Simulation Modeling of Automatic Production Lines with Intermediate Buffers

M. Heshmat, Mahmoud A. El-Sharief, M. G. El-Sebaief

Abstract— A production line is an important class of manufacturing system when large quantities of identical or similar products are to be produced. The performance of a production line is highly influenced by machine failures. When a machine fails, it is then be unavailable during a certain amount of time required to repair it. Analysis of production lines divides into three types: analytical, approximation and simulation models. The analytical and approximation models have assumptions which make these models unrealistic such as reliable workstations, certain processing distribution, the first workstation cannot be starved and the last workstation cannot be blocked. The main problems in production lines treatment are the calculation of throughput and average levels of buffers because of the great size of state space. An analytical model is reviewed to clarify the limitations to use such treatment in real production lines. Simulation modeling of production lines is considered very important for designers interested in: Workload Allocation Problem (WAP), Server Allocation Problem (SAP), and Buffer Allocation Problem (BAP). This paper studies and analyzes the performance keys, which effect on production lines. A simulation model is developed by using ARENA software and used to analyze and test several bottlenecks that are causing severe congestions in different areas on the production line and could resolve all of these bottlenecks. In this paper, an actual cement production line is studied. After a simulation time of 13 days, a simulation results show the line bottlenecks, workstations utilization, buffer capacities and the line production rate. The outputs clarify redesign of allocation of buffers, which verify an optimum size could be made; it might be taken into consideration when designers implement such lines. Finally modified better workstations utilization, buffer capacities and the line production rate with an increase about 15% of the production rate and economizing of 37 % from buffer capacities.

Index Terms— Production lines; Buffer allocation; Simulation; Modeling; Case study; ARENA; throughput .

1 INTRODUCTION

A production line is an important class of manufacturing system when large quantities of identical or similar products are to be produced (mass production). The performance of a production line is highly influenced by machine failures. When a machine fails, it is then be unavailable during a certain amount of time required to repair it. When a machine is in a failure status, the number of parts in the upstream buffer tempted to be increased while the number of parts in the downstream buffer tempted to be decreased. If this status persists, the upstream buffer may become full and as a consequence the upstream machine may be blocked which of course, would negatively affect the rate of production. Similarly, the downstream buffer may become empty and, therefore, the downstream machine may be starved. Figure 1 depicts a production line with a k stations and k-1 intermediate buffers.

A great deal of literature has been devoted to the modeling and analyses of production lines since the early 1950’s because of their economic importance as well as academic interest. A comprehensive survey on mathematical models by Dallery and Gershwin,[1], Buzacott and Hanifin [2] and Papadopoulos et al. [3]. Simulation is considered the powerful tool to model a production line with unreliable machines and stochastic variable intermediate buffers to identify the line performance. Papadopoulos et al. [4] stated that “Simulation of production lines is a powerful tool in obtaining the performance measures where analytical methods are either difficult or impossible to use”. Hosseinpour et al. [5] presented a comprehensive literature review on importance of simulation in manufacturing as a very helpful work tool in industrial field to test the system's behavior. Simulation is low cost, secure and fast analysis tool with many different system configurations [5], Hosseinpour et al. [5]investigated the application of simulation that used to address in manufacturing which provides this paper with the following:

- Location and size of inventory buffers,
- Evaluation of the effect of a new piece of equipment on an existing manufacturing system,
- Throughput analysis,
- Bottleneck analysis,
- Times parts spend in queues,
- Queue size,
- Utilization of equipment or personnel.

Kelton et al. [6]presented the concepts of simulation using ARENA to help the modeler reaching the ability to carry out effective simulation modeling. ARENA is based on SIMAN modeling language, and has an object-oriented design to any application area. Many papers have used ARENA software to study production lines, identify the bottlenecks, and resolve it in the design phase or in a standing line. Seraj[7]studied a Rusk production line to increase its capacity using a simulation ARENA model. He simulated the old line to find congestions and bottlenecks then he replaced a machine with a new one and increased the production rate by 50%.

http://www.ijser.org
Hecker et al. [8] analyzed and optimized a bakery production line using ARENA; a one-shift period data was collected, then formulated the model and simulated it, followed by validation of the simulation results with respect to the real data. As equipment utilization affects directly on the line productivity, achieving a possible highly utilization will increase the line productivity, therefore, increase the line performance. This would be achieved based on a perfect preventive and predictive maintenance schedule. Gonca et al. [9] simulated a production line by using an ARENA-based simulation model to select a preventive maintenance schedule, which gives the best utility and performance values.

In this paper, an actual cement production line as a real case study is studied for verification and validation the proposed algorithm. Actual data is collected about each workstation including production capacities, processing times and the intermediate buffer capacities as mentioned in the following sections. One-year failure history data is recorded about each machine from preventive and predictive maintenance department and using ARENA Input Analyzer the most appropriate probability distribution of each unreliable machine is detected. A block diagram of the cement line is established and all needed data is introduced. After a simulation replication time of 13 days, a simulation results are obtained; these results show the line bottlenecks, workstations utilization, buffer capacities and the line production rate. A verification and validation of the model has been done. To resolve the bottlenecks, an improvement was done by rebuilding a modified simulation model, which verifies better performance keys. These keys might be taken into consideration when designers implement such line. Finally, the modified workstations utilization and, buffer capacities increase the line production rate by more than 15% of the production rate and economizing buffer capacities.

2. MODELING OF PRODUCTION LINES

Modeling of production lines divides into three types: analytical, approximation and simulation models. Buzacott and Hanifin [2] have compared seven analytical models of automatic production lines with buffers, but these models have assumptions which make these models unrealistic such as reliable workstations, certain processing distribution, the first workstation cannot be starved and the last workstation cannot be blocked.

Buzacott model assumptions are:

- Operation dependent failure
- Geometric distribution for up and down time
- The probability of two failures or two repairs are negligible

Using Markov chain, an equation (equations (1-4)) is derived for calculating the line efficiency. It is considered an important indicator for line performance and production rate under the previous assumptions.

\[
E = \frac{2 - b(1+x) + l(1+x)}{(1+2x)(2-b(1+x)) + l(1+x)(1+x)} \quad (1)
\]

\[
x = \frac{b}{l} \quad (2)
\]

\[
L = \frac{N}{D} \quad (3)
\]

\[
b = \frac{1}{D} \quad (4)
\]

Where,

- \( D \) = mean down time of station \( Mi \) in minutes
- \( T \) = mean up time in minutes measure for operation dependent failure
- \( N \) = Buffer capacity
- \( E \) = Efficiency

2.1 ANALYTICAL MODELING OF THE BUFFER ALLOCATION PROBLEM

The objective is to maximize the line throughput, subject to a given total buffer space. That is, equations 5.7 [10].

\[
\text{maximize}(N_1, N_2, \ldots, \ldots, N_{n-1}) \quad (5)
\]

Subject to

\[
\sum_{i=1}^{n-1} N_i = K \quad (6)
\]

\[
N_i \geq 0 \quad (7)
\]

The quantity \( N_i \) represents the feasible buffer allocation to the \( i \)th allocation zone \( f(N_1, N_2, \ldots, \ldots, N_{n-1}) \) is the throughput of the production line to be maximized. \( K \) is the total buffer capacity. The number of feasible allocations of \( K \) buffer slots among the \( (n-1) \) intermediate buffer locations increases dramatically with \( K \) and \( n \) and is given by[10]

\[
\frac{(k+n-2)!}{(k+1)(k+2)\ldots\ldots(k+n-2)} \quad (8)
\]

Demir et al. [11] presented an integrated approach to solve the buffer allocation problem in unreliable production lines to maximize the throughput rate of the line with minimum total buffer size.

2.2 MODELING OF TWO-MACHINE PRODUCTION LINES

To find out the size of the problems of mathematical analysis of production lines, an analysis of two machine production line
with phase-type distribution is considered [1].

Assumptions:-

- The processing time distribution of each machine is given in the form of a continuous phase-type distribution.
- The blocking mechanism is blocking before service (BBS).
- The system behavior is a discrete state, continuous time Markov process.

Let $S_i$ be the number of phases of (phase-type distribution) $P_{Hi}$. The behavior of such a system can be characterized by a discrete state, continuous time Markov process. Analyzing this system then reduces to that of calculating the steady-state probabilities of this Markov process.

The state of the Markov process can be expressed as $(n, j_1, j_2)$, where $n$ is the number of parts currently present in the buffer, and $j_1$ is the current phase of service of machine $M_i$, $i = 1, 2$. $n$ can take integer values from 0 to $N$. $j_1$ can take integer values from 1 to $S_1$ except when machine $M_1$ is blocked in which case $j_1 = 0$. Similarly, $j_2$ can take integer values from 1 to $S_2$, except when machine $M_2$ is starved, in which case $j_2 = 0$. The state space is partitioned according to the values of $n$. Let $p$ denote the steady-state probability vector and let $P_n$ denotes the portion of this vector that corresponds to a buffer content of $n$, so,

$$
P = \begin{pmatrix}
p_0 \\
p_1 \\
p_2 \\
p_3 \\
\vdots \\
p_n
\end{pmatrix}
$$

(9)

Note that $P_n$, (equation 9) where $n=1...N-1$ is size $S_1S_2$ while $P_0$ and $P_N$ are of size $S_1$ and $S_2$ respectively. Let $QT$ (Equations (10-12)) denotes the infinitesimal generator of the Markov process. The steady state probability vector $p$ of the Markov process is the solution of the equation $pQT=0$; or, equivalently,

$$
Q^T p = 0
$$

(10)

$1^T p = 1.0$

(11)

Matrix $Q^T$ is a block tridiagonal matrix with the following special structure:

$$
Q^T = \begin{pmatrix}
B_0 & A_0 & 0 & 0 & \cdots & 0 \\
C_0 & B & A & 0 & \cdots & . \\
0 & C & B & A & \cdots & . \\
0 & 0 & C & B & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \ddots \\
0 & \cdots & A & 0 & \cdots & C \\
0 & \cdots & C_N & B_N
\end{pmatrix}
$$

(12)

Where $A$, $B$, and $C$ are square matrices of size $(S_1S_2, S_1)$, and $B_0$ and $B_N$ are square matrices of size $(S_1, S_1)$ and $(S_2, S_2)$; $A_0$, $C_0$, and $C_N$ are matrices of size $(S_1, S_1)$, $(S_1, S_1)$, $(S_2, S_2)$, and $(S_1S_2, S_1S_2)$.

$QT$ has this special structure because the Markov process associated with a two-machine flow line is a generalized birth-death process. Transitions can only occur between states that are neighbors of each other with respect to the value of $n$. That is, the only possible transitions from a state $(n, j_1, j_2)$ are to a state $(n', j'_1, j'_2)$ such that either $n' = n$, or $n' = n-1$, or $n' = n+1$. In addition, transition rates are independent of $n$, for $1 < n < N-1$. Because of the special block tridiagonal structure of $QT$, eq. (12) can be decomposed into the following set of equations (13-17):

$$
B_0 p_0 + A_0 p_1 = 0,
$$

(11)

$$
C_0 p_0 + B p_1 + A p_2 = 0,
$$

(12)

$$
C p_{n-1} + B p_n + A p_{n+1} + a = 0, 1 < n < N - 1
$$

(13)

$$
C p_{N-2} + B p_{N-1} + A p_N = 0,
$$

(14)

$$
C p_{N-1} + B p_N = 0
$$

(15)

Two solution techniques that make use of the special structure of the matrix $QT$ have received special attention. They are known as the recursive technique and the matrix geometric technique [1]. It is important to note that the direction of exact solution depends on the model assumptions such as processing distribution type, boundary conditions and solution techniques precision. All of these assumptions are reliable only for the two machines problem and deviate from the real cases, so simulation techniques is chosen for applying in this work to measure the key performance of production lines virtually.

3. PROBLEM STATEMENT OF PRODUCTION LINE

The main problems in production lines treatment are the calculation of throughput and average levels of buffers because of the large size of the state space. Each machine can be in one or two states: operational or under repair. Buffer $B_i$ can be in the Ni+1 state, where $n_i = 0, 1, 2, \ldots, N_i$, and where $ni$ is the amount of material in $Bi$ and $Ni$ is its capacity. Consequently, the Markov Chain representation of a k-machine in the production line with $k-1$ buffers has a state space of cardinality

$$
2k!\prod_{i=0}^{k-1} (N_i + 1).
$$

(16)

As an example, a production line with 20 machines and 19 buffers with capacity 10 parts for each. Therefore, the number of states for this production line is over $6.41 \times 10^{25}$ states. This state space is too large to allow brute force calculation [12]. Designers of such production lines want to optimize either the production rate, or the profit of the line. However, material flow may be disrupted by machine failures. The inclusion of buffers increases the average production rate of the line by limiting the propagation of disruptions, but at the cost of additional capital investment, floor space of the line, and inventory [13]. It is clear from Table 1 that the number of states increases significantly with an increase in the size of the buffer and in the number of stations. This places strict limits on the size of the system for which exact results can be obtained.
Table 1. Number of states for only one phase of processing and one repair distribution of the ith station (P=1, R=1) and identical buffer capacities[4]

<table>
<thead>
<tr>
<th># of Stations</th>
<th>Buffer Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>112</td>
</tr>
<tr>
<td>5</td>
<td>418</td>
</tr>
<tr>
<td>6</td>
<td>1,550</td>
</tr>
<tr>
<td>7</td>
<td>5,922</td>
</tr>
<tr>
<td>8</td>
<td>21,728</td>
</tr>
<tr>
<td>9</td>
<td>81,090</td>
</tr>
</tbody>
</table>

4. COMPARISON OF ANALYTICAL MODEL BEHAVIOR WITH REAL WORLD BEHAVIOR

Hanifin [7] carried out a simulation of the effect of buffers using the actual data from the line. His simulated efficiency was significantly lower than that predicted using the exact formula for unequal stages (see Fig. 2). There was a difference in the predicted efficiency for the line without a buffer, which can be attributed to the existence of time dependent failures in the real world and in the simulation model. The difference between the analytical and simulation model’s predictions of increased efficiency due to buffers value is far more significant. This can be attributed to two main reasons. First, the real data does not fit the assumptions of the analytic model because of the exponential down time distributions. Secondly, there is serial correlation in the up times of the finish section.

Fig. (2) Comparison of efficiencies predicted by analytical and simulation models[2]

5. SIMULATION APPROACH

In this section, the proposed model is explained with its methodology. Assumptions will be stated clearly in the simulation policy then the model steps are performed. This methodology depends on a precise and longtime data collection, which leads to accurate results. Section 6 clarifies the ARENA simulation model. This model simulates the actual cement line.

6. APPLICATION CASE STUDY

This section presents the studying of production line performance keys, which applied on a real cement production line as a case study. Similar to the author paper [14], a supplementary simulation model is developed by ARENA software and used to analyze and test several bottlenecks that are causing severe congestions in different areas on the production line. Workstation failure data is collected along one year to all machines to obtain the machines failure behaviors. After a simulation time of 13 days, a simulation results show the line bottlenecks, workstations utilization, buffer capacities and the line production rate. To resolve the bottlenecks, a simulation model is rebuilt with 13 days simulation time and 15 replications. The outputs clarify resolving of allocation of buffers, which verify reliable size. These sizes might be taken into consideration when designers implement such lines. Finally modified workstations utilization, buffer capacities that lead to an increase of the line production rate by about more than 15% of the production rate and economizing of 37% of buffer capacities.

6.1 SIMULATION POLICY

The policy of the simulation package ARENA is:

- The entity which the simulation package operates on the capacity of the arrival truck which, is unloaded to the crasher both limestone and clay. The two entities are summed as a single entity before the two Raw Mills to complete the cycle.
- Steady state simulation models are appropriate for the analysis of systems, which in theory could run indefinitely so a 5 hours warm-up period is taken.
- It might be appropriate to consider the product as a discrete unit in particular the trucks come in a discrete truck, also the customer van come out the same discrete value.

6.2 METHODOLOGY

The production line should be studied in details which given by [4]:

- All workstations should be analyzed; processes, resources, material, and timings should be identified and documented.
- All data related to activities and resources should be identified and collected.
- A simulation model that truly represent the real production line and simulate its behavior, should be developed, and validated.
- Once, a valid model is built, a simulation experiment should be conducted to search for a feasible solution to maximize the capacity of the production line and
optimize the buffer allocation within the existing con-

straints.

### 6.3 The Collected Input Data

The probability distributions with their parameters of major activities are collected from the actual production line for a complete year. These data include the failure of each machine during this year, which is entered to ARENA Input Analyzer to produce the best distribution of failure. The failure data includes the predictive and preventive maintenance schedule. The probability distributions with their parameters are scheduled in Table 2. The appropriate failure distributions, which resulted from ARENA Input Analyzer, are scheduled in Table 3.

#### Table 2. Probability distributions with their parameters for each workstation

<table>
<thead>
<tr>
<th>Activity</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truck arrival</td>
<td>EXPO (7) min.</td>
</tr>
<tr>
<td>Crasher processing time</td>
<td>EXPO (6) min.</td>
</tr>
<tr>
<td>Stacker processing time</td>
<td>EXPO (7) min.</td>
</tr>
<tr>
<td>Reclaimer processing time</td>
<td>EXPO (8) min.</td>
</tr>
<tr>
<td>Raw Mill Capacity</td>
<td>9000 ton/day</td>
</tr>
<tr>
<td>Kiln Capacity</td>
<td>7000 ton/day</td>
</tr>
<tr>
<td>Cooler processing time</td>
<td>EXPO(6) min.</td>
</tr>
<tr>
<td>Cement Mill Capacity</td>
<td>8000 ton/day</td>
</tr>
<tr>
<td>Packing machine</td>
<td>EXPO(8) min.</td>
</tr>
<tr>
<td>Disposal truck arrival</td>
<td>EXPO(10) min.</td>
</tr>
</tbody>
</table>

#### Table 3. Failure time distributions according to ARENA Input Analyzer

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Failure time Distribution (hrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crasher</td>
<td>LOGN (1.04, 2.26)</td>
</tr>
<tr>
<td>Stacker</td>
<td>EXPO (7)</td>
</tr>
<tr>
<td>Reclaimer</td>
<td>EXPO (8)</td>
</tr>
<tr>
<td>Raw Mill 1</td>
<td>EXPO(7.4)</td>
</tr>
<tr>
<td>Raw Mill 2</td>
<td>EXPO(7.18)</td>
</tr>
<tr>
<td>Kiln</td>
<td>GAMM(15.8, 0.718)</td>
</tr>
<tr>
<td>Cooler</td>
<td>EXPO (10)</td>
</tr>
<tr>
<td>Cement Mill 1</td>
<td>LOGN(3.72, 8.12)</td>
</tr>
<tr>
<td>Cement Mill 2</td>
<td>LOGN(4.69, 10.1)</td>
</tr>
<tr>
<td>Cement Mill 3</td>
<td>LOGN(4.23, 9.1)</td>
</tr>
<tr>
<td>Packing machine</td>
<td>EXPO(8)</td>
</tr>
</tbody>
</table>

#### 6.4 Verification and Validation

The animation method is used to show the movement of entities inside the model and to insure that the movement is similar to what the designer think which called Face Validity [15]. Validation of the ARENA model is done by comparing the model output with the real system output which called statistical validation or walkthrough validation [15].

The number of cement trucks produced per day from the model is109 trucks, which is equivalent to 5450 tons while the real system production rate per day is 104, which are equal 5200 tons per day, which is considered valid. The nature of this production system is a steady state because it works continuously for 24 hours a day and 7 days a week, except the crashing, stacking and reclaiming workstations, which works only two shifts and take 18 hours daily.

#### 6.5 Performance Measures and the Results

This section clarifies the output results about the standing cement production line which include the intermediate buffer capacities and total production rate. The following performance measures of the line were determined:

- Throughput (jobs exiting from the production line per unit time).
- Utilization of each workstation (the limit of the time average of the number of busy machines over time divided by the total number of machines in the station).
- Average buffer level for each intermediate buffer.
- Average work-in-process, WIP, excluding the buffer before the first station.
- Average job-waiting time at each of the intermediate buffers.
The WIP for the line is 39,859.9 trucks
The line production rate = \frac{\text{Number of entities out}}{\text{Total simulation time}} = \frac{1311 \text{ (Table 4.)}}{13 \text{ days}} = 100.8 \text{ ton/day}

Figures 6 and 7 clarify the line performance keys indicators. It is clear that the maximum buffer size is located before the kiln and the cooler. The simulated kiln buffer capacity is 2.1571 trucks, which equalize 107.855 tons whereas the standing value is 500 ton, so a 78.429% could be saved.

Table 4. Number of entities out

To resolve the model bottlenecks, a simulation model is rebuilt with 13 days simulation time and 15 replications as in figure 4. The kiln processing time modified from EXPO (10) to EXPO (8) by increasing its capacity to reach 9000 ton/day without any change in the other equipment parameters merely increasing the third shift of Crashers, Stackers and Reclaimers to work all day like the other equipment of line because they work only two shifts in the standing line. If it is done, the daily production capacity will increase to 128 trucks per day, which is equivalent to 6400 tons instead of 109 trucks per day, which equivalent to 5450 tons per day with an increase of 950 tons per day. In addition, that represents about 15% extra production, which would lead to a profit, covers the kiln extension cost after one-month production.

6.6.1 Utilization percent

Figure 5 shows the utilization percent for each workstation after the improvement made in the model. It is clear that the utilization percent of the cement mills increased because of the more kiln capacity.

6.6.2 Buffer capacities

Table 5 and figures (6, 7, and 8) show the buffer capacities. It is clear that after resolving the buffer capacities are reduced.
Table 6. Buffer capacities resulted from the modified model

<table>
<thead>
<tr>
<th>Buffer size (ton)</th>
<th>Packing</th>
<th>Cement mill</th>
<th>Cooler</th>
<th>Kiln</th>
<th>Raw mill</th>
<th>Clay Reclaimer</th>
<th>Limestone Reclaimer</th>
<th>Clay Stacker</th>
<th>Limestone stacker</th>
<th>Clay Crusher</th>
<th>Limestone Crusher</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25</td>
<td>42</td>
<td>55</td>
<td>25</td>
<td>55</td>
<td>55</td>
<td>57</td>
<td>51</td>
<td>57</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7 CONCLUSIONS

The goal of this study was achieved by measuring the performance of a cement production line. The production line was thoroughly analyzed and found to have bottlenecks that were causing congestion in the kiln area on the line. Simulation was used to analyze this bottleneck and resolve it, so Simulation is the best tool that can be used in such a study because one can search for a good feasible solution without disrupting its operation. The production capacity could be increased by 15.4% if an extension is added to the kiln and it may need an extra cement mill. The line performance would be increased by improving the preventive maintenance schedule to increase the machines utilization, which leads to extra productivity increase.
REFERENCES


