to M_2 in the following form

$$\alpha_2 = \min\{WIB[B_1], \xi_2\}.$$

Therefore, the process is repeated updating the respective indices.

From the definitions of $\alpha(t)$ and $\beta(t)$, the following results are obtained:

$$\omega(t) = \begin{cases} \frac{195}{7}, & 0 \le t \le 15 \\ -91.875, & 15 < t \le 30 \end{cases}$$

$$E(X_t) = \begin{cases} 0, & 0 \le t \le 15 \\ 0, & 15 < t \le 30 \end{cases}$$

$$\sigma^2_{X_t} = \begin{cases} \infty, & 1 \le t \le 15 \\ 0, & 15 < t \le 30 \end{cases}$$

$$P_0 = \begin{cases} 1, & 1 \le t \le 15 \\ 1, & 15 < t \le 30 \end{cases}$$

Table 5 summarizes some interesting results for this case.

Table 5. Average values for the non-homogeneous birth–death process for 100 h of operation.

$\boldsymbol{\Theta}$	M_1	B_{12}	M_2	B_{23}	M_3	B_{34}	M_4	B_{45}	M_5	B_{56}	M_6	$B_{6,fp}$
h		10		10		10		10		10		10
ξ	8		8		8		8		8		8	
$\bar{\alpha}(t)$	12.000		7.096		7.096		7.096		7.096		7.096	
$\bar{\beta}(t)$		16.000		16.000		16.000		16.000		16.000		16.000
s	7.096		7.096		7.096		7.096		7.096		7.096	
WIP	4.93		0		0		0		0		0	
WIB		12.000		7.096		7.096		7.096		7.096		7.096
$\mathcal{S}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$		$\mathcal{B}$

Again, the expected final quantity of production of this model is given by

$$M_{tot} = \sum_{i=1}^{6} WIB_i = 47.48, \tag{57}$$

and $s_i = \min\{\alpha(t), \beta(t)\}, i = 1, \ldots, n$ represents the amount of product coming from machine i at time t.

4.3. A Brief Discussion of the Results Obtained

In the takes all approach, the results obtained in the analyzed cases show some important differences. In the first one, it is assumed that the totality of the material contained in the buffer is taken to be processed. This implies that the subsequent work station is capable of channeling the material from which it is fed, placing it in tandem while the worksation absorbs it. This situation is typical of a numerical control center (NCC) with robotic systems for picking, classifying and placing the product on the equipment platform: for example, in car door assembly and welding, LE TV assembly, inspection and packaging lines, or vehicle engine cover assembly lines.

The numerical values obtained for the occupation, expected production and leisure totally depend on the λ and μ rates. In our example, these quantities were obtained experimentally in companies of the mechanical metal industry. Naturally, various values for these constants can be found reported in the literature. In the Conclusion, we cite several authors that report figures similar to ours. Although the estimators $\bar{y}$ and $\bar{z}$ are approximate values of the average production volume obtained during the planning horizon, they provide an indicator of the most likely values to be obtained under the conditions imposed on the system. The disadvantage of using them is the huge variance associated with them. In a model with multiple runs, it may be more appropriate to fit a theoretical density function to the samples created.

In a Markovian model, its simulation is simple due to the simplicity of the state matrix associated with the model. For example, taking advantage of the flow balance property from Equations (1) and (2) and knowing that the property $\pi_i p_{ij} = \pi_j p_{ji}$ for $i, j \in \Omega$ is fulfilled, then the Metropolis–Hasting algorithm is a useful tool for analyzing the behavior of the stochastic process $\zeta(t)$.

For the case of the pure birth–death and the non-homogeneous birth–death process, the ease of the model lies in their theoretical importance and the extensive literature on the subject. The classical values as the $\hat{Q}$ matrix uniquely determine the asymptotic behavior of the system, which allowed us to identify, under stable-state conditions of the system, the moments of leisure, bottlenecks and its stability.

Knowing the rates of birth (arrival of pieces to the buffer) and death (exit of pieces from the buffer), the system is practically characterized. Its simulation is also feasible for different initial (boundary) conditions in its parameters using the same Metropolis–Hasting algorithm. This could be a future line of research.

Particularly fascinating is the non-homogeneous birth–death model. Its beauty lies in the fact that arrival and departure rates depend on the instant of time in which they are observed. In practical models, this is an almost necessary condition that must be fulfilled to guarantee greater reality as in our proposal. Developing a simulation model of this process is almost a necessity. This would allow the analysis of more system variables such as operating costs (fixed and variable), stochastic demand and level of customer service.

In relation to the instances proposed in this approach, the results are very similar in terms of the average expected amount of production. In our case, the proposed models, in Equations (55) and (56) are estimated values of the production rate in flexible manufacturing systems associated with metalworking processes. Another important result found in this proposal is the Work in Process (WIP) and the introduction of a new variable called Work in Buffer (WIB). Our approach constitutes one more contribution to the literature since these metrics are mandatory in flexible manufacturing processes.

Finally, we want to express that the application of these models to real manufacturing cases is ready to be carried out. It is only necessary to obtain the magnitudes involved in the models through statistical sampling applied to the historical files that are prepared in the area of Maintenance Engineering in any manufacturing company. In Conclusions, several relevant papers are mentioned where models successfully have been applied to real engineering cases. The restrictions of our proposal is that it is limited only to those manufacturing models that satisfy the working hypotheses.

5. Conclusions

In this paper, we have carried out an analysis of the dynamics in the buffers of a mass production system whose workstations are aligned. By using probability theory, we were able to approximate three models that represent three different but common situations in the manufacturing industry. The results obtained show that under certain working hypotheses of the model, it is possible to make good approximations with a minimum of uncertainty.

We have analyzed three common situations. In the first one, we assumed that the total amount of product existing in each buffer is placed in the subsequent machine, and it is capable of operating simultaneously the volume of material received. In the second case, we assumed that there is a constant rate of access of products to the manufacturing line (a common situation in automated systems). With this in mind, we proposed a homogeneous birth–death model. The proposed rates are approximate magnitudes that are normally found in the metalworking industry, which is why they are very close to reality. The capacity of the buffers are also magnitudes that are directly related to the manufacturing capacities in the antecedent and consequent machines. Finally, the third case was built under the assumption that the corresponding arrival and departure rates of the products in the system buffers depend on the time in which the manufacturing line is operating. This assumption is not so unrealistic, since at certain times of the day, the rates vary due to factors such as worker fatigue, machine wear, shift changes, and set ups. We used a variant of the discrete-state stochastic process called the non-homogeneous birth–death process.

Although the assumptions under which this proposal is built vary substantially in each case, the results suggest that there is a great similarity in the expected amount of production in a planning horizon of 100 h. However, the idle levels and the number of bottlenecks are significantly different in each case. For this reason, we assume that the indicators of the models faithfully express their behavior. Measurements such as the unavailability of equipment, the expected variability of demand, and the probability distributions associated with the quantity of products stored in the buffers are engineering indicators of high value in the industry when the assumptions under which they were built are respected.

Although we may feel tempted to prepare a comparative table between the results obtained in each model, this would not be fully representative (except in the expected average production), since the construction assumptions change significantly from model to model. The proposed technique allows an exploration very close to the operation of a real system and, therefore, the results obtained are highly reliable.

In future lines of research in this topic, greater detail in aspects such as the time that the product remains in each work station should be included. In this case, a technique based on the age-dependent branching stochastic processes (also called regenerative processes) could be implemented as in [28,29].

Finally, another important aspect to address is the case of interconnected lines with derivations and stoppages.

In relation to the supposed reliability that should be had in a manufacturing line, reliability is defined as the probability that a system (in this case, a manufacturing line) operates satisfactorily at the time it is required. As a measure of probability, rarely does a system reach the value 1. That is, even the most reliable systems have an associated probability of failure. Therefore, buffer systems exist in both automated and non-automated manufacturing lines. For example, in automated lines, buffers automatically handle production stops and allow machines on the line to continue. The following papers refer to buffers in automated manufacturing lines: [30–36].

As for mean time to repair (MTTR), this is the ratio of the total time spent on unplanned maintenance to the number of times a piece of equipment has failed during a specified period. MTTR does not consider parts lead time and is not used for scheduled maintenance. The concept assumes a metric with a single meaning. However, the letter R can stand for repair, recovery, response or resolution, and each case is specific to the company where it is used [37]. The variability of times lies in the fact that each company is a particular case; each facility and its topology is different. MTTR provides information to the extent of the data on which it is based. Therefore, a careful service history must be kept of the machines involved.

An estimate of the MTTR must be made considering the correct workflow and timestamps that collect data from multiple sources. For example, a workflow can be made up of the following activities: (a) Equipment operator enters a work order, (b) A scheduler releases the work order, (c) A technician adds his time to the work order, (d) The technician executes the requested work, and (e) Work order completed. An ordinary sequence of times to determine this magnitude is as follows [38]:

$$\text{Incident} \to \text{Failure detected} \to \text{Diagnosis} \to \text{Repair} \to \text{Recovery} \to \text{Availability} \to \text{Incident}$$

In any case, if the MTTR is too large, then the replacement of the damaged part is considered. In any case, this consumes a time that is never standard for all the companies that make use of this concept. Some authors who report their experience in the use of high values for an MTTR are the following: [39–42]. In particular, Ref. [43] i ncludes the idea of using buffers and MTTR.

In a manufacturing line, we could have two failure causes. In the first one is the operator or another instance that decides to stop the system. In the second, the failure is considered due to some technical aspect. In this work, we were only interested in failures of the technical type.

To date, there is a significant increase in the complexity of production technologies and the equipment used in them. In this case, mathematical models help to understand the complexity of the systems and their interactions with other components in a manufacturing line. A typical example of this is trying to understand how bottlenecks arise in the production process. Traditionally, in some industrial companies, mathematical models are used to analyze the system and make operational or resource decisions. The advantage of a mathematical model of a new system can be reflected in cost and flexibility to adapt the decisions made and due to its relative accuracy, which influences how and where they can be used [44]. Some significant examples showing the advantages and delimitations of mathematical models applied to industrial situations can be verified in [45,46].

Author Contributions: G.P.-L. developed the mathematical model of the proposal. Also obtained the information for the creation of the instances and elaborated their computational runs. F.V.-M. supervised and adapted the mathematical model. Also reviewed the results of the computational runs and adapted them to the proposed models. M.A.M.-B. obtained the information used in the model and elaborated the experimental designs that guarantee the reliability of the results. J.M.-V. supervised the development, construction, style, spelling and computational runs of the document as well as its technical content. All authors have read and agreed to the published version of the manuscript.

Funding: The funds for the publication of this document were provided by the Instituto Tecnológico y de Estudios Superiores de Monterrey, Mexico.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chapter 3 Mathematical Modeling of Production Systems. Available online: http://www.productionsystemsengineering.com/pdfs/PSE_Chapter3.pdf (accessed on 10 February 2022).
2. Ilija, D.; Alen, M.; Aleksandra, N.; Anet, R.J.; Photis, P.; Aberham, H.F.; Kamal, K.; Rallou, T.; Heiko, B.; Alexandrina, S.; et al. Sirbu Alexandrina, Mihnea Moisescu Alexandru, Tomasevic Igor, Vrabič Brodnjak Urška, Charalambides Maria, Tonda Alberto. Cross-European initial survey on the use of mathematical models in food industry. *J. Food Eng.* **2019**, *261*, 109–116.
3. Shadkam, E.; Ghavidel, F. Presentation and Implementation Multi-Objective Mathematical Models to Balance the Assembly Line. *Int. J. Res. Ind. Eng.* **2021**, *10*, 56–66.
4. Usubamatov, R.; Sartov, T. A mathematical model for productivity rate of APSL segmented on sections with buffers of limited capacity. *Int. J. Adv. Manuf. Technol.* **2017**, *92*, 3859–3869. [CrossRef]
5. Grznár, P.; Gregor, M.; Mozol, Š.; Krajčovič, M.; Dulina, Ľ.; Gašo, M.; Major, M. A System to Determine the Optimal Work-in-Progress Inventory Stored in Interoperation Manufacturing Buffers. *Sustainability* **2019**, *11*, 3949. [CrossRef]
6. Chen, J.; Jia, Z.; Dai, Y. Real-Time Performance Analysis of Batch-Based Serial Flexible Production Lines with Geometric Machines. In Proceedings of the 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), Vancouver, BC, Canada, 22–26 August 2019.
7. Fragapane, G.; Ivanov, D.; Peron, M.; Sgarbossa, F.; Strandhagen, J.O. Increasing flexibility and productivity in Industry 4.0 production networks with autonomous mobile robots and smart intralogistics. *Ann. Oper. Res.* **2022**, *308*, 125–143. [CrossRef]
8. Rossit, D.A.; Tohmé, F.; Frutos, M. An Industry 4.0 approach to assembly line resequencing. *Int. J. Adv. Manuf. Technol.* **2019**, *105*, 3619–3630. [CrossRef]
9. Shi, L.; Men, S. Optimal buffer allocation in production lines. *IIE Trans.* **2003**, *35*, 1–10. [CrossRef]
10. Ouzineb, M.; Mhada, F.Z.; Pellerin, R.; Hallaoui, I. Optimal planning of buffer sizes and inspection station positions. *Prod. Manuf. Res.* **2018**, *6*, 90–112. [CrossRef]
11. Imseitif, J.; Tang, H.; Smith, M. Throughput Analysis of Manufacturing Systems with Buffers Considering Reliability and Cycle Time Using DES and DOE. In Proceedings of the 25th International Conference on Production Research Manufacturing Innovation: Cyber Physical Manufacturing, Chicago, IL, USA, 9–14 August 2019.
12. Ameen, W.; AlKahtani, M.; Mohammed, M.K.; Abdulhameed, O.; El-Tamimi, A.M. Investigation of the effect of buffer storage capacity and repair rate on production line efficiency. *J. King Saud Univ.-Eng. Sci.* **2018**, *30*, 243–249. [CrossRef]
13. Chomnawung, Y.; Prombanpong, S.; Klavohm, C. A Buffer Analysis in a Transfer Production Line. *MATEC Web Conf.* **2016**, *68*, 06005. [CrossRef]
14. Tolga, T.; Abhijit, G. Optimal Buffer Allocation in Production Lines Using an Automata Search. Available online: https://web.mst.edu/~gosavia/tezcan_gosavi.pdf (accessed on 7 February 2022).
15. Schuler, F.; Darabi, H. Buffer clustering policy for sequential production lines with deterministic processing times. *Int. J. Ind. Eng. Comput.* **2016**, *7*, 555–572. [CrossRef]
16. Duan, J.; Xie, N.; Li, L. Optimal Buffer Allocation in Multi-Product Repairable Production Lines Based on Multi-State Reliability and Structural Complexity. *KSII Trans. Internet Inf. Syst. Korean Soc. Internet Inf.* **2020**, *14*, 1579–1602.
17. Smith, J.M.; Daskalaki, S. Buffer Space Allocation in Automated Assembly Lines. *Oper. Res.* **1988**, *36*, 343–358. [CrossRef]
18. Nourelfath, M.; Nahas, N.; Ait-Kadi, D. Optimal design of series production lines with unreliable machines and finite buffers. *J. Qual. Maint. Eng.* **2005**, *11*, 121–138. [CrossRef]
19. D'Apice C.; Göttlich, S.; Herty, M.; Piccoli, B. *Modeling, Simulation, and Optimization of Supply Chains: A Continuous Approach*; Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, USA, 2010; pp. 5–9, ISBN 978-0-898717-00-6.
20. Hongler, M.-O. Stochastic Buffered Flows. In *Chaotic and Stochastic Behaviour in Automatic Production Lines*; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 1994; pp. 57–76.

21. Pérez-Lechuga, G.; Gress, E.S.H.; Karelyn, A.A.; Orán, M.G.M. Some Efficiency Measures in the Operation of Flexible Manufacturing Systems: A Stochastic Approach, Society for Industrial and Applied Mathematics. In Proceedings of the 2009 SIAM Conference on Mathematics for Industry, San Francisco, CA, USA, 9–10 October 2009; Society for the Industrial and Applied Mathematics (SIAM): Philadelphia, PA, USA, 2009; pp. 31–36.
22. Pérez-Lechuga, G.; Venegas-Martínez, F.; Martínez-Sánchez, J.F. Mathematical Modeling of Manufacturing Lines with Distribution by Process: A Markov Chain Approach. *Mathematics* **2021**, *9*, 3269. [CrossRef]
23. Karlin, S. Linear Growth, Birth-and-Dead Processes. *J. Math. Mech.* **1958**, *1*, 643–662.
24. Kendall, D.G. On the Generalized "Birth-and-Death" Process. *Ann. Math. Stat.* **1948**, *19*, 1–15. doi: 10.1214/aoms/1177730285. [CrossRef]
25. Bharucha-Reid, A.T. *Elements of the Theory of Markov Processes and Their Applications*; Dover Publications Inc.: New York, NY, USA, 2010; pp. 89–93, ISBN 9780486695396.
26. Lamens, A.; Consael, R. Sur le processus non-homogène de naissance et de mort. *Acad. R. Belg. Bull. Cl. Sci. Ser.* **1957**, *43*, 597–605.
27. Vvedenskaya, N.D.; Logachov, A.V.; Suhov, Y.M.; Yambartsev, A.A. A local large deviation principle for inhomogeneous birth-death processes. *Probl. Inf. Transm.* **2018**, *54*, 263–280. [CrossRef]
28. Bellman, R. A Survey of the Mathematical Theory of Time-Lag, Retarded Control, and Hereditary Processes. RAND Monograph. EE.UU. 1954. Available online: https://www.rand.org/content/dam/rand/pubs/reports/2009/R256.pdf (accessed on 24 April 2022).
29. Stidham, S., Jr. Regenerative processes in the theory of queues, with applications to the alternating-priority queue. *Adv. Appl. Probab.* **1972**, *4*, 542–577. Available online: https://www.jstor.org/stable/1425993 (accessed on 5 May 2022). [CrossRef]
30. Savsar, M. Modeling of a two-stage merge production with buffer storage. *JKAU Eng. Sci.* **1992**, *4*, 67–81. [CrossRef]
31. McClain, J.O.; Moodie, D.R. A Comment on "Buffer Space Allocation in Automated Assembly Lines". *Oper. Res.* **1991**, *39*, 857–860. [CrossRef]
32. Pravin, K.J. A linear programming approach to capacity estimation of automated production lines with finite buffers. *Int. J. Prod. Res.* **1987**, *25*, 851–867.
33. Zhou, M.; DiCesare, F. Petri net modelling of buffers in automated manufacturing systems. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **1987**, *26*, 157–164. [CrossRef] [PubMed]
34. Prombanpong, S.; Kaewyu, J.; Thanadulthaveedech, N.; Matwangsang, M. A buffer design for mitigation downtime effect in an automated transfer line. *Int. J. Innov. Manag. Technol.* **2013**, *4*, 155.
35. Xia, B.; Chen, J.; Zhang, Z. An Exact Method for the Analysis of a Two-Machine Manufacturing System with a Finite Buffer Subject to Time-Dependent Failure. *Math. Probl. Eng.* **2015**, *2015*, 916193. [CrossRef]
36. Faria, J.; Matos, M.A.; Pinto, N.E.M. Optimal design of work-in-process buffers. *Int. J. Prod. Econ.* **2006**, *99*, 144–155. [CrossRef]
37. The Path to Better Incident Management Starts Here. Available online: https://www.atlassian.com/incident-management/kpis/common-metrics (accessed on 24 April 2022).
38. Reliability Metrics 101: Mean Time to Repair (MTTR). Available online: https://www.maxgrip.com/resource/reliability-metrics-101-mean-time-to-repair-mttr/ (accessed on 24 April 2022).
39. Atighi, I.; Soleimannejad, J.; Akbarinasab, A.; Moradpour, S. Providing a Practical Model to Reduce Maintenance Costs: A Case Study in Golgohar Company. *World Acad. Sci. Eng. Technol. Int. J. Econ. Manag.* **2019**, *13*, 1457–1460. Available online: https://zenodo.org/record/3593154#.YmWPO_PMI6U (accessed on 7 February 2022).
40. Naraphorn Paoprasert, N.; Htet Lin, W.Y.; Muneekaew, T. Assessing Risk Priority Numbers of Failures in the Screw Tightening Machine of a Hard Disk Drive Production System. *J. Mach. Eng.* **2022**, *22*, 124–137. Available online: http://jmacheng.not.pl/Application-of-Reliability-Centered-Maintenance-in-Screw-Tightening-Machine-of-Hard,145272,0,2.html (accessed on 7 February 2022). [CrossRef]
41. Mohammad, Z.; Aggarwal, G. OEE Evaluation of Long Life Food Product Line. *Int. J. Emerg. Technol. Eng. Res. (IJETER)* **2016**, *4*, 61–66. Available online: https://ijeter.everscience.org/Manuscripts/Volume-4/Issue-8/Vol-4-issue-8-M-13.pdf (accessed on 5 April 2022).
42. Nahas, N.; Nourelfath, M. Joint optimization of maintenance, buffers and machines in manufacturing lines. *Eng. Optim.* **2018**, *50*, 37–54. [CrossRef]
43. Stanley, B.G. The future of manufacturing systems engineering. *Int. J. Prod. Res.* **2018**, *56*, 224–237. [CrossRef]
44. Symonov, D. Usage of mathematical modeling on the industrial enterprises. In Proceedings of the Conference: DO Desenvolvimento Mundial Como Resultado de Realizações em Ciência e Investigação Científicaat, Lisboa, Portugal, 9 October 2020. [CrossRef]
45. Pérez Lechuga, G. Optimal logistics strategy to distribute medicines in clinics and hospitals. *J. Math. Ind.* **2018**, *8*, 2. [CrossRef]
46. Pérez-Lechuga, G.; Aguilar-Velázquez, S.L.; Cisneros-López, M.A. A model for the location and scheduling of the operation of second-generation ethanol biorefineries. *J. Math. Ind.* **2019**, *9*, 3. [CrossRef]