Chapter 5

Continuous Improvement of Bernoulli Lines

Motivation: It is not uncommon that, due to unscheduled downtime, machining lines in many industries operate at 60%-70% of their capacity. Although assembly systems are typically more efficient (often operating at 80%-90% of their capacity), the losses are still significant. In this situation, continuous improvement is a major tool for production systems management.

Typically, continuous improvement projects are developed using common sense, managerial intuition and, in some cases, discrete event simulations. Due to the "soft" nature of these approaches, they often do not result in an actual productivity improvement. The purpose of this chapter is to present analytical methods for designing continuous improvement projects in Bernoulli lines with predictable results. The development is based on the analytical method and recursive equations derived in Chapter 4.

Overview: Two approaches to design of continuous improvement projects are developed. They are referred to as constrained and unconstrained improvability.

Constrained improvability addresses the issue of improving a production system by re-allocating its limited resources, e.g., buffer capacity or workforce. The main question here is: Can or cannot a production system be improved by utilizing more efficiently its limited resources? If it is possible, the system is called improvable under constraints; otherwise, it is unimprovable. Section 5.1 presents criteria, which allow to determine whether the system is improvable and provides a characterization of unimprovable allocations.

Constrained improvability is related to optimality. Indeed, an unimprovable system is, in fact, optimal. We use, however, the term "improvable" to indicate that the goal is not necessarily to render the system optimal but rather to determine whether it can be improved and indicate actions that lead to this improvement. Moreover, given the lack of accurate information on the factory floor, the optimality may not be practically achievable, whereas continuous improvement, being robust with respect to inaccurate information, may.

Unconstrained improvability addresses the issue of bottleneck identification and elimination by allocating additional resources (such as additional buffer capacity, machine improvement or replacement, etc.).

The concept of bottleneck (BN) is not well understood, and, as a result, it is not unusual that in practice an improvement or replacement of a machine, viewed as the BN, leads to no improvement of the production system as a whole. So, what is a BN? Often, BN is understood as the machine with the smallest production rate in isolation. In other cases, the machine with the largest work-in-process in front of it is viewed as the BN. It is possible to show, however, that neither may be the BN in the sense of being the most impeding for the production rate of the system. This happens because the above intuitive conceptualizations are local in nature and do not take into account the total system properties, such as the location of machines in the production line, capacity of the buffers, types of interactions among the machines and buffers, etc. In Section 5.2, we introduce "system-based" definitions of bottleneck machines (BN-m) and bottleneck buffers (BN-b) in terms of their effect on the production rate of the line.

The main practical results of this chapter are the criteria (referred to as *indicators of improvability*), which allow factory floor personnel to determine if the system is improvable (in the constrained or unconstrained case) and define actions that must be taken to achieve this improvement. In addition, we define the notion of *buffering potency* and introduce the method of *measurement-based management* of production systems.

5.1 Constrained Improvability

5.1.1 Resource constraints and definitions

Consider a serial production line with M Bernoulli machines defined by parameters p_i, $i = 1, \dots, M$, and $M-1$ buffers with capacities N_i, $i = 1, \dots, M-1$, which operates according to conventions (a)-(e) of Subsection 4.2.1.

Assume that N_i's and p_i's are constrained as follows:

$$\sum_{i=1}^{M-1} N_i = N^*, \tag{5.1}$$

$$\prod_{i=1}^{N} p_i = p^*, \tag{5.2}$$

where N^* and p^* are positive numbers with p^* satisfying $p^* < 1$. Constraint (5.1) implies that the total buffer capacity cannot exceed N^*. Constraint (5.2) can be interpreted as a bound on the machine efficiency or workforce. Indeed, in many systems, assignment of the workforce (both machine operators and skilled trades for repair and maintenance) defines the machine efficiency and, thus, p_i's.

Therefore, we refer to (5.1) and (5.2) as the *buffer capacity* (BC) and *workforce* (WF) *constraints*, respectively.

Let, as before, $\widehat{PR} = \widehat{PR}(p_1, \ldots, p_M, N_1, \ldots, N_{M-1})$ denote the production rate of the system, calculated using (4.30)-(4.36).

Definition 5.1 *A serial production line with Bernoulli machines is:*

- *improvable with respect to BC if there exists a sequence* $N_1', \ldots, N_{M-1}'$ *such that*
$$\sum_{i=1}^{M-1} N_i' = N^* \tag{5.3}$$
and
$$\widehat{PR}(p_1, \ldots, p_M, N_1', \ldots, N_{M-1}') > \widehat{PR}(p_1, \ldots, p_M, N_1, \ldots, N_{M-1}); \tag{5.4}$$
otherwise, it is unimprovable with respect to BC;

- *improvable with respect to WF if there exists a sequence* $p_1', \ldots, p_M'$ *such that*
$$\prod_{i=1}^{M} p_i' = p* \tag{5.5}$$
and
$$\widehat{PR}(p_1', \ldots, p_M', N_1, \ldots, N_{M-1}) > \widehat{PR}(p_1, \ldots, p_M, N_1, \ldots, N_{M-1}); \tag{5.6}$$
otherwise, it is unimprovable with respect to WF;

- *improvable with respect to BC and WF simultaneously if there exist sequences* $N_1', \ldots, N_{M-1}'$ *and* $p_1', \ldots, p_M'$ *such that*
$$\sum_{i=1}^{M-1} N_i' = N^*, \qquad \prod_{i=1}^{M} p_i' = p* \tag{5.7}$$
and
$$\widehat{PR}(p_1', \ldots, p_M', N_1', \ldots, N_{M-1}') > \widehat{PR}(p_1, \ldots, p_M, N_1, \ldots, N_{M-1}); \tag{5.8}$$
otherwise, it is unimprovable with respect to BC and WF simultaneously.

Conditions for various types of improbability are given next.

5.1.2 Improvability with respect to WF

Necessary and sufficient conditions: Below we provide both theoretical and practical conditions of improvability with respect to WF.

Theorem 5.1 *A Bernoulli line defined by assumptions (a)-(e) of Subsection 4.2.1 is unimprovable with respect to WF if and only if*

$$p_i^f = p_{i+1}^b, \qquad i = 1, \ldots, M-1, \tag{5.9}$$

where p_i^f *and* p_i^b *are the steady states (4.35) of the recursive aggregation procedure (4.30).*

Proof: See Section 20.1.

Recall that, as it has been shown in Chapter 4, for each buffer b_i, an M-machine line can be represented as a two-machine system with virtual machines characterized by p_i^f and p_{i+1}^b (see Figure 4.10). Thus, Theorem 5.1 implies that the necessary and sufficient condition of unimprovability with respect to WF is that both virtual machines are identical. Using the property (4.21) of two-machine lines with identical machines, we obtain

Corollary 5.1 *Under condition (5.9),*

$$\widehat{WIP}_i = \frac{N_i(N_i+1)}{2(N_i+1-p_i^f)}, \qquad i = 1, \ldots, M-1. \tag{5.10}$$

Proof: See Section 20.1.

Since $0 < p_i^f < 1$, expression (5.10) implies that

$$\frac{N_i}{2} < \widehat{WIP}_i < \frac{N_i+1}{2}, \qquad i = 1, \ldots, M-1. \tag{5.11}$$

Based on this, we formulate the practical

WF-Improvability Indicator 5.1: *A Bernoulli line is practically unimprovable with respect to WF if each buffer is, on average, close to being half full.*

Although this indicator may seem somewhat unexpected, it is, in retrospect, quite natural. Indeed, buffer b_i is intended to protect m_i from blockage and m_{i+1} from starvation. From the point of view of m_i, buffer b_i should be all the time empty; from the point of view of m_{i+1} it should be full. The compromise is – the buffer is half full (with some correction depending on the machines' efficiency as indicated in (5.10) by p_i^f).

Unimprovable allocation of p_i: To characterize the unimprovable (i.e., optimal) p_i's, introduce the notation

$$\widehat{PR}^* := \max_{p_1,\ldots,p_M;\prod_{i=1}^M p_i = p^*} \widehat{PR}(p_1, \ldots, p_M, N_1, \ldots, N_{M-1}). \tag{5.12}$$

Clearly, $\widehat{PR}^*$ is the largest production rate that can be attained in the system under constraint (5.2). It turns out that $\widehat{PR}^*$ can be calculated as follows:

Consider the recursive procedure

$$x(n+1) = (p^*)^{\frac{1}{M}} \prod_{i=1}^{M-1} \Big(\frac{N_i + x(n)}{N_i + 1}\Big)^{\frac{2}{M}}, \qquad x(0) \in (0,1). \tag{5.13}$$

Theorem 5.2 *Assume $\sum_{i=1}^{M-1} N_i^{-1} \le M/2$. Then recursive procedure (5.13) is convergent and its steady state, x^*, is:*

$$x^* = \lim_{n\to\infty} x(n) = \widehat{PR}^*. \tag{5.14}$$

Proof: See Section 20.1.

Thus, the maximum production rate can be calculated without the knowledge of the optimal WF allocation. Moreover, this $\widehat{PR}^*$ can be used to calculate the optimal allocation of p_i's, as shown below.

Theorem 5.3 *The sequence p_1^*, …, p_M^*, $\prod_{i=1}^{M} p_i^* = p^*$, which renders the serial production line with Bernoulli machines unimprovable with respect to WF, is given by*

$$p_1^* = \Big(\frac{N_1 + 1}{N_1 + \widehat{PR}^*}\Big)\widehat{PR}^*, \tag{5.15}$$

$$p_i^* = \Big(\frac{N_{i-1} + 1}{N_{i-1} + \widehat{PR}^*}\Big)\Big(\frac{N_i + 1}{N_i + \widehat{PR}^*}\Big)\widehat{PR}^*, i = 2, \ldots, M-1, \tag{5.16}$$

$$p_M^* = \Big(\frac{N_{M-1} + 1}{N_{M-1} + \widehat{PR}^*}\Big)\widehat{PR}^*. \tag{5.17}$$

Proof: See Section 20.1.

Thus, for any given sequence N_1, …, N_{M-1} and p^* of (5.2), the optimal allocation of work can be calculated using (5.13)-(5.17).

Taking into account that $\widehat{PR}^* < 1$, expressions (5.15)-(5.17) lead to a "flat" inverted bowl phenomenon:

Corollary 5.2 *If all buffers are of equal capacity, i.e., $N_i =: N$, $i = 1$, …, $M-1$, then*

$$p_1^* = p_M^* < p_2^* = p_3^* = \ldots = p_{M-1}^*. \tag{5.18}$$

Although the flat inverted bowl allocation of optimal p_i's does take place, it does not lead, as it was mentioned before, to a significantly larger PR as compared with the uniform allocation. Indeed, consider a system with five Bernoulli machines, $N_i = 2$, $i = 1, \ldots, 4$, and $p^* = (0.9)^5 = 0.5905$. For this system, the optimal p_i allocation, according to (5.15)-(5.17) is

$$p_1^* = p_5^* = 0.8655,$$
$$p_2^* = p_3^* = p_4^* = 0.9237,$$

leading to $\widehat{PR}^* = 0.8110$. If p_i's were distributed in the uniform manner, i.e.,

$$p_i = 0.9, \qquad i = 1, \ldots, 5,$$

$\widehat{PR}_{unif} = 0.8050$, which is only 0.74% less than the production rate for the inverted bowl allocation. To reinforce this conclusion, consider one more example: $M = 10$, $p^* = (0.95)^{10} = 0.5987$, and $N_i = 3$, $i = 1, \ldots, 9$. Then the optimal allocation (5.15)-(5.17) becomes

$$\begin{aligned} p_1^* &= p_{10}^* = 0.9335, \\ p_i^* &= 0.9542, \qquad i = 2, \ldots, 9 \end{aligned}$$

and the optimal $\widehat{PR}^* = 0.9132$, while the uniform allocation

$$p_i = 0.95, \qquad i = 1, \ldots, 10$$

results in $\widehat{PR}_{unif} = 0.9104$, which is only 0.31% less than $\widehat{PR}^*$.

Continuous improvement by WF re-allocation: While expressions (5.15)-(5.17) can be used for design of new production systems, continuous improvement projects for existing operations can be designed using WF-Improvability Indicator 5.1. This can be carried out as follows:

WF-Continuous Improvement Procedure 5.1:

(1) By calculations off-line (using the aggregation procedure (4.30) and formula (4.37)) or by measurements on the factory floor, evaluate the average buffer occupancy, WIP_i, $i = 1, \ldots, M-1$.

(2) Determine the buffer for which $|WIP_i - (N_i + 1)/2|$ is the largest. Assume this is buffer k.

(3) If $WIP_k - (N_k + 1)/2$ is positive, re-allocate a sufficiently small amount of work, ϵp_k, from m_k to m_{k+1}; if $WIP_k - (N_k + 1)/2$ is negative, re-allocate ϵp_{k+1} from m_{k+1} to m_k (observing, of course, the constraint $p_k p_{k+1} = const$).

(4) Return to step (1).

(5) Continue this process until $\max_i |WIP_i - (N_i + 1)/2|$ is sufficiently small.

To illustrate this procedure, consider a four-machine Bernoulli line with $N_i = 5$ and $p = [0.9675, 0.9225, 0.8780, 0.8372]$. The $\widehat{PR}$ of this line is 0.8281. Running WF-Continuous Improvement Procedure 5.1 with $\epsilon = 0.01$, we arrive at the following WF allocation $p = [0.8875, 0.9125, 0.9163, 0.8841]$. The resulting $\widehat{PR} = 0.8707$, which is a 5.14% improvement.

5.1.3 Improvability with respect to WF and BC simultaneously

Necessary and sufficient conditions: For the case of simultaneous improvability with respect to WF and BC, these conditions can be formulated as follows:

Theorem 5.4 *A Bernoulli line defined by assumptions (a)-(e) of Subsection 4.2.1 is unimprovable with respect to WF and BC simultaneously if and only if*

$$p_1 = p_i^f = p_i^b = p_M, \qquad i = 2, \dots, M-1. \tag{5.19}$$

Proof: See Section 20.1.

Corollary 5.3 *Under condition (5.19),*

$$\begin{aligned} \widehat{BL}_1 &= \widehat{ST}_M, & (5.20)\\ \widehat{BL}_i &= \widehat{ST}_i, \quad i = 2, \dots, M-1. & (5.21)\end{aligned}$$

In addition,

$$N_i = N, \quad i = 1, \dots, M-1, \tag{5.22}$$

and

$$\widehat{WIP}_i = \frac{N(N+1)}{2(N+1-p_1)}, \qquad i = 1, \dots, M-1. \tag{5.23}$$

Proof: See Section 20.1.

Thus, if a line is unimprovable with respect to WF and BC simultaneously, all internal machines are starved and blocked with equal frequency, the first machine is blocked as frequently as the last machine is starved, all buffers are of equal capacity, have equal steady state occupancy, and this occupancy can be evaluated using (5.23). Clearly,

$$\frac{N}{2} < \widehat{WIP}_i < \frac{N+1}{2}, \qquad i = 1, \dots, M-1.$$

Unimprovable allocation of p_i and N_i: To characterize the unimprovable (i.e., optimal) p_i and N_i allocations, introduce

$$\widehat{PR}^{**} := \max_{\substack{N_1,\dots,N_{M-1};\, \sum_{i=1}^{M-1} N_i = N^* \\ p_1,\dots,p_M;\, \prod_{i=1}^{M} p_i = p^*}} \widehat{PR}(p_1, \dots, p_M, N_1, \dots, N_{M-1}). \tag{5.24}$$

Theorem 5.5 *Let N^* be a multiple of $M-1$. Then the sequences p_1^*, …, p_M^* and N_1^*, …, N_{M-1}^*, which render the serial production line with Bernoulli machines unimprovable with respect to WF and BC simultaneously, are given by*

$$\begin{aligned} N_i^* &= \frac{N^*}{M-1} = N_{opt}^*, \qquad i = 1, \dots, M-1, & (5.25)\\ p_1^* &= p_M^* = \Big(\frac{N_{opt}^*+1}{N_{opt}^*+\widehat{PR}^{**}}\Big)\widehat{PR}^{**}, & (5.26)\\ p_i^* &= \Big(\frac{N_{opt}^*+1}{N_{opt}^*+\widehat{PR}^{**}}\Big)^2\widehat{PR}^{**}, \qquad i = 2, \dots, M-1. & (5.27)\end{aligned}$$

Proof: See Section 20.1.

Thus, based on the capacity N^*_{opt}, the value of $\widehat{PR}^{**}$ can be calculated using (5.13), and then p_i^*'s can be evaluated from (5.26) and (5.27). This result can be useful at the design stage of serial production lines. Note that the optimal p_i allocation is again a flat inverted bowl, while optimal N_i's are uniform.

5.1.4 Improvability with respect to BC

Necessary and sufficient conditions:

Theorem 5.6 *A Bernoulli line defined by assumptions (a)-(e) of Subsection 4.2.1 is unimprovable with respect to BC if and only if the quantity*

$$\min_{i=1,\ldots,M} p_i \Big(\min \Big\{ \frac{p_i^f}{p_i^b}, \frac{p_i^b}{p_i^f} \Big\} \Big) \tag{5.28}$$

is maximized over all sequences $N_1', \ldots, N_{M-1}'$ such that $\sum_{i=1}^{M-1} N_i' = N^*$.

Proof: See Section 20.1.

Unfortunately, this result is of little practical importance, since its interpretation and utilization are much less obvious than those of (5.9) and (5.19). However, numerical experiments with (5.28) show that in most cases the unimprovable sequence $N_1^*, \ldots, N_{M-1}^*$, $\sum_{i=1}^{M-1} N_i^* = N^*$, can be determined by a criterion with a clearer interpretation:

Numerical Fact 5.1 *The production rate ensured by the buffer capacity allocation defined by Theorem 5.6 is almost always the same as the production rate defined by the allocation that minimizes*

$$\max_{i=2,\ldots,M-1} |\widehat{WIP}_{i-1} - (N_i - \widehat{WIP}_i)| \tag{5.29}$$

over all sequences $N_1', \ldots, N_{M-1}'$ such that $\sum_{i=1}^{M-1} N_i' = N^*$.

This implies that a production line is practically unimprovable with respect to BC if the occupancy of each buffer b_{i-1} is as close to the availability of buffer b_i as possible. In other words, b_{i-1} and b_i offer practically equal protection of machine m_i against starvations and blockages, respectively. This is illustrated in Figure 5.1 where the occupancy of b_{i-1} (i.e., $\widehat{WIP}_{i-1}$) and the availability of b_i (i.e., $N_i - \widehat{WIP}_i$) are indicated by shaded areas. Based on the above, we formulate the practical

BC-Improvability Indicator 5.1: *A Bernoulli line is practically unimprovable with respect to BC if the average occupancy of each buffer is as close as possible to the average availability of its downstream buffer.*

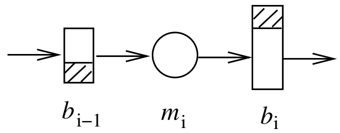


Figure 5.1: Illustration of Numerical Fact 5.1

Continuous improvement by BC re-allocation: Similar to WF, the above improvability indicator can be used in the following

BC-Continuous Improvement Procedure 5.1:

(1) By calculations off-line (using (4.30) and (4.37)) or by measurements on the factory floor, evaluate the average occupancy of each buffer, WIP_i, $i = 1$, ..., $M-1$.
(2) Determine the buffer for which $|WIP_i-(N_{i+1}-WIP_{i+1})|$, $i = 1, \ldots, M-2$, is the largest. Assume this is buffer k.
(3) If $WIP_k - (N_{k+1} - WIP_{k+1})$ is positive, transfer a unit of capacity from b_k to b_{k+1}; if $WIP_k - (N_{k+1} - WIP_{k+1})$ is negative, re-allocate a unit of capacity from b_{k+1} to b_k.
(4) Return to step (1).
(5) Continue this process until arriving at a limit cycle and choose the buffer capacity allocation on the limit cycle, which maximizes PR.

As an illustration, consider an 11-machine serial line with $p_i = 0.8$, $i \neq 6$ and $p_6 = 0.6$. Assume that the total buffer capacity is 24 and determine the unimprovable buffer capacity allocation. Based on the above procedure and calculating $\widehat{WIP}_i$ using (4.30), (4.37), the following result is obtained:

$$\begin{aligned} N_1 &= 1, \\ N_2 &= N_3 = N_4 = 2, \\ N_5 &= 5, \\ N_6 &= 4, \\ N_7 &= N_8 = N_9 = N_{10} = 2. \end{aligned}$$

This allocation ensures $\widehat{PR} = 0.5843$, which is almost the maximum possible in this system.

Note that if, following Goldratt's *Theory of Constraints*, all available buffer capacity were placed in front of the "bottleneck" machine m_6, i.e.,

$$\begin{aligned} N_i &= 1, \quad i \neq 5, \\ N_5 &= 15, \end{aligned}$$

the resulting $\widehat{PR} = 0.427$, which is 27% lower than that of the unimprovable allocation. If, keeping in mind the property of reversibility (see Section 4.3), the "bottleneck" machine is protected on both sides, i.e., the allocation is

$$\begin{aligned} N_i &= 1, \quad i \neq 5, 6, \\ N_5 &= N_6 = 8, \end{aligned}$$

the resulting $\widehat{PR} = 0.491$, which is still 16% lower than that of the unimprovable allocation obtained using BC-Continuous Improvement Procedure 5.1.

As a final note, it should be pointed out that in some cases, the unimprovable allocation may not appear in the limit cycle, but a few steps before the limit cycle is reached. However, the production rate obtained using BC-Continuous Improvement Procedure 5.1 is typically very close to the production rate ensured by the unimprovable allocation.

PSE Toolbox: The unimprovable allocations of WF and BC, as well as WF- and BC-Continuous Improvement Procedures, are implemented in the **Continuous Improvement** function of the toolbox. For a description and illustration of these tools, see Subsections 19.4.1–19.4.4.

5.2 Unconstrained Improvability

5.2.1 Definitions

Bottleneck machine: Consider a serial production line with M Bernoulli machines defined by parameters p_i, $i = 1, \dots, M$, and $M - 1$ buffers with capacities N_i, $i = 1, \dots, M - 1$. Assume that the line operates according to conventions (a)-(e) of Subsection 4.2.1.

Let, as before, PR, denote the production rate of the system, i.e.,

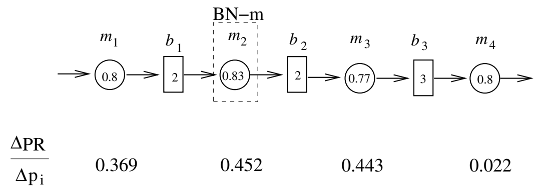
$$PR = PR(p_1, \dots, p_M, N_1, \dots, N_{M-1}).$$

Definition 5.2 *Machine m_i, $i \in \{1, \dots, M\}$, is the bottleneck machine (BN-m) of a Bernoulli line if*

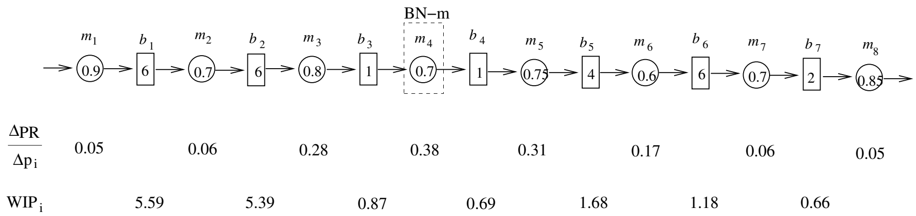
$$\frac{\partial PR}{\partial p_i} > \frac{\partial PR}{\partial p_j}, \quad \forall j \neq i. \tag{5.30}$$

Due to the monotonicity properties of PR with respect to p_i's (see Section 4.3), both derivatives in (5.30) are positive. Thus, Definition 5.2 implies that m_i is the BN-m if its infinitesimal improvement leads to the largest increase of the production rate, as compared with a similar improvement of any other machine in the system.

A machine with the smallest p_i is not necessarily the BN-m in the sense of Definition 5.2. Indeed, consider the production lines shown in Figure 5.2, where the numbers in the circles and the rectangles are p_i and N_i, respectively, and the row of numbers under the machines represent the estimates of partial derivatives $\frac{\partial PR}{\partial p_i}$ evaluated by numerical simulations. Clearly, the bottleneck machines are m_2 (in Figure 5.2(a)) and m_4 (in Figure 5.2(b)), none of which corresponds to the worst machine (i.e., the machine with the smallest p_i). In fact, m_2 in Figure 5.2(a) is the best machine in the system.



(a). The best machine is the bottleneck.



(b). The worst machine is not the bottleneck.

Figure 5.2: Examples of bottleneck machines in Bernoulli lines

Similarly, a machine with the largest work-in-process in front of it is not necessarily the bottleneck. An example is given in Figure 5.2(b), where m_2 has the largest WIP to be processed, while the BN-m is m_4.

While Definition 5.2 provides a formal characterization of the BN-m, its practical application is not straightforward because the partial derivatives involved could be neither measured on the factory floor nor calculated analytically with any acceptable accuracy. Indeed, while PR itself is often measured on the factory floor, its sensitivity to p_i's is not and could hardly be expected to be measured, since it would require increasing efficiency of each machine and evaluating the resulting increase in PR. Such a procedure is hardly possible in most practical situations. Although the analytical calculations, described in Chapter 4, lead to acceptable estimates $\widehat{PR}$, no analytical methods for evaluating $\partial PR/\partial p_i$ are available. Therefore, to make Definition 5.2 practical, it has to be reformulated in terms of quantities, which are either available through measurements on the factory floor or through analytical calculations or both. It is shown in Subsections 5.2.2 and 5.2.3 that this can be accomplished using the probabilities of blockages and starvations, BL_i and ST_i, defined in Chapter 3 and evaluated in Chapter 4.

The above arguments notwithstanding, the machine with the smallest p_i is, in fact, the bottleneck in the sense of (5.30) if the production line is WF-unimprovable. This follows from the following

Theorem 5.7 *In WF-unimprovable Bernoulli lines defined by assumptions* *(a)-(e) of Subsection 4.2.1,*

$$\frac{\partial \widehat{PR}}{\partial p_i} p_i = const, \qquad i = 1, \ldots, M. \tag{5.31}$$

Proof: See Section 20.1.

Thus, the BN-m in WF-unimprovable lines can be identified quite easily.

Bottleneck buffer: While the term "bottleneck machine" is widely used in practice, the term "bottleneck buffer" is not. This happens, perhaps, because thinking locally, one usually pays more attention to the efficiency of the machines, as part-producing devices, and less to buffers, as "shock absorbers" of perturbations. Nevertheless, the notion of a bottleneck buffer could and, moreover, must be introduced in order to explore all means of system improvements.

Definition 5.3 *Buffer b_i, $i \in \{1, \ldots, M-1\}$, is the bottleneck buffer (BN-b) of a Bernoulli line if*

$$\begin{aligned} &PR(p_1, \ldots, p_M, N_1, \ldots, N_i + 1, \ldots, N_{M-1}) \\ &\quad > PR(p_1, \ldots, p_M, N_1, \ldots, N_j + 1, \ldots, N_{M-1}), \quad \forall j \neq i. \end{aligned} \tag{5.32}$$

In other words, BN-b is the buffer, which leads to the largest increase of the PR if its capacity is increased by 1, as compared with increasing any other buffer in the system.

The buffer with the smallest capacity is not necessarily the BN-b. An example is shown in Figure 5.3, where the numbers under each buffer correspond to the PR of the system obtained by simulations when the capacity of this buffer is increased by one. Clearly, the BN-b is b_1 while the smallest buffer is b_3.

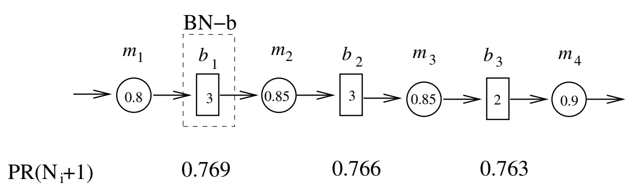


Figure 5.3: Example of bottleneck buffer in a Bernoulli line

To identify the BN-b using Definition 5.3, one would have to experiment with the system by increasing each buffer and measuring the resulting production rate, which is hardly possible in practice. It turns out that this is also unnecessary: as it is shown below, BL_i and ST_i can be used to identify not only the BN-m but also the BN-b.

5.2.2 Identification of bottlenecks in two-machine lines

Theorem 5.8 *For a two-machine Bernoulli line defined by assumptions (a)-(e) of Subsection 4.1.1, the inequality*

$$\frac{\partial PR}{\partial p_1} > \frac{\partial PR}{\partial p_2} \quad (respectively, \frac{\partial PR}{\partial p_1} < \frac{\partial PR}{\partial p_2}) \tag{5.33}$$

takes place if and only if

$$BL_1 < ST_2 \quad (respectively, BL_1 > ST_2).$$

Proof: See Section 20.1.

There are three benefits offered by this theorem. First, it provides a relationship between the "non-measurable" and "non-calculable" partial derivatives of PR and "measurable" and "calculable" probabilities of blockages and starvations. Second, it offers a possibility of identifying the BN-m without even knowing parameters of the machines and buffer, but just by measuring ST_2 and BL_1. Third, it offers a simple graphical way of representing the BN-m. To illustrate this, consider the production line of Figure 5.4, where the two rows of numbers under the machines represent ST_i and BL_i. Place an appropriate inequality sign between ST_2 and BL_1 and turn the inequality into an arrow by adding a line within the sign of the inequality. According to Theorem 5.8, the machine, to which the arrow is pointed, is the BN-m. As it turns out, this procedure can be extended to $M > 2$-machine lines as well.

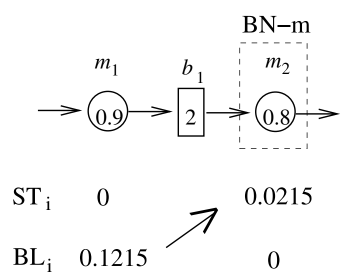


Figure 5.4: Arrow-based method of bottleneck identification

Note that for the case of $M = 2$,

- the problem of BN-b does not arise;
- the machine with the smallest p_i is the BN-m since, as it follows from the results of Section 4.1, $ST_2 > BL_1$, if and only if $p_1 < p_2$.

5.2.3 Identification of bottlenecks in $M > 2$-machine lines

Bottleneck Indicator: As it was discussed in Chapter 4, ST_i and BL_i in lines with $M > 2$ machines cannot be calculated exactly, and only estimates, $\widehat{ST}_i$

and $\widehat{BL}_i$, are available. Therefore, the rules for BN identification in $M > 2$-machine lines are formulated either in terms of ST_i and BL_i, which may be available from factory floor measurements, or in terms of $\widehat{ST}_i$ and $\widehat{BL}_i$, which may be calculated using (4.30) and (4.39), (4.40)). Note that the application of the former requires no knowledge of the machine and buffer parameters.

Consider the production lines shown in Figures 5.5 and 5.6 with two rows of numbers under each machine, the first one indicating ST_i and the second BL_i. Place arrows directed from one machine to another in the same manner as in Subsection 5.2.2, i.e., according to the following rule

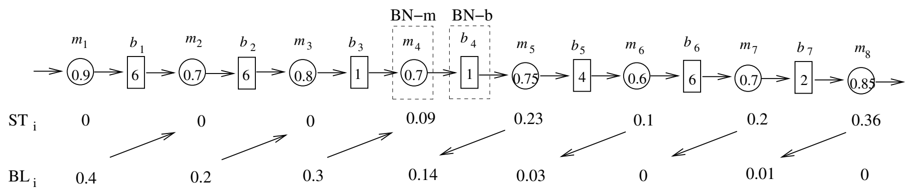


Figure 5.5: Illustration of a Bernoulli line with a single bottleneck machine

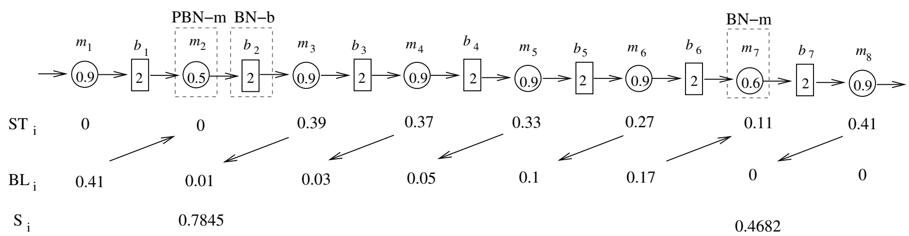


Figure 5.6: Illustration of a Bernoulli line with multiple bottleneck machines

Arrow Assignment Rule 5.1: *If $BL_i > ST_{i+1}$, assign the arrow pointing form m_i to m_{i+1}. If $BL_i < ST_{i+1}$, assign the arrow pointing from m_{i+1} to m_i.*

Bottleneck Indicator 5.1: *In a Bernoulli line with $M > 2$ machines,*

- *if there is a single machine with no emanating arrows, it is the BN-m;*
- *if there are multiple machines with no emanating arrows, the one with the largest severity is the Primary BN-m (PBN-m), where the severity of each (local) BN-m is defined by*

$$\begin{aligned} S_i &= |ST_{i+1} - BL_i| + |ST_i - BL_{i-1}|, \qquad i = 2, \ldots, M-1, \\ S_1 &= |ST_2 - BL_1|, \\ S_M &= |ST_M - BL_{M-1}|; \end{aligned} \tag{5.34}$$

- *the BN-b is the buffer immediately upstream of the BN-m (or PBN-m) if it is more often starved than blocked, or immediately downstream of the BN-m (or PBN-m) if it is more often blocked than starved.*

Thus, according to this indicator, m_4 and b_4 are the bottlenecks in Figure 5.5, and m_2 and b_2 are the PBN-m and BN-b in Figure 5.6.

Bottleneck Indicator 5.1 is justified below using both numerical and analytical approaches.

Numerical justification: It is carried out by calculations and simulations. The calculation approach consists of calculating $\widehat{ST}_i$ and $\widehat{BL}_i$, identifying BN-m and BN-b using Bottleneck Indicator 5.1, and then verifying the conclusions using the calculated quantities $\Delta\widehat{PR}/\Delta p_i$, $\widehat{PR}(N_i+1)$ and Definitions 5.2 and 5.3. The simulation approach is carried out analogously but using ST_i, BL_i, and $\Delta PR/\Delta p_i$, $PR(N_i+1)$ evaluated numerically based on Simulation Procedure 4.1. In both calculation and simulation approaches, Δp_i was selected as 0.03 and $\frac{\Delta\widehat{PR}}{\Delta p_i}$ (or $\frac{\Delta PR}{\Delta p_i}$) was evaluated as

$$\frac{\Delta\widehat{PR}}{\Delta p_i} = \frac{\widehat{PR}(p_1,\dots,p_i+\Delta p_i;\dots,p_M) - \widehat{PR}(p_1,\dots,p_i,\dots,p_M)}{\Delta p_i}.$$

In the majority of cases analyzed, Bottleneck Indicator 5.1 identified BNs correctly. Typical examples are shown in Figures 5.7 (single BN case) and 5.8 (multiple BN case). Some counterexamples, however, have also been discovered. Two of them are shown in Figures 5.9 and 5.10 for single and multiple BN cases, respectively.

To investigate the "frequency" of the counterexamples, the following statistical experiment was carried out: 5000 five-machine serial lines were constructed by selecting p_i's and N_i's randomly and equiprobably from the sets

$$\begin{aligned} p_i &\in \{0.75, 0.80, 0.85, 0.90.0.95\}, \\ N_i &\in \{1,2,3\}, \end{aligned}$$

respectively. For each of these lines, BNs were identified using the calculation and simulation approaches, and the percent of correct and incorrect BN identifications were evaluated. The results are shown in Figures 5.11 and 5.13 for calculation and simulation approaches, respectively. Figure 5.11(a) indicates that among the 5000 lines, investigated by calculations, about 80% had a single BN-m. Within those, Bottleneck Indicator 5.1 identified the BN-m and BN-b correctly in over 95% and 87% of cases, respectively (Figures 5.11(b) and (c)). The lower accuracy of BN-b identification is, perhaps, due to the fact that, unlike p_i's, buffer capacity cannot be increased infinitesimally. This conjecture is supported by Figure 5.11(d), which shows that the BN-b, identified by Definition 5.3, is one of the buffers around the BN-m in over 91% of cases. Thus, Figures 5.11(b)-(d) indicate that Bottleneck Indicator 5.1 is a sufficiently reliable tool for BN-m and BN-b identification. Note that, as it is shown in Figure 5.12(a), the BN-m is the worst machine of the system in only 62% of cases analyzed; thus, assuming that the worst machine is the bottleneck leads to a much lower frequency of correct BN identification.

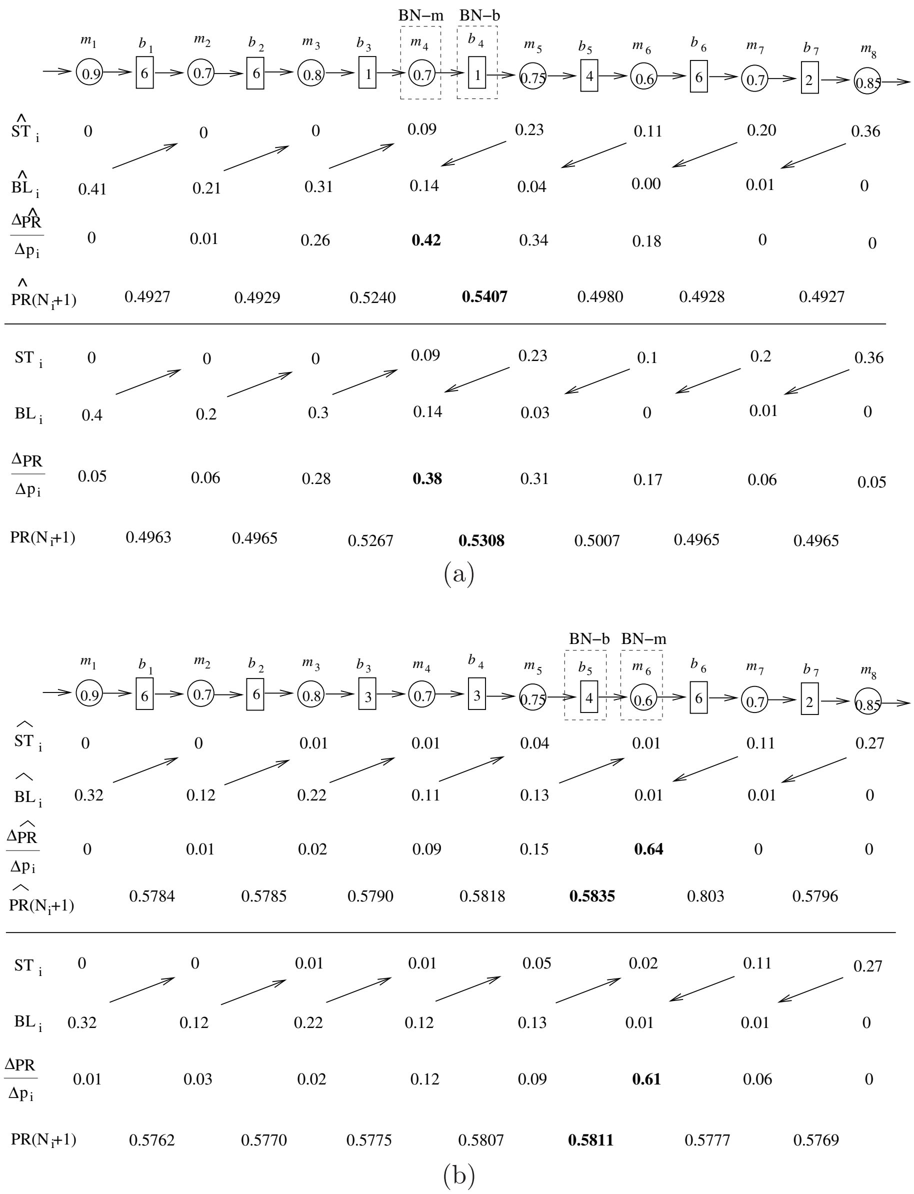
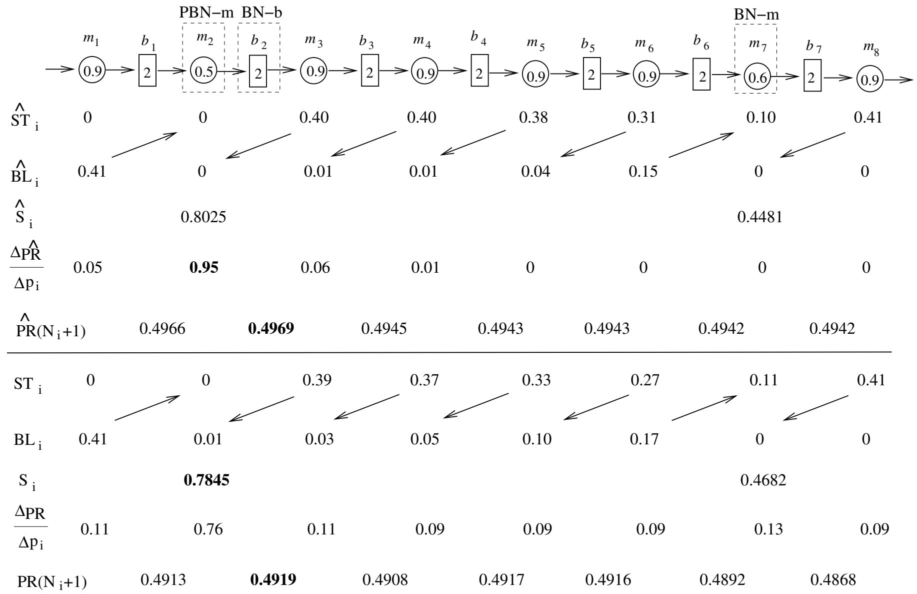
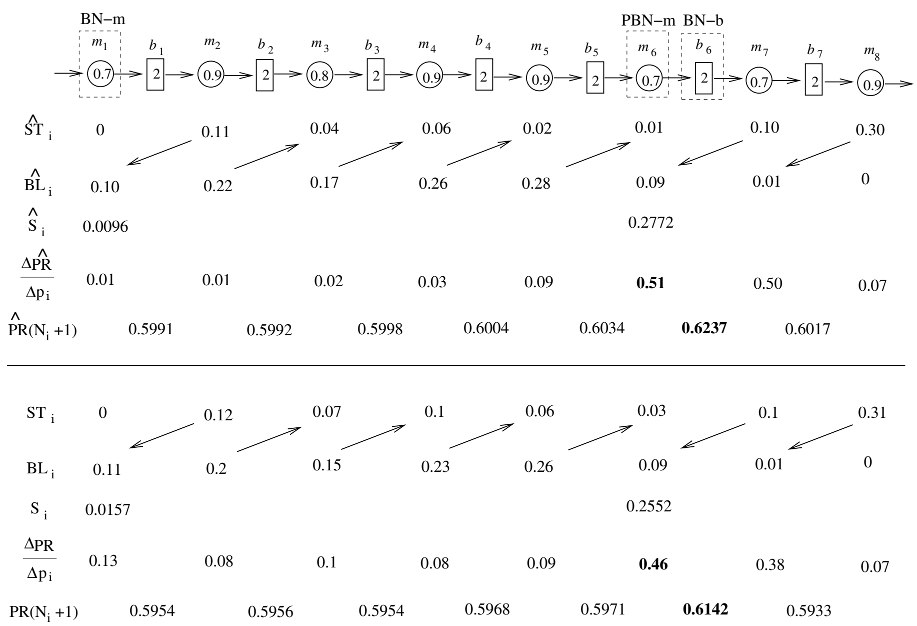


Figure 5.7: Examples of bottleneck identification using Bottleneck Indicator 5.1; single bottleneck case



(a)



(b)

Figure 5.8: Examples of bottleneck identification using Bottleneck Indicator 5.1; multiple bottlenecks case

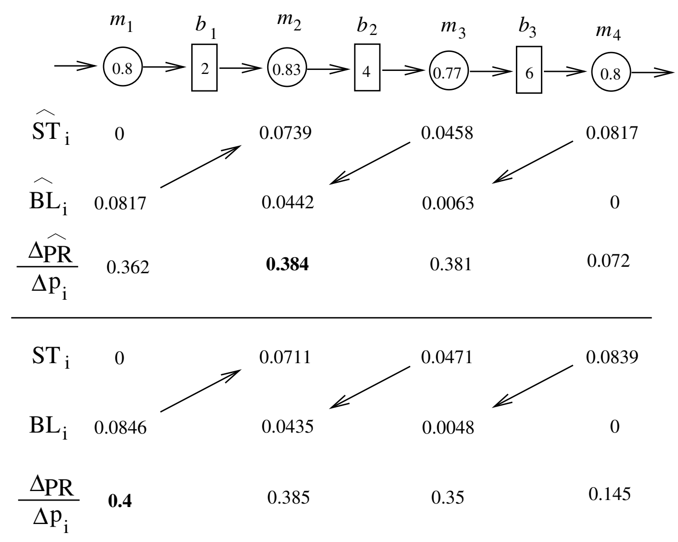


Figure 5.9: Counterexample for Bottleneck Indicator 5.1, single bottleneck case

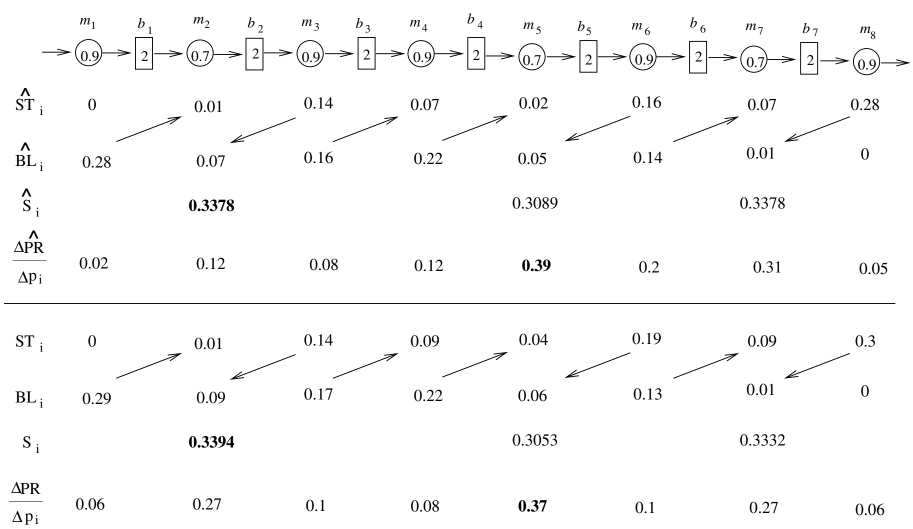
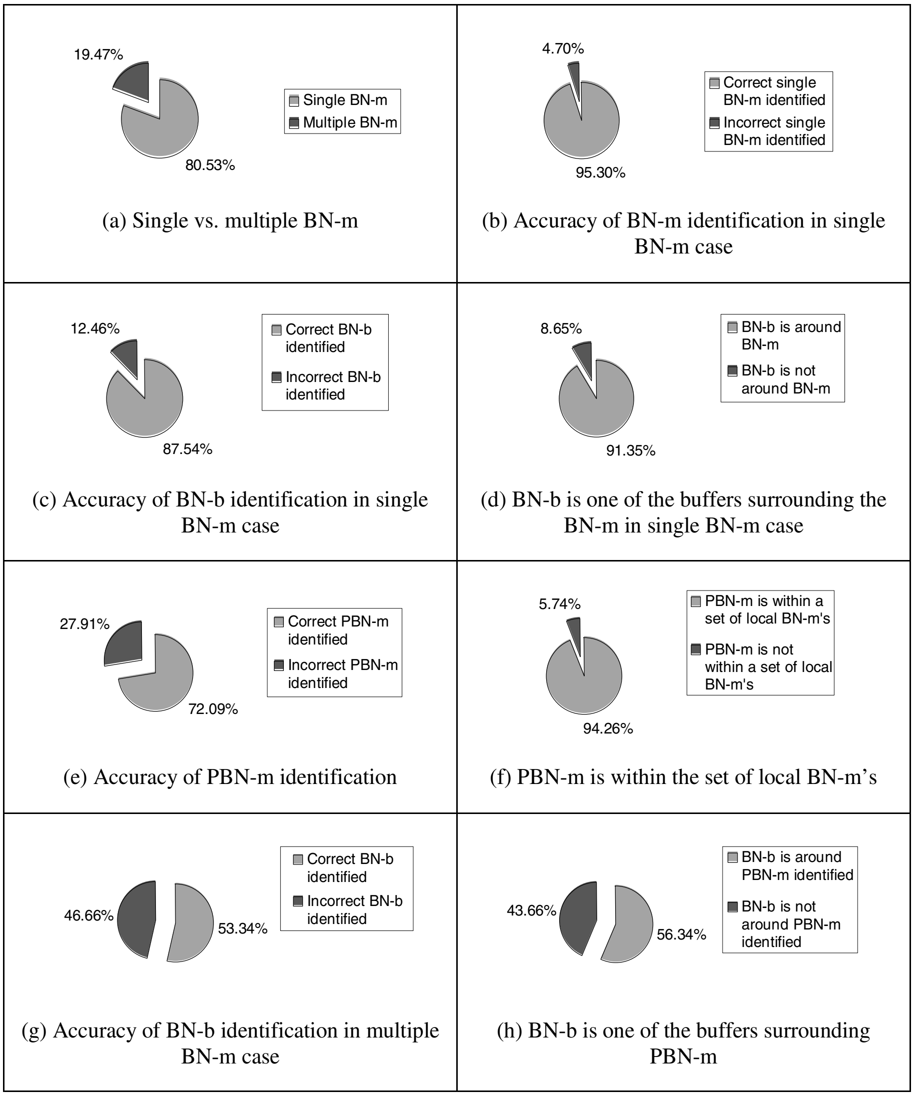


Figure 5.10: Counterexample for Bottleneck Indicator 5.1, multiple bottlenecks case



(a) Single vs. multiple BN-m

(b) Accuracy of BN-m identification in single BN-m case

(c) Accuracy of BN-b identification in single BN-m case

(d) BN-b is one of the buffers surrounding the BN-m in single BN-m case

(e) Accuracy of PBN-m identification

(f) PBN-m is within the set of local BN-m's

(g) Accuracy of BN-b identification in multiple BN-m case

(h) BN-b is one of the buffers surrounding PBN-m

Figure 5.11: Accuracy of Bottleneck Indicator 5.1 using calculation data

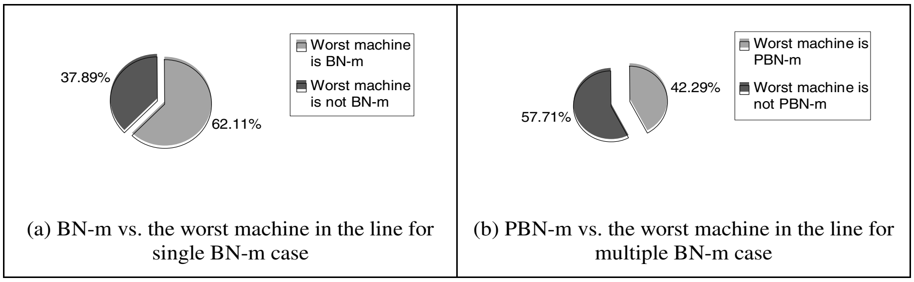


Figure 5.12: Frequency of worst machine being BN-m using calculation data

The frequency of correct PBN identification using Bottleneck Indicator 5.1 for multiple bottlenecks case is illustrated in Figure 5.11(e). Although it is relatively low (about 72%), the true PBN is within the set of local BN-m's in over 94% of cases (Figure 5.11(f)). The frequency of correct BN-b identification is illustrated in Figures 5.11(g) and (h)). Clearly, accuracy of PBN-m identification is lower than that of just a bottleneck. This is, perhaps, due to the fact that the bottleneck severity (5.34) has been defined in an ad-hoc manner and better definitions might be possible. Nevertheless, we conclude that Bottleneck Indicator 5.1 is a useful tool for bottleneck identification, especially taking into account that in only 42% of cases analyzed the worst machine was indeed the PBN (see Figure 5.12(b)).

Similar results were obtained using the simulation approach. These results are illustrated in Figures 5.13 and 5.14.

Thus, the conclusion from this investigation is that Bottleneck Indicator 5.1 provides a sufficiently accurate method for identifying BNs using either measured quantities ST_i and BL_i or calculated ones $\widehat{ST}_i$ and $\widehat{BL}_i$. We also remark that, in our experience, even when Bottleneck Indicator 5.1 leads to an incorrect BN identification, the true BN has just a slightly higher $\Delta PR/\Delta p_i$ and $PR(N_i + 1)$ than those identified by the indicator. Based on the above, we formulate

BN-Continuous Improvement Procedure 5.1:

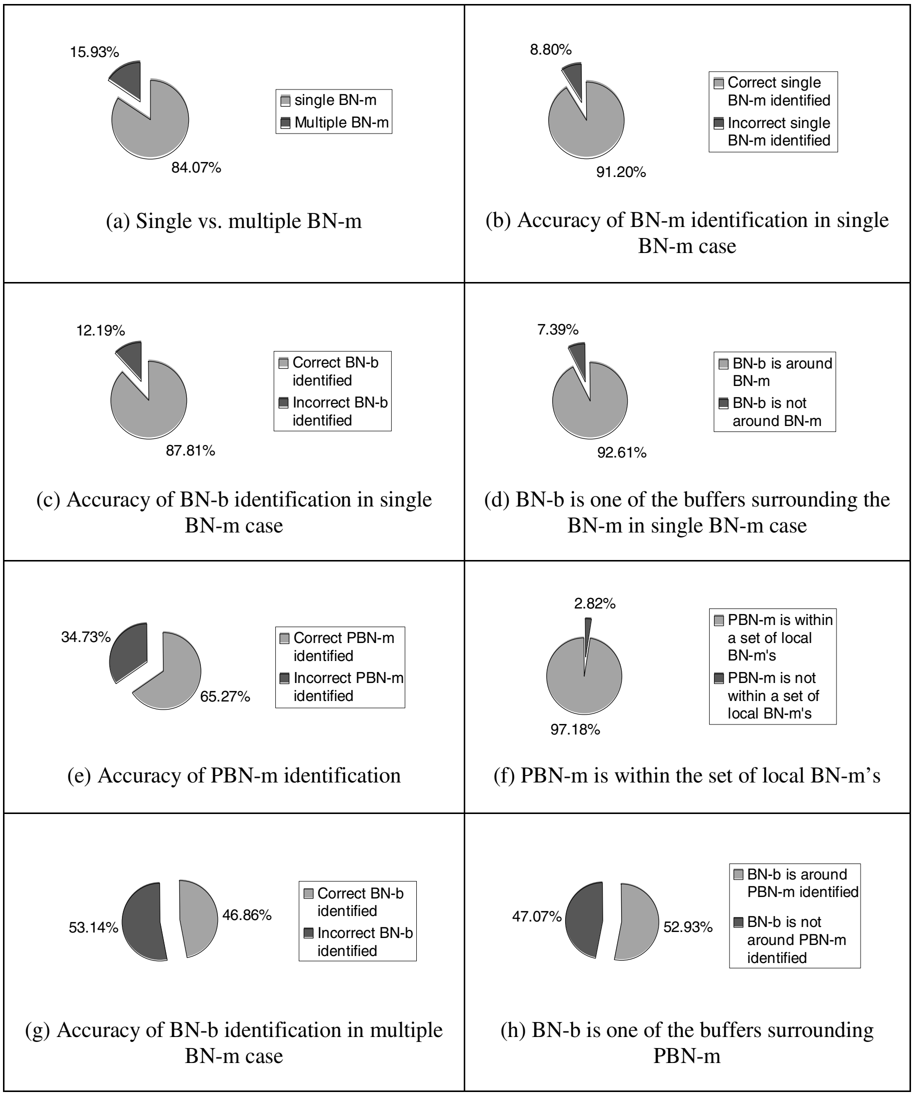
(1) By off-line calculations (using (4.30), (4.39), and (4.40)) or by measurements on the factory floor, evaluate the probabilities (or frequencies) of blockages and starvations of each machine.

(2) Using Bottleneck Indicator 5.1, identify the BN-m (or PBN-m) and BN-b.

(3) Take actions to increase the efficiency, p_i, of this machine (for instance, by improved preventative maintenance, assigning additional work force and the like).

(4) If the above, for one reason or another, is impossible, increase the BN-b or both buffers around the BN-m (or PBN-m).

(5) Return to step (1).



(a) Single vs. multiple BN-m

(b) Accuracy of BN-m identification in single BN-m case

(c) Accuracy of BN-b identification in single BN-m case

(d) BN-b is one of the buffers surrounding the BN-m in single BN-m case

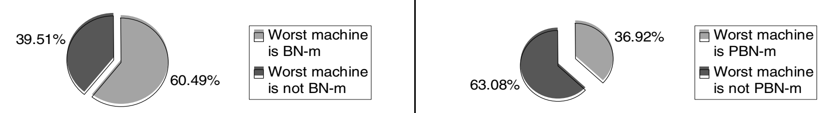
(e) Accuracy of PBN-m identification

(f) PBN-m is within the set of local BN-m's

(g) Accuracy of BN-b identification in multiple BN-m case

(h) BN-b is one of the buffers surrounding PBN-m

Figure 5.13: Accuracy of Bottleneck Indicator 5.1 using simulation data



(a) BN-m vs. the worst machine in the line for single BN-m case

(b) PBN-m vs. the worst machine in the line for multiple BN-m case

Figure 5.14: Frequency of worst machine being BN-m using simulation data

Our industrial experience, gained through numerous case studies, indicates that *BN-Continuous Improvement Procedure 5.1* (and its generalization for production lines with exponential and general models of machine reliability - see Parts III and IV) *is one of the most efficient ways of managing production systems.*

Analytical justification: An analytical justification is available only for the case of a single bottleneck machine. It is based on the following

Hypothesis 5.1 *Inequalities*

$$BL_{j-1} > ST_j, \quad j = 2, \ldots, M$$

and

$$BL_j < ST_{j+1}, \quad j = 1, \ldots, M-1$$

imply, respectively, that

$$\epsilon_{j1} := P_{j-1}(0) \approx Q(p^f_{j-1}, p^b_j, N_{j-1}) \ll 1, \tag{5.35}$$

and

$$\epsilon_{j2} := (1 - p^b_{j+1})P_j(N_j) \approx Q(p^b_{j+1}, p^f_j, N_j) \ll 1. \tag{5.36}$$

The following lemma states that Hypothesis 5.1 indeed holds, at least for N_j sufficiently large:

Lemma 5.1 *In a Bernoulli line defined by assumptions (a)-(e), for any* $0 < \epsilon_0 \ll 1$*, there exists* N^* *such that if* $N_j > N^*$*,* $j = 1, \ldots, M-1$*, then*

$$\epsilon = \max(\epsilon_{j1}, \epsilon_{j2}) < \epsilon_0.$$

Proof: See Section 20.1.

Theorem 5.9 *Under Hypothesis 5.1, the BN-m is downstream of* m_j *if* $\widehat{BL}_j > \widehat{ST}_{j+1}$ *and upstream of* m_j *if* $\widehat{BL}_{j-1} < \widehat{ST}_j$.

Proof: See Section 20.1.

Thus, this theorem confirms Bottleneck Indicator 5.1 as far as a single BN-m is concerned.

PSE Toolbox: The method of BN-m and BN-b identification, using both calculated and measured data, is implemented in the **Bottleneck Identification** function of the toolbox. For a description of these tools, see Subsections 19.5.1 and 19.5.3.

5.2.4 Potency of buffering

As it has been shown above, the worst machine often is not the BN of the system. Why does this happen? Clearly, this is because of an inappropriate buffer capacity allocation. To formalize the notion of buffering quality, introduce

Definition 5.4 *The buffering of a production system is*

- *weakly potent if the BN-m is the worst machine in the system (i.e., the machine with the smallest efficiency); otherwise, it is not potent;*
- *potent if it is weakly potent and its production rate is sufficiently close to the BN-m efficiency (e.g., within 5% of the BN machine efficiency in isolation);*
- *strongly potent if it is potent and the system has the smallest possible total buffer capacity (i.e.,* $N^* = \sum_{i=1}^{M-1} N_i$ *is the smallest possible to ensure the desired production rate).*

For serial lines with Bernoulli machines, one can determine if the buffering is weakly potent or not using the method of BN-m identification described in this section. To determine if it is potent, the method of PR calculation, described in Chapter 4, can be used. To investigate the notion of strong potency, methods that allow one to calculate the smallest total buffers capacity, which is necessary to ensure the desired production rate, must be available; these methods are described in Chapter 6.

For serial lines with exponential and general models of machine reliability, similar techniques are discussed in Part III, and for assembly systems in Part IV.

Along with its practical utility, the notion of buffering potency has conceptual significance. Indeed, production systems consist of two distinct entities: machines and buffers. The quality of the machines is characterized by their efficiency. In practice, machine efficiency is often monitored, and continuous improvement efforts are largely centered on its modification. In contrast, the quality of buffering is rarely monitored and even more rarely viewed as a resource for continuous improvement. The quantification provided by Definition 5.4 brings buffering to the same level of monitoring potential as that for machine efficiency.

To illustrate the importance of buffering potency, consider the automotive ignition module assembly system described in Sections 3.2 and 3.10. Its throughput losses due to machines and due to MHS are analyzed in Table 5.1. Here, the losses due to machines are evaluated as the difference between the nominal throughput (600 parts/h) and the isolation throughput of the worst machine; the losses due to MHS are evaluated as the difference between the isolation throughput of the worst machine and the actual throughput of the system. As follows from these data, out of roughly 240 parts/h lost, 80 parts/h are attributed to the machines and 160 parts/h to MHS. Clearly, this MHS is non-potent. (Chapter 16 shows that the worst machine is not the bottleneck of the system.) Interestingly, this 1:2 ratio has been observed in other production systems as well. Thus, ensuring potency of MHS is an important resource of production system improvements. A detailed analysis of the automotive ignition module assembly system is described in Chapter 16.

Table 5.1: Losses analysis in automotive ignition module assembly system

Month	May	June	July	Aug.	Sept.	Oct.
Isolation TP of the slowest machine (parts/h)	522	534	468	498	540	492
Losses due to machine (parts/h)	78	66	132	102	60	108
TP of the system (parts/h)	337	347	378	340	384	383
Losses due to MHS (parts/h)	185	187	90	158	156	109

5.2.5 Designing continuous improvement projects

Based on the methods developed in this chapter, the design of continuous improvement (CI) projects can be carried out following the procedure illustrated in Figure 5.15. As it is shown in this figure, after modeling the production system at hand and model validation (as described in Chapter 3), methods of constrained and/or unconstrained improvability (Chapter 5) can be used to determine possible avenues for system improvement, and the most efficient one is identified (using the performance evaluation techniques of Chapter 4). In this manner, rigorous improvement projects with quantitatively predicted results can be designed. Following their implementation and evaluation on the factory floor, the process must be repeated anew, in a never ending quest for improvement.

5.3 Measurement-based Management of Production Systems

The process of designing continuous improvement projects described above requires mathematical modeling of the production systems at hand. In particular,

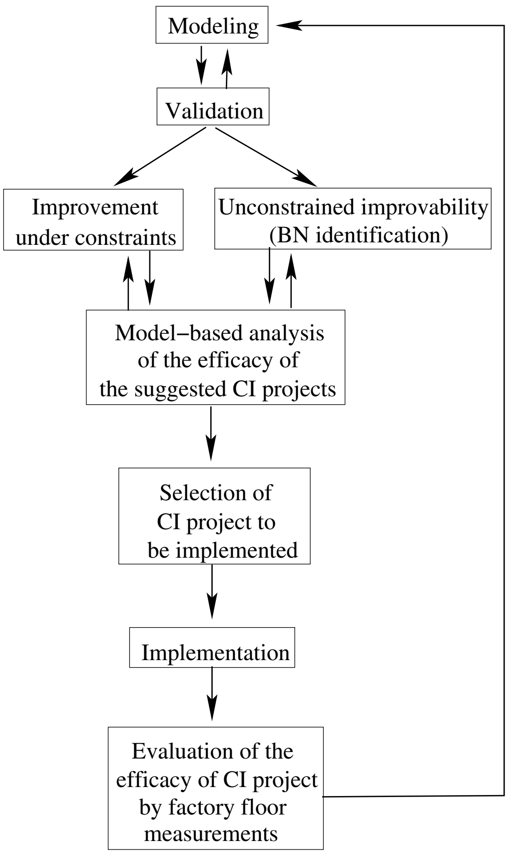


Figure 5.15: Procedure for designing continuous improvement projects

it requires a relatively detailed block-diagram of the system and identification of machine and buffer parameters. This information may be difficult to obtain on the factory floor and, more importantly, maintain on a daily basis. Therefore, a simpler method is desirable to exercise daily managerial duties. The bottleneck identification technique of Section 5.2 leads to such a simpler method. It is illustrated in Figure 5.16 and is referred to as *Measurement-based Management* (MBM) of production systems.

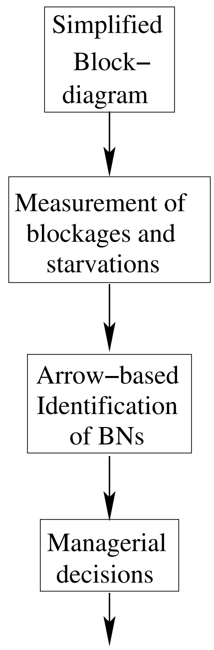


Figure 5.16: Procedure for Measurement-based Management

MBM consists of the following:

- *Simplified block-diagram.* Every manager must maintain a simplified block-diagram of his/her production system. This diagram should include all major operations and their interconnections but may omit conveyors, buffers and other elements of material handling. For example, an automotive assembly plant manager may have the simplified block-diagram as shown in Figure 5.17. For a paint shop manager, the simplified block-diagram may be as the one of Figure 5.18.

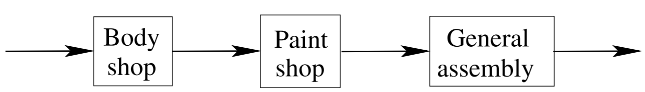


Figure 5.17: Simplified block diagram of automotive assembly plant

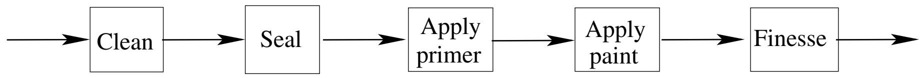


Figure 5.18: Simplified block diagram of automotive paint shop

- *Measurements of blockages and starvation for each block of the simplified block-diagram.* These measurements must be carried out on a continuous basis, either automatically or manually – whichever method is available. Then, estimates of the frequencies of blockages and starvations can be computed as

$$BL_i = \frac{\text{Time of blockages of operation } i}{\text{Total time of observation}},$$
$$ST_i = \frac{\text{Time of starvations of operation } i}{\text{Total time of observation}}.$$

  In practice, these calculations must be carried out on a daily or weekly basis, depending on the scale of the system. For instance, a plant manager may have these data on a weekly basis, while a paint shop manager on a daily basis or even for each shift separately.

- *Identification of bottlenecks.* Using the measured data and the method of Subsection 5.2.3, the bottleneck block must be identified. This is illustrated in Figures 5.19 and 5.20 for the assembly plant and paint shop, respectively.

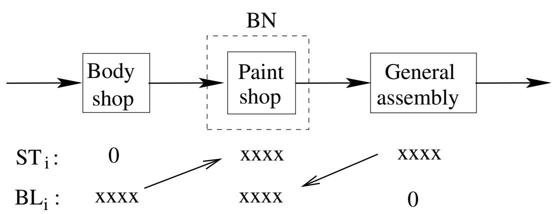


Figure 5.19: Illustration of bottleneck identification for automotive assembly plant

- *Managerial decisions.* Using the information derived above, managers should develop and implement actions, which would improve the performance of the bottleneck block.

Similar to a physician, who cannot treat a patient without taking vital signs, no production system should be "treated" without measuring its "vital signs." It is shown in this chapter that:

The "vital signs" that characterize a production system as a whole are blockages and starvations.

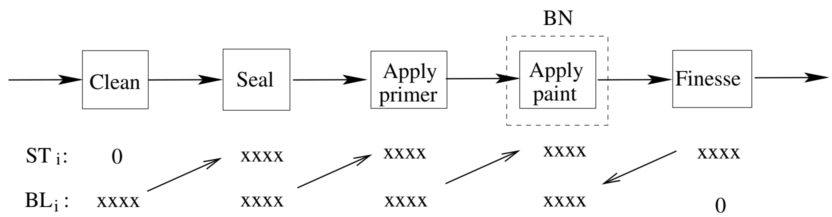


Figure 5.20: Illustration of bottleneck identification for automotive paint shop

5.4 Case Studies

5.4.1 Automotive ignition coil processing system

The mathematical model of this system is constructed in Subsection 3.10.1 and validated in Subsection 4.4.1. Below, we analyze the possibilities of its performance improvement using the methods developed in this chapter.

Based on the model for Period 1 and expressions (4.39), (4.40), we calculate the probabilities of blockages and starvations of all machines in the system. The results are shown in Figure 5.21. Arranging the arrows according to Arrow Assignment Rule 5.1, and using Bottleneck Indicator 5.1, we find that the BN-m and BN-b are m_{9-10} and b_{9-10}. Increasing the capacity of the BN-b by 1 leads to $\widehat{PR} = 0.8475$ and $\widehat{TP} = 476.7$ parts/hour. In addition, increasing the

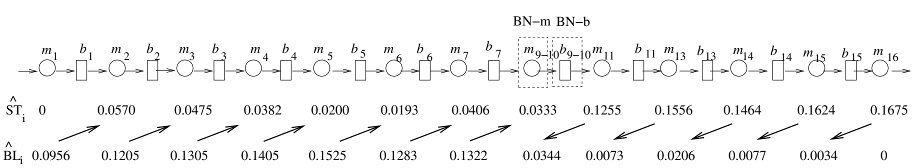


Figure 5.21: Bottleneck identification in coil processing system (Period 1)

efficiencies of m_{9-10} by 10% leads to $\widehat{PR} = 0.8976$ and $\widehat{TP} = 505$ parts/hour.

After these improvements, the bottleneck shifts to m_6 and m_1 with m_6 being the primary bottleneck machine and b_5 the bottleneck buffer (Figure 5.22). Increasing the capacity of b_5 by 1, we obtain $\widehat{PR} = 0.9092$ and $\widehat{TP} = 511.4$ parts/hour. Thus, these two steps of improvement result in significant recovery of losses in the ignition coil processing system. A similar conclusion is obtained using the model for Period 2.

Using the B-exp transformation of Subsection 3.9.4, we obtain the transformed machine average uptime and buffer capacity as follows:

$$T^{tr}_{up,9-10} = 123.11 \text{ min}, \quad N^{tr}_5 = 17, \quad N^{tr}_7 = 10.$$

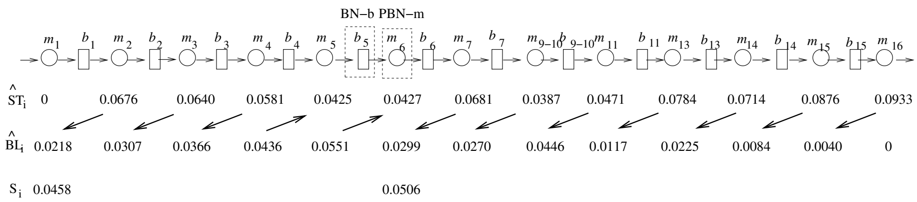


Figure 5.22: Bottleneck identification in improved coil processing system (Period 1)

5.4.2 Automotive paint shop production system

The mathematical model of this system is constructed in Subsection 3.10.2 and validated in Subsection 4.4.2. Below, we analyze its performance improvement.

Based on the model for Month 1 and expressions (4.39) and (4.40), we calculate the probabilities of blockages and starvations of all machines. The results are shown in Figure 5.23. Using Bottleneck Indicator 5.1, we find that the BN-m is m_3.

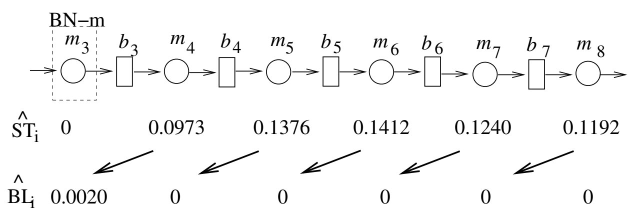


Figure 5.23: Bottleneck identification for paint shop system (Month 1)

The main reason for m_3 to be the bottleneck is starvation by empty carriers. Assuming the empty carriers are always available so that the starvation probability P_{st} can be eliminated, we obtain $\widehat{PR} = 0.9411$ and $\widehat{TP} = 59.29$ jobs/hour. Again, since all buffers are large, machine m_3 is the system bottleneck due to its relatively low reliability (see Figure 5.24). Increasing the efficiency of m_3 by 4% leads to $\widehat{PR} = 0.9558$ and $\widehat{TP} = 60.22$ jobs/hour, and machine m_4 becomes the new bottleneck (Figure 5.25). Thus, these improvements result in almost complete recovery of losses in the paint shop system. Similar conclusions are obtained using the models for Months 2 - 5.

5.5 Summary

- Production systems can be improved in a constrained or unconstrained scenario. In the constrained case, the system is improvable if its work

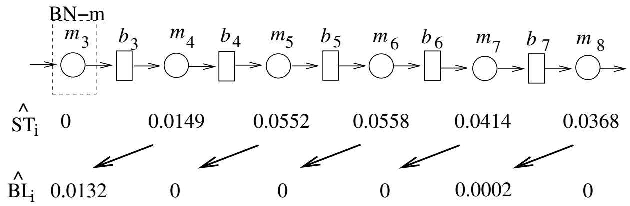


Figure 5.24: Bottleneck identification for paint shop system without starvation by empty carriers (Month 1)

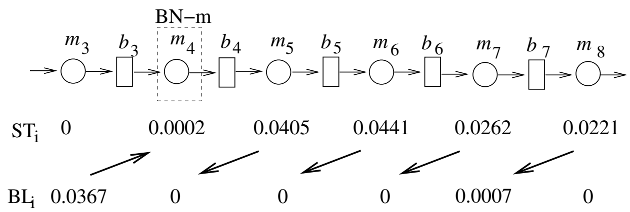


Figure 5.25: Bottleneck identification for improved paint shop system (Month 1)

force (WF) and/or buffer capacity (BC) can be re-allocated among various operations so that the production rate is increased. In the unconstrained case, improvement involves identifying the machine and/or buffer that impedes the system performance in the strongest manner, followed by improving either this machine, or this buffer, or both.

- Production lines with Bernoulli machines are unimprovable with respect to WF re-allocation if each buffer is, on average, close to being half full.
- Production lines with Bernoulli machines are unimprovable with respect to WF and BC re-allocation simultaneously if all buffers are of equal capacity and, on average, are close to being half full.
- Production lines with Bernoulli machines are unimprovable with respect to BC re-allocation if the average occupancy of buffer i is close to the average availability of buffer $i+1$, $i = 1, \ldots, M-2$.
- A machine is the bottleneck machine (BN-m) of a Bernoulli line if increasing its efficiency has the largest effect on the production rate of the line.
- A buffer is the bottleneck buffer (BN-b) of a Bernoulli line if increasing its capacity has the largest effect on the production rate of the line.
- Both BN-m and BN-b can be identified during normal system operation by measuring machine blockages and starvations and using Bottleneck Indicator 5.1.

- The buffering of a production line is potent if the BN-m is indeed the machine with the smallest efficiency.
- Identifying the BN-m can be used as a basis for Measurement-based Management of production systems.

5.6 Problems

Problem 5.1 Consider a five-machine line with Bernoulli machines. Assume $N_1 = N_4 = 1$, $N_2 = N_3 = 3$ and the product of all p_i's of the machines is $(0.9)^5$.

(a) Design an unimprovable (i.e., optimal) system with respect to WF.

(b) Does the inverted bowl phenomenon take place? Explain why it does or does not.

Problem 5.2 Consider the same system as in Problem 5.1.

(a) Using WF-Continuous Improvement Procedure 5.1, obtain an unimprovable system.

(b) Does it coincide with the one obtained in Problem 5.1? Explain why it does or does not.

(c) Calculate $\widehat{PR}$ of the design obtained and compare it with that ensured by the design of Problem 5.1.

Problem 5.3 Consider again a five-machine production line where the sum of N_i's is 8 and the product of p_i's is $(0.9)^5$.

(a) Design a system that is unimprovable with respect to both WF and BC simultaneously.

(b) Compare it with the designs obtained in Problems 5.1 and 5.2 and comment on the differences in structures and production rates of each design.

Problem 5.4 Consider a four-machine Bernoulli line, where the sum of N_i's is 5.

(a) Assume each p_i is 0.8. Using BC-Continuous Improvement Procedure 5.1, find the unimprovable allocation of N_i's. Comment on the shape of this allocation as a function of i.

(b) Assume now that $p_1 = p_3 = 0.7$ and $p_2 = p_4 = 0.9$. Again, using BC-Continuous Improvement Procedure 5.1, determine the unimprovable allocation of N_i's and compare it with the one obtained in (a). Comment on the reason for the differences.

Problem 5.5 Consider the Bernoulli model of the production system analyzed in Problem 3.3.

(a) Determine if this system is WF-improvable. If so, calculate the unimprovable work allocation and the resulting production rate.

(b) Determine if this system is BC-improvable. If so, calculate the unimprovable buffer capacity allocation and the resulting production rate.

(c) Calculate the simultaneously unimprovable WF and BC allocations and the resulting production rate.

(d) Compare the production rates obtained among themselves and with that obtained in Problem 4.9 for the original system.

(e) Which, if any, of the above improvement projects would you recommend for implementation?

Problem 5.6 Consider the Bernoulli model constructed in Problem 3.4 and analyzed in Problem 4.11. Repeat steps (a)-(e) of Problem 5.5.

Problem 5.7 Consider a five-machine serial line with Bernoulli machines. Assume $N_i = 2$, $i = 1, \ldots, 4$, and p_i's are as follows: $[0.9, 0.7, 0.9, 0.9, 0.7]$.

(1) Determine the BN-m and BN-b.

(2) Change the buffer capacities around the BN-m so that the bottleneck moves to another machine. What is the new BN-m? Why did the bottleneck move?

(3) What is the BN-m when all buffers are infinite?

Problem 5.8 Consider the production system analyzed in Problem 5.5.

(a) Identify its BN-m (or PBN-m) and BN-b and determine if the buffering is potent.

(b) Based on this information, design the best, from your point of view, continuous improvement project that results in 10% increase of the production rate of the system (as compared with the original one).

(c) Using the B-exp transformation of Chapter 3, return to the exponential description and formulate measures, which would have to be carried out in order to implement this continuous improvement project.

Problem 5.9 Repeat steps (a)-(c) of Problem 5.8 for the system considered in Problem 5.6.

Problem 5.10 Derive a WF-Improvability Indicator for Bernoulli lines with the symmetric blocking convention.

Problem 5.11 Derive a BC-Improvability Indicator for Bernoulli lines with the symmetric blocking convention.

Problem 5.12 Derive a Bottleneck Indicator for identifying BN-m and BN-b in Bernoulli lines with the symmetric blocking convention.

5.7 Annotated Bibliography

The notion of improvability of serial lines with Bernoulli machines has been introduced in

[**5.1**] D.A. Jacobs and S.M. Meerkov, "A System-theoretic Property of Serial Production Lines: Improvability," *International Journal of System Science*, vol. 26, pp. 95-137, 1995.

A detailed study of improvability under constraints has been carried out in

[**5.2**] D.A. Jacobs and S.M. Meerkov, "Mathematical Theory of Improvability for Production Systems," *Mathematical Problems in Engineering*, vol. 1, pp. 95-137, 1995.

The bowl phenomenon was discovered in

[**5.3**] F.S. Hillier and R.W. Bolling, "The Effect of Some Design Factors on the Efficiency of Production Lines with Variable Operation Times," *Journal of Industrial Engineering*, vol. 27 , pp. 351-358, 1966.

The Theory of Constraints is described in a "manufacturing novel:"

[**5.4**] E.M. Goldratt and J. Cox, *The Goal*, North River Press Inc, Croton-on-Hudson, NY, 1984,

see also

[**5.5**] E.M. Goldratt and R.E. Fox, *The Race*, North River Press Inc, Croton-on-Hudson, NY, 1986.

These are simple and useful readings; they provide a good foundation for qualitative understanding of production systems. A drawback is that they are based just on common sense, without analytical investigations. Is it possible to design, say, Boeing 777, just using common sense? Equally it is impossible to design and manage a production system using only common sense; quantitative methods are necessary.

The definitions of bottleneck machine (5.30) and bottleneck buffer (5.32) have been introduced and investigated in

[**5.6**] C.-T. Kuo, J.-T. Lim and S. M. Meerkov, "Bottlenecks in Serial Production Lines: A System-theoretic Approach," *Mathematical Problems in Engineering*, vol. 2, pp. 233-276, 1996.

More details can be found in

[**5.7**] C.-T. Kuo, *Bottlenecks in Production Systems: A Systems Approach*, Ph.D Thesis, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, 1996.