International Journal of Mechanical, Industrial and Aerospace Sciences
ISSN: 2517-9950
Vol:3, No:5, 2009

Modeling and Simulation of a Serial Production Line with Constant Work-In-Process

Mehmet Savsar

Abstract**—** This paper presents a model for an unreliable production line, which is operated according to demand with constant work-in-process (CONWIP). A simulation model is developed based on the discrete model and several case problems are analyzed using the model. The model is utilized to optimize storage space capacities at intermediate stages and the number of kanbans at the last stage, which is used to trigger the production at the first stage. Furthermore, effects of several line parameters on production rate are analyzed using design of experiments.

Keywords**—**Production line simulator, Push-pull system, JIT system, Constant WIP, Machine failures.

I. Introduction

SIMULATION has been extensively used in modeling and analyzing production control systems. A particular type of production control, which has become a common trend in industry, is just in time (JIT) or "pull" system of production control. In a JIT system, production is initiated according to demand for finished products at each stage to produce what is needed at the right time and in the right quantity. Alternative to a purely pull system is the hybrid push-pull system, where the production at the first stage is scheduled according to the demand for the products in the last stage. Withdrawal of finished products from the last stage triggers the production at the first stage by an information signaling card, called a kanban. Intermediate operations are performed by a push system. Push-pull systems are commonly used in electronics assembly operations. Several studies have been carried out on the implementation and efficiency of JIT systems.

[1]-[6] have analyzed JIT systems from different perspectives using simulation as well as other meta modeling approaches, including neural network models. [7]-[10] studied a hybrid push-pull system and presented a control algorithm for multi-stage, multi-line production systems. [11] compared three pull control policies, namely the kanban, base stock, and generalized kanban.

The effects of kanbans and other factors on JIT system performance have been investigated mostly for pull types of production control strategy. Countless number of other JIT applications and related models can be seen in the literature. Most of the literature deals with the efficiency of JIT systems under different operational conditions. Either mathematical models are developed based on restrictive assumptions or simulation models are utilized in the analysis of JIT systems. In relation to the effects of intermediate buffer capacities on a push type of serial production line and optimum allocations of buffers on the line, several papers have been published. In particular, papers related to buffer allocations include [12]-[22].

In this paper, we developed a discrete mathematical model to analyze a push-pull system of production with constant work-in-process (CONWIP). When a final product is withdrawn from the finished products inventory in the last stage, a kanban is signaled to the first stage to start the production. Fig. 1 illustrates operation of such a system.

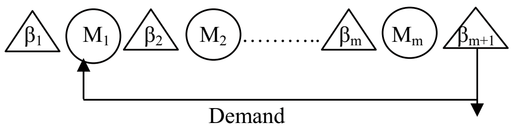

Fig. 1 A Push-pull production control system

Successive operations are carried out by completion of each product at each station (Mi) and its delivery to the succeeding station or its buffer store (βi), if the station is busy. It is assumed that the intermediate buffer sizes, which represent maximum work-in process at each stage, are limited in capacity. Thus, when the storage of finished units in the final products inventory reaches a specified maximum capacity, the last station stops its production. Similarly, when an intermediate buffer βi is filled up to its maximum capacity, the preceding station, Mi-1, stops its production or the completed part stays on the station until a part is removed from the succeeding buffer. The capacity of buffer i is zi.

In production systems, which operate according to demand, equipment availability is important since machine failures significantly delivery time of products. While equipment failures due to wear-outs can be eliminated, failures due to random causes could not be eliminated. When optimizing

Mehmet Savsar is professor and chairman of the Department of Industrial & Management Systems Engineering at Kuwait University, College of Engineering & Petroleum, , P.O. Box 5969, Safat 13060, Kuwait (phone:+965 24987873; fax: +965 24816137; e-mail: mehmet@kuniv.edu.kw or msavsar@yahoo.com).

numbers of kanbans and the sizes of in-process buffers, it is necessary to consider equipment reliability in model development. In the following section, we present a discrete mathematical model, which is based on the flow of discrete parts or batches from stage to stage. The model is used to analyze the behavior of the push-pull system under various operational conditions including equipment availabilities and randomness in demand.

II. Mathematical Model for the Push-Pull System

The basic principal of the discrete model is to determine the total time a batch of parts n spends in station i, the time instant at which batch n is completed in station i, and the time instant at which batch n leaves the station i. Storages $\beta_2,\ldots\ldots,\beta_m$ are called intermediate buffer storages, having finite capacity z_i, i=2,…,m. Initial input storage is assumed to have unlimited capacity for the raw material, while the final output storage has limited capacity for completed batches of products with attached kanbans. The final buffer (m+1) is assumed to be the finished products storage with time between part departures being equal to time between demand. The following notations are used in the formulation:

π_{in}= Time duration that n^{th} batch stays on the i^{th} station not considering imposed stoppages due to equipment failures; i=1,2,…..,m.

m = Number of stations on the line.

τ_{in} = Processing time of batch n on station i (this may be a random variable with certain distribution)

μ_{in} = Repair time of the i^{th} station required for correction of a failure during processing of the n^{th} batch. Time to failures and the repair times are assumed to follow certain distributions, which are generated and incorporated into the model when the model is solved iteratively.

ψ_{in} = Instant of time at which processing of the n^{th} batch is completed on the i^{th} station.

Δ_{in} = Instant of time at which n^{th} batch departs from the i^{th} station.

Δ_{0n} = Instant of time at which n^{th} batch enters the first station.

η_{in} = Instant of time at which i^{th} station is ready to process the n^{th} batch.

$\Omega_{m+1,n}$= Instant of time at which n^{th} batch departs from the final buffer m+1.

δ_n = Mean time between demand for batches n-1 and n from the final buffer. This time may also be a random variable with certain distribution.

A part stays in a station for three reasons: (i) The part is being machined; (ii) The machine has failed during machining of the part and a repair is taking place; (iii) The successive buffer is full and the part can not be transferred to the next station due to an imposed stoppage. The residence time of the nth part on the ith station, π_{in}, without considering imposed stoppages is given as follows:

$$\pi_{in} = \tau_{in} + \mu_{in} \quad (1)$$

The discrete mathematical model of the push-pull system consists of calculating part completion times, ψ_{in}, and part departure times, Δ_{in}, in an iterative fashion. The following formulation is developed for ψ_{in} and Δ_{in} to be used in iterative calculations.

Processing of batch n cannot be started on station i until the previous batch, n-1 leaves station i. Therefore the time instant at which i^{th} station is ready to begin the n^{th} batch, denoted by η_{in}, is given by $\eta_{in} = \Delta_{i,n-1}$. If, $\Delta_{i-1,n}<\eta_{in}$, then the n^{th} batch must wait in buffer β_i, since it has left station i-1 before station i is ready to accept it. Therefore, processing of the nth batch in the i^{th} station will start at the instant η_{in}. If however, $\Delta_{i-1,n} \geq \eta_{in}$, then processing of the n^{th} batch in the i^{th} station can start immediately at the time instant $\Delta_{i-1,n}$. Considering both cases above, one gets the relation for the ready time of the nth batch to be processed in the ith station as follows:

$$\eta_{in} = max[\Delta_{i-1,n}, \Delta_{i,n-1}] \quad (2)$$

Since the n^{th} batch will stay in station i for a period of π_{in} time units, its processing will be completed by the time instant ψ_{in} given by:

$$\psi_{in}= max[\Delta_{i-1,n}, \Delta_{i,n-1}] + \pi_{in}= \eta_{in}+\pi_{in} \quad (3)$$

where i=2,3,……,m.

In case of the first station, a kanban must arrive before the batch can be processed. The arrival of a kanban from storage m+1 is modeled as follows:

Let ρ=n-L_2 where, L_2=Total number of batches initially in stations S_2 and storages $\beta_2,\ldots.., \beta_m$. Then,

$$\psi_{1n}= max[\Delta_{1,n-1}, \Omega_{m+1,\rho}] + \pi_{1n} \quad (4)$$

Time instant at which nth batch is ready to enter the first station is assumed to be $\Delta_{0n}<\Delta_{1,n-1}$ since we assumed that there are always batches of parts available in front of the first station. However, a kanban must arrive from the buffer (m+1) to start the process at station 1. Now, it remains to determine the time instant at which nth batch departs from the i^{th} station, Δ_{in}. It is found by considering two cases.

Let $k = n - z_{i+1} - 1$ (5)

In the first case, $\psi_{i,n} < \Delta_{i+1,k}$ (6)

which indicates that the n^{th} batch has been completed on the i^{th} station before processing of the $(n-z_i+1)^{th}$ batch has started on the $(i+1)^{th}$ station. Since buffer i+1, which is between station i and i+1 with capacity z_i+1, is full and station i has completed the n^{th} batch, the n^{th} batch may leave the i^{th} station only at the instant of time at which the $(n-z_i+1)^{th}$ batch of the $(i+1)^{th}$ station has started processing.

Therefore, $\Delta_{i,n} = \Delta_{i+1,k}$ (7)

In the second case, $\psi_{i,n} > \Delta_{i+1,k}$ (8)

which indicates that, at the instant ψ_{in} there are free spaces in buffer β_i+1 and therefore part n can leave machine i immediately after it is completed; that is, $\Delta_{in}= \psi_{in}$ holds under this case.

Considering both cases above, we have the following relations for Δ_{in}:

$$\Delta_{i,n} = \psi_{i,n} \text{ if } n \leq z_{i+1} + 1 \quad (9)$$

$$\Delta_{i,n} = \max[\psi_{i,n}, \Delta_{i+1,k}] \quad (10)$$

if $n > z_{i+1}+1$; i =1, 2, 3,……, m-1, where k = n–z_{i+1}–1 and,

$$\Delta_{m,n} = \max\lfloor \psi_{m,n}, \Omega_{m+1,k} \rfloor \quad (11)$$

Where, $\Omega_{m+1,k}$= Departure time if the k[th] batch from the final buffer m+1. Departure time of a finished product from the final buffer (m+1) depends on two conditions and calculated as follows:

In the first case, $\Delta_{m,n} \leq \Omega_{m+1,n-1}+\delta_n$. In this case,

$$\Omega_{m+1,n}= \Omega_{m+1,n-1}+ \delta_n. \quad (12)$$

In the second case, $\Delta_{m,n} > \Omega_{m+1,n-1}+\delta_n$. In this case,

$$\Omega_{m+1,n}= \Delta_{m,n}. \quad (13)$$

Combining both cases above, the following general relation is obtained for $\Omega_{m+1,n}$:

$$\Omega_{m+1,n}= max[\Delta_{m,n}, \Omega_{m+1,n-1}+ \delta_n] \quad (14)$$

In real-world situations, intermediate buffers may contain batches of parts that are often left from a previous shift or day. Therefore, it is important to start the iterations with some initial conditions as follows:

Let: l_i=Number of batches initially in buffer β_i and station S_i and, $L_k = \sum_{i=k}^{m+1} l_i =$ Total number of batches on stations $S_k,\ldots,S_m$ and storages $\beta_k,.., \beta_{m+1}$.

When iterations are started, one can assume that, at the initial time instant t=0, parts 1,2,…, L_{i+1} are already processed on station S_i, since these batches are initially in the line right after station S_i. Therefore, given the initial values of $l_1,\ldots,l_{m+1}$, the initial values of ψ_{in} and Δ_{in} for n=1,2, …,L_{i+1} are expressed by:

$$\Psi_{in} = \Delta_{in} = 0; \quad i = 1, 2, \ldots, m. \quad (15)$$

In order to carry out iterative computations, a simulation procedure is developed and implemented on the computer to determine several production line performance measures, which include average number of batches completed by the line during a specified period, average number of batches completed by each station during the same time, percentage of time for which each station is up and down due to imposed or inherent stoppages.

III. Computations of the Model

The discrete model is coded into a simulation program and implemented on the computer to calculate system performance measures. In addition to the variables described for the discrete model, the simulation allows several distributions, including: exponential, uniform, Weibull, normal, log normal, Erlang, gamma, beta and constant values to be specified for failure and repair times of the equipment in each station. Iterative simulation model basically calculates the time instant at which each part enters a station, duration of its stay, and the time it leaves the station. This is continued until, for example one shift, which is the specified simulation time T_{sim}, is completed. In order to obtain reliable results, several simulation runs have to be obtained and the average performance measures should be calculated. The results of each iterative simulation are utilized with statistical tests to determine if the specified conditions are met to stop the number of simulation iterations (i.e., shifts). If the conditions are not met, simulation iterations are continued with further runs. For each simulation realization, calculations of π_{in}, ψ_{in}, Δ_{in}, and $\Omega_{m+1,n}$ are performed iteratively with the consideration given to equipment failures and repairs as the parts flow through the system.

Reliable results cannot always be obtained from a single simulation realization. Therefore, additional runs have to be performed and the results tested statistically until the error in the line production rate is less than an ε value with a probability, both of which are predefined. This is accomplished by comparing the average production output rate from simulation ($\overline{Q}$) to the expected value ($\overline{\overline{Q}}$) using the confidence interval calculation given below. Here,

$$\overline{Q}=\frac{\overline{N}}{T_{sim}} \text{ and } \overline{N} = \sum_{i=1}^{n} (N_i)/n \quad (16)$$

where N_i=production output obtained from simulation run i and n is the number of simulation runs.

$$\Pr\left[1-\frac{Z_{\alpha/2}\sqrt{V(\overline{Q})}}{\overline{Q}} < \frac{\overline{\overline{Q}}}{\overline{Q}} < 1+\frac{Z_{\alpha/2}\sqrt{V(\overline{Q})}}{\overline{Q}}\right] = 1-\alpha \quad (17)$$

The aim is to have an estimated output rate, $\overline{Q}$, as close to the actual mean output rate $\overline{\overline{Q}}$ as possible. To achieve this, $Z_{\alpha/2}\sqrt{V(\overline{Q})}/\overline{Q}$ is minimized by obtaining more runs. As this value gets closer to 0, $\overline{Q} \to \overline{\overline{Q}}$ with probability 1 – α. An ε value is entered by the user; the simulation program calculates $Z_{\alpha/2}\sqrt{V(\overline{Q})}/\overline{Q}$ after each iteration; compares this quantity with ε and terminates the program if it is less than ε. If it is not less than ε after a maximum number of runs, the program is still terminated to avoid excessive computation. The iterative simulation model allows one to determine various parameters and dependent variables with significant effects on productivity and other performance measures. Estimation indices are obtained for such variables as the total, inherent, and imposed time losses due to failures and stoppages for each station as follows:

$Q_{ni} = 60/\tau_i$ is the nominal productivity of station i, where τ_i is the cycle time for station i; $Q_{ri} = 60\overline{N_i}/T_{sim}$ is the relative productivity of station i; $K_{loss}(i) = 1 - Q_{ri}/Q_{ni}$ is the total loss factor of station i; $K_{inh}(i) = 1 - \bar{t}_{ri}/(\bar{t}_{ri} + \bar{t}_{fi})$ is inherent loss factor of station i; and $K_{imp}(i) = K_{loss}(i) - K_{inh}(i)$ is the imposed loss factor for station i, i = 1, 2, ….., m. The terms $\bar{t}_{fi}$ and $\bar{t}_{ri}$ are mean times to failure and to repairs, of station i respectively. After determining these loss factors, they are compared for all stations. The station with the highest total loss factor is then chosen for improvement. If $K_{imp}(i)>K_{inh}(i)$, stoppages are mainly due to blocking and starvation; therefore it is necessary to increase the capacity of buffers immediately preceding and succeeding it. If $K_{inh}(i)>K_{imp}(i)$, stoppages are mainly caused by inherent failures, that is breakdowns; therefore, the reliability of station i, should be increased or its mean repair time should be decreased in order to gain improvement in total line productivity. After the suggested changes are made, iterative simulation is repeated to see the effects of the proposed changes in the design.

IV. Storage Space Allocation

Optimum storage space or buffer allocation problem has been studied by several researchers with respect to allocation of a fixed amount of total buffer space on a serial production line. The problem can be stated as follows: Given a total amount of acceptable buffer space of Z units, allocate this total space to individual buffers $S_2,\ldots,S_{m+1}$, the quantities $z_2, z_3,\ldots,z_{m+1}$ respectively such that the total production output rate of the line, Q(z), is maximized. The problem is stated as follows:

Choose $z_2, z_3,\ldots,z_{m+1}$ so as to

Maximize Q(z)

$$\text{Subject to: } \sum_{i=2}^{m+1} z_i \le Z \tag{18}$$

$z_i \geq 0$ and integer (i=2,3,…..m+1)

This problem has been discussed by [7] for production lines with exponential processing times in all stations. The optimization model is a linearly constrained integer nonlinear programming problem that is difficult to solve due to the fact that Q(z) has to be evaluated by either continuous time Markov chains or by some other stochastic processes approximation. [7] evaluated Q(z) for the serial line using Markov chains approach and indicated that the number of states are too large and exceeds well over 20,000 equations to be solved to obtain the value of Q(z) for a given buffer size combination. Even if it was practical to solve the problem, it would not be still applicable to the cases with equipment failures and non-exponential process times. The buffer allocation model is applied to the push-pull system in this paper. However, we obtain the solution for Q(z) using the iterative solution procedure presented above. This procedure is not restricted to exponential process times and all reliable equipment, since it is based on simulation. A fixed number of buffers are specified and the iterative computations are performed to determine optimum combinations by evaluating all buffer combinations. The optimum corresponds to the maximum production output rate. For small size problems, such as lines with up to 8 stations and up to 10 buffer capacities, computational time is in the order of minutes, depending on the accuracy required. However, for larger problems, such as more than m=10 stations and more than Z=15 buffer capacities, computational time is relatively large since number of possibilities evaluated is large. If a small accuracy with 5% error is acceptable, large problems can also be solved in a reasonable time.

V. Simulated Case Problems

The model is illustrated by several case problems. Table I and Table II are the input data and the output results obtained for a 5-station line with all stations available 85% of the time. Processing times, failure distributions, their parameters, repair distributions and their parameters are shown in Table I.

TABLE I
Input Data for Production Line Simulation

Station	Process Time	No. of Failures	Failure Distrib.	Repair Distrib.
1	1.0	1	Expo(85)	Normal (15, 2.25)
2	1.0	1	Expo(85)	Normal (15, 2.25)
3	1.0	1	Expo(85)	Normal (15, 2.25)
4	1.0	1	Expo(85)	Normal (15, 2.25)
5	1.0	1	Expo(85)	Normal (15, 2.25)

The outputs for 2000 time units of simulation with α=0.05, ε=0.005, and maximum iterations=100 are shown in Table II. The results include relative production rate of each station and various loss factors due to equipment failures as discussed in section III. The output also includes suggestions for line improvement.

TABLE II
Units

Average Line Output (Parts/Time Unit.)=0.564975
Standard Dev. of Line Output Rate=0.00281367
Optimum Buffer Allocations for Buffers m=2 to 6
are as Follows: 0 2 1 0 0

Station	Relative Prod. Rate	Imposed Loss Factors	Inherent Loss Factor
1	0.564	0.150	0.285
2	0.565	0.151	0.285
3	0.565	0.148	0.287
4	0.566	0.144	0.290
5	0.566	0.143	0.291

Suggestions: Station No. 1 has the Maximum Total Loss Factor. Down Time is Mainly Imposed. Increase the Capacity of Storage Adjacent to This Station. Also Increase Reliability and Productivity of Adjacent Stations and Try Simulation Again.
Error<Epsilon Is Reached at Iteration = 100
Maximum Iteration Is Reached At Iteration = 100
Total Computation Time = 104.39 Seconds

In the second case problem, a push-pull production line with 5 serial stations is considered as before. All stations are assumed to be reliable, except one station which was placed at the beginning (B), in the middle (M), or at the end (E) of the line to see the effects of unreliable station at different segments of the line. Failure and repairs for the unreliable station are as in Table I. Standard deviation of the repair times was taken as 15% of the mean. Since availability is A=MTBF/(MTBF+MTTR), selected parameters represented 85% equipment availability for the particular station, whose effect on the line was investigated. In other words, we wanted to see how the buffers would be allocated if the unreliable station was at the beginning, at the middle or at the end of the line. The system was analyzed by the simulation over a period of 2000 time units. All intermediate buffer combinations, which add up to less than or equal to total buffer capacity Z (Z=0,1,23,5,10) are evaluated in order to determine line productivity, as shown in Fig. 2, and optimum buffer combinations (z_2, z_3, z_4, z_5), which resulted in maximum production rate, as shown in Table III, for three different locations of the unreliable equipment at the beginning (B), in the middle (M), and at the end (E) of the line.

The corresponding production output rates are given as the percentage of nominal rate, which would be 100 if the line was all reliable, balanced with constant processing times. An important observation related to buffer size allocation can be seen in Table III. If the line has an unreliable station at the start or at the end, the buffer capacity available should be located immediately after the last station, except in the case of 5 and 10 buffer sizes, in which case one unit is allocated to buffer 4, after station 3 in the center of the line. A similar observation is seen in the case if the center station is unreliable. In this case if there is one buffer space to be allocated on the line, it is preferred to be at buffer 3, immediately after station 2.

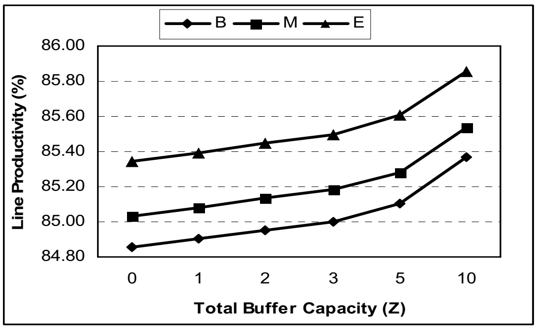

Fig. 2 Line productivity as a function of Z

Additional buffer sizes are mostly allocated to the end of the line after station 5, except in the case of 2 and 10 buffer sizes, in which case one space is allocated after the center station. The main reason that the buffer spaces are mostly allocated to the end of the line could be due to the fact that the line is operated as a push-pull system and therefore the first station can not start processing a part unless a part or batch is withdrawn from the last station. Buffer spaces after the last station helps increasing part availability during demand. The

TABLE III
BUFFER CAPACITY DISTRIBUTION TO STATIONS

	Beginning	Middle	End
Z	z_2 z_3 z_4 z_5 z_6	z_2 z_3 z_4 z_5 z_6	z_2 z_3 z_4 z_5 z_6
0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
1	0 0 0 0 1	0 1 0 0 0	0 0 0 0 1
2	0 0 0 0 2	0 1 0 1 0	0 0 0 0 2
3	0 0 0 0 3	0 1 0 0 2	0 0 0 0 3
5	0 0 0 0 5	0 1 0 0 4	0 0 1 0 4
10	0 0 0 0 10	0 1 0 1 8	0 0 1 0 9

same three cases were evaluated for a purely push type of production line, where the last station does not have a limit on outputting its product and the first station can start without waiting for a part withdrawal from the last station. The results, which are not shown here, are almost opposite of what is obtained for the push-pull system and the buffer allocation is preferred to be immediately after the first station near the start of the line in all cases.

VI. EFFECTS OF LINE CONFIGURATIONS, MAINTENANCE POLICIES AND LINE PARAMETERS ON LINE PERFORMANCE

In order to see effects of various production related parameters and factors on line performance measure, such as the production rate, several experiments were set up and results were obtained. In particular, the following production line factors were taken into consideration:

1. Production line length (3, 5 ,and 9 stations);
2. Buffer capacities between stations (0, 2, 4, 6, 8);
3. Process time variability measured by its coefficient of variation (CV_{pt}=0, 0.2, 0.5, 0.7);
4. Demand interval variability (CV_{dm}=0, 0.2, 0.5, 0.7);
5. Type of maintenance applied (Design out maintenance resulting in full reliability [REL], reliability centered maintenance [CM-PM], corrective maintenance [CM])

Process time at each station was assumed to be normally distributed with mean of 3.0 time units and varied according to the coefficient of variation (CVpt) selected. Similarly, time interval between the demands for withdrawal of products from the finished products storage was assumed to be normally distributed with mean of 3.0 time units and also varied according to the coefficient of variation (CV_{dm}) selected. The production lines are simulated over 2400 time units. 10 runs are carried out for each combination and average values are recorded. Figures 3-5 illustrates the production output rate as a function of various line configuration and factors mentioned above. CM-i, CM-PM-i, and REL-i represent two levels of maintenance and full reliability case for each station i.

In order to compare effects of corrective maintenance (CM) only to the CM with preventive maintenance (PM), reliability centered maintenance (RMC) concept was incorporated into the model. Under RMC, equipment is subjected to PM just before a failure is expected. Mean time between failures (MTBF) must be determined in advance. In this case, it is assumed that failures due to wear outs are eliminated and only random failures remain. This idea can be implemented analytically if time between failures are uniformly distributed. This concept has been explained in detail by Savsar[23].

Following is a mathematical procedure to separate random failures from the wear-out failures. This separation is needed in order to be able to see the effects of maintenance on the productivity and availability of a line when simulating the system.

Let f(t) = Probability distribution function (pdf) of time between failures.

F(t) = Cumulative probability distribution function (cdf) of time between failures.

R(t) = Reliability function (Probability that the equipment survives by time t).

h(t) = Hazard rate (or instantaneous failure rate).

Hazard rate $h(t)$ consists of two components, the first due to random failures and the second due to wear-out failures as:

$$h(t) = h_1(t) + h_2(t) \quad (19)$$

$h_1(t)$ = Hazard rate due to random failures.

$h_2(t)$ = Hazard rate due to wear-out failures.

Since the equipment failures are either due to chance causes or wear-outs, reliability of the equipment, which is the probability that equipment survives by time t, can be expressed as follows:

$$R(t) = R_1(t)\, R_2(t) \quad (20)$$

where, $R_1(t)$ = Reliability due to chance causes (or random failures) and $R_2(t)$ = Reliability due to wear-outs.

Since the hazard rate due to random failures is independent of time and therefore constant, we let $h_1(t)=\lambda$. Thus, the reliability of the equipment due to random failures with constant hazard rate would be as follows:

$$R_1(t) = e^{-\lambda t} \quad (21)$$

$$h(t) = \lambda + h_2(t) \quad (22)$$

It is known that

$$h(t) = f(t)/R(t) = f(t)/[1-F(t)] = \lambda + h_2(t) \quad (23)$$

$$h_2(t) = h(t) - h_1(t) = f(t)/[1-F(t)] - \lambda \quad (24)$$

$$f_2(t) = h_2(t)R_2(t) = [\frac{f(t)}{1-F(t)} - \lambda][\frac{1-F(t)}{e^{-\lambda t}}] = \frac{f(t)}{e^{-\lambda t}} - \frac{\lambda}{e^{-\lambda t}}[1-F(t)]$$

$$F_2(t) = 1 - R_2(t) = 1 - \frac{1-F(t)}{e^{-\lambda t}} = \frac{e^{-\lambda t} - R(t)}{e^{-\lambda t}}$$

$$f_2(t) = \frac{dF_2(t)}{dt} \quad (25)$$

$$R_2(t) = R(t)/R_1(t) = [1-F(t)]/\, e^{-\lambda t} \quad (26)$$

$$h_2(t) = f_2(t)/R_2(t) \quad (27)$$

These derivations show that, total time between failures, f(t) can be separated into two distributions, time between failures due to random causes [$f_1(t)$] and time between failures due to wear-outs [$f_2(t)$]. Since the failures due to random causes could not be eliminated, we must concentrate on the failures due to wear-outs in order to eliminate them by appropriate maintenance policies. By the procedure described above, it is possible to separate the two types of failures and develop the best maintenance policy to eliminate the wear-out failures. This separation is analytically possible for uniform distribution. However, it is not possible analytically for other distributions. It is assumed that when a preventive maintenance policy is implemented, failures due to wear-outs are eliminated and only failures due to random causes remain. These random failures are assumed to follow exponential distribution with constant hazard rate since they are completely random with unknown causes and the memoryless property of exponential is applicable.

For uniformly distributed time between failures, t, in the interval $0<t<\mu$, probability distribution function of time between failures without introduction of PM is given by:

$$f(t) = 1/\mu\,. \quad (28)$$

If we let $\alpha=1/\mu$, then, reliability is given as $1-\alpha t$ and the total failure rate is given as:

$$h(t)=f(t)/R(t)=\alpha/(1-\alpha t). \quad (29)$$

Let us assume that hazard rate due to random failures is a constant given by $h_1(t)=\alpha$, then the hazard rate due to wear-out failures could be determined by:

$$h_2(t)=h(t)-h_1(t)=\alpha/(1-\alpha t)-\alpha=\alpha^2 t/(1-\alpha t) \quad (30)$$

The corresponding time to failure probability density functions for each type of failure rate is:

$$f_1(t) = \alpha \times e^{(-\alpha t)} \qquad 0 < t < \mu \quad (31)$$

$$f_2(t) = \alpha^2 \times t \times e^{(\alpha t)}\,, \quad 0 < t < \mu \quad (32)$$

The reliability function for each component would be is as follows:

$$R_1(t) = e^{(-\alpha t)} \qquad 0 < t < \mu \quad (33)$$

$$R_2(t) = (1-\alpha t)\times e^{\alpha t}, \quad 0 < t < \mu \quad (34)$$

$$R(t) = R_1(t)\times R_2(t) \quad (35)$$

When the preventive maintenance (PM) is introduced, failures due to wearouts are eliminated and thus the machinery fails only due to random causes, which are exponentially distributed as given by $f_1(t)$. Sampling for the time to failures in simulations is thus based on exponential distribution with mean μ and *a* constant failure rate of $\alpha=1/\mu$. In case of CM without PM, in addition to the random failures, wear-out failures are also present and thus the time between equipment failures is uniformly distributed between 0 and μ as given by f(t). The justification behind this assumption is that uniform distribution implies an increasing failure rate with two components, namely, failure rate due to random failures and failure rate due to wearout failures as given by $h_1(t)$ and $h_2(t)$ respectively. Initially when t = 0, failures are due to random effect with a constant rate $\alpha=1/\mu$. As the equipment operates, wearout failures come into play and thus the total failure rate h(t) increases with time t. Sampling for the time between failures in simulation is based on a uniform distribution with mean $\mu/2$ and an increasing rate, h(t).

In the simulation experiments considered, time to failure is assumed uniformly distributed between 0 and 200 time units with a mean of 100 time units for all stations for the case of CM only. In the case of PM, wearouts are eliminated and time to failure extends; it becomes exponentially distributed with a mean of 200 time units. Time to repair was assumed normally distributed with mean of 15 time units and standard deviation of 3 time units.

Fig. 3 illustrates simulation results for the case when CV=0.0 for process time (CVpt) and the demand (CV_{dm}). As it is seen in fig. 3, 800 units (2400/3.0) are produced on any length of line if the line is fully reliable and there is no other source of variability. However, if the line is under failures with CM only, production rate is significantly reduced when line length is increased. When PM is introduced in addition to CM, production rate is between the CM and REL cases. It can be seen from figure 3 that between the cases of unreliable lines, the lowest production rate is for a 9-station line with CM only, while the highest rate is for a 3-station line with CM and PM together.

Fig. 4 shows the results for CVpt=0.0 and CV_{dm}=2.1; fig. 5 shows the results for CVpt=2.1 and CV_{dm}=0.0; fig. 6 shows the results for CVpt=2.1 and CV_{dm}=2.1. As it can be seen from figs. 3-6, as the process time and demand variability increase, production rate decreases. It is also clear from figs. 5 and 6 that as the process time becomes variable, the production rate can no longer reach to the maximum level of 800 units even for the reliable line. In all cases however, as the line length increases and the buffer capacities decrease, production rate decrease. Also, in all cases CM only results in lower production rate.

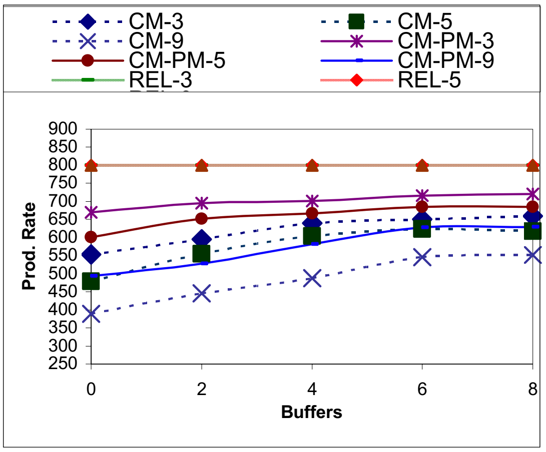

Fig. 3 Line production rate under different factors (CV_{pt}=CV_{dm}=0.0)

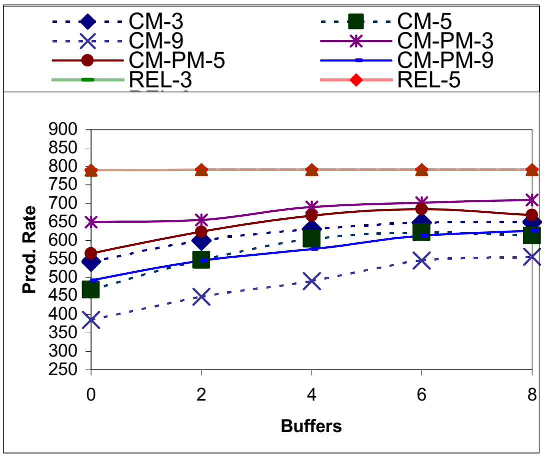

Fig. 4 Line production rate under different factors (CV_{pt}=0.0; CV_{dm}=2.1)

VII. EXPERIMENTAL DESIGN

In order to see significance of the effects of significant factors on line production rate, a general factorial design was set up with five factors each at three levels. Thus, line lengths of 3, 5, and 7; buffer capacities of 0, 2, and 6; process time CV of 0, 0.2, and 0.7; demand CV of 0, 0.2, and 0.7; and maintenance policies of REL, CM, and CM-PM cases were

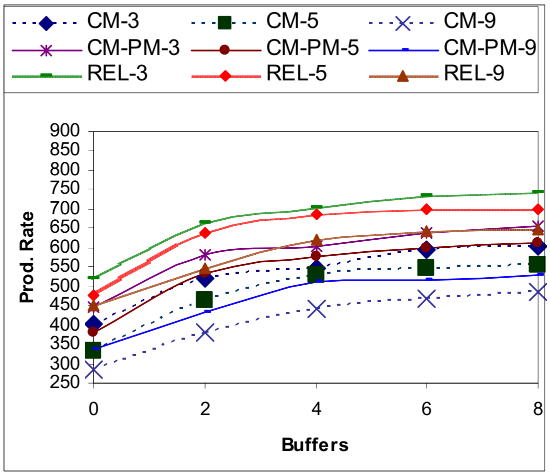

Fig. 5 Line production rate under different factors (CV_{pt}=2.1; CV_{dm}=0.0)

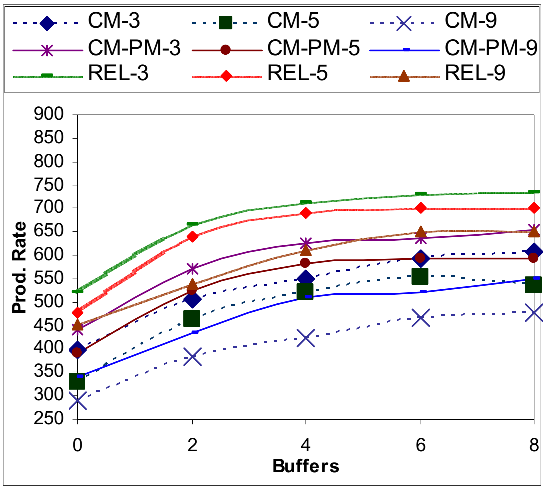

Fig. 6 Line production rate under different factors (CV_{pt}=2.1; CV_{dm}=2.1)

considered. The ANOVA results shown in Table IV indicate that following factors are significant: A: Line length: B: Buffers; D: Demand CV: E: Maintenance policy; and three interactions AE, BD, and DE. 97.56% of the variation is explained by these significant terms. It is interesting that process time variation was not a significant factor for this model. In the general factorial design model, the factors are considered as qualitative and therefore the model is hierarchical. The production rate is given as function of significant factors and their interactions.

Fig. 7, the normal probability plot for the residuals shows that the normality assumption is valid.

TABLE IV
ANOVA FOR SELECTED FACTORIAL MODEL RESPONSE: PRODUCTION RATE

Source	Sum of Squares	DF	Mean Square	F Value	Prob. > F
Model	4.3E+6	20	2.1E+5	443.5	< 0.0001significant
A	4.2E+5	2	2.1E+5	440.7	< 0.0001
B	7.9E+5	2	3.95E+5	822.2	< 0.0001
D	8.4E+5	2	4.2E+5	870.26	< 0.0001
E	1.94E+6	2	9.7E+5	2020.21	< 0.0001
AE	69418.9	4	17354.7	36.09	< 0.0001
BD	1.13E+5	4	28248.6	58.75	< 0.0001
DE	88468.0	4	22117.0	46.00	< 0.0001
Residual	1.07E+5	222	480.83		
Corrected Tot.	4.4E+6	Total DF: 242			

The Model F-value of 443.5 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise. Values of "Prob > F" less than 0.05 indicate model terms are significant. In this case A, B, D, E, AE, BD, DE are significant model terms. Values greater than 0.1 indicate the model terms are not significant.

Other ANOVA related statistics are as follows: Std. Dev.=21.93; R-Squared=0.9756; Mean=591.67; Adj R-Squared=0.9734; C.V.=3.71; Pred R-Squared=0.9707; PRESS=1.279E+5; Adeq Precision=91.49.

The "Pred R-Squared" of 0.9707 is in reasonable agreement with the "Adj R-Squared" of 0.9734. "Adeq Precision" measures the signal to noise ratio. A ratio greater than 4 is desirable. The ratio of 91.49 indicates an adequate signal. Final equation, which relates the production rate to the coded values of the significant factors, is given as follows:

Production Rate=591.67+47.08*A[1]+7.34*A[2]-73.22*B[1] +7.29*B[2]+59.28*D[1]+20.67*D[2]-95.47*E[1]-24.07*E[2] +14.88*A[1]E[1]+0.56*A[2]E[1]+12.91*A[1]E[2]+2.19* A[2]E[2]+37.60*B[1]D[1]-11.48*B[2]D[1]-3.48*B[1]D[2]+ 3.98 * B[2]D[2]-20.44*D[1]E[1]-6.75*D[2]E[1]-5.98* D[1]E[2]-1.61*D[2]E[2]

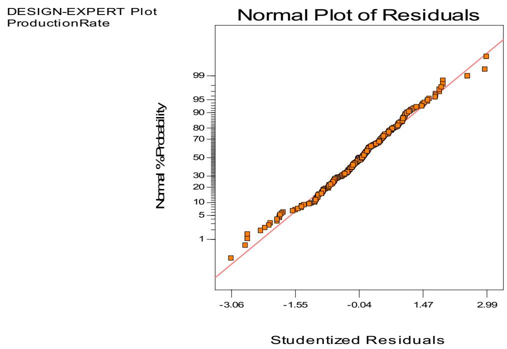

Fig. 7 Normal probability plot of residuals

VII. CONCLUDING REMARKS

This paper has presented an iterative mathematical model and a computer simulation procedure for a multi-stage production flow line operated according to demand at the last station, while using a push system at the intermediate stations. Based on the discrete mathematical model, simulation process incorporates a three-stage procedure which allows the user to enter a set of data describing the system under study, simulate the system iteratively until selected statistical criteria are satisfied, obtain the output, and apply specific recommendations for productivity improvement until satisfied production output is achieved. The simulation model is very useful in estimating production line productivity for realistic systems. It allows the line designer or managers to evaluate effects of storage capacity and repair/maintenance policies on productivity of a system.

The model was utilized to see the optimum allocations of storage unit capacities along the line if the equipment were subject to random failures. If all the equipment had similar failure rates, it was observed that the optimum allocation of buffer storages followed a bowl shape, meaning that more buffer spaces were allocated to the center stations. If only one station was subject to failures, most of the buffers were allocated to the final storage to achieve maximum production output irrespective to the location of the unreliable station being either at the start, at the middle, or at the end of the line. As a future study, the suggested iterative model can be incorporated into interactive computer software to be effectively utilized by engineers and managers.

Simulation model was utilized to investigate the effects of line configurations, maintenance policies, buffer capacities, process time variability, and demand variability on production rate of the line. A factorial design was set up to investigate the significant factors that affect the production rate. It was found that line length, buffer capacities, maintenance policies, and demand variability had significant effects on production rate.

REFERENCES

[1] Chu, C. and Shih, W. "Simulation Studies in JIT Production", International Journal of Production Research, 30 (11), 1992, pp. 2573-2586.
[2] Fukukawa, T. and Hong S.C., "The Determination of Optimal Number of Kanbans in a Just-In-Time Production System", Computers Industrial Engineering, 24 (4), 1993, pp. 551-559.
[3] Savsar, M. and Aljawini, A., "Simulation Analysis of Just-In-Time Production Systems", International Journal of Production Economics, 42, 1995, pp. 67-78.
[4] Savsar, M., "Effects of Kanban Withdrawal Policies and Other Factors on the Performance of JIT Systems: A Simulation Study", Int. Journal of Prod. Res., 34 (10), 1996, pp. 2879-2899.
[5] Savsar, M., "Simulation Analysis of a Push-Pull System for an Electronic Assembly Line", Int. Journal of Prod. Economics, 51, 1997, pp. 205-214.
[6] Savsar, M. and Choueiki, H. M., "A Neural Network Procedure for Kanban Allocation in JIT Production Control Systems", Int. Journal of Prod. Research, 38 (14), 2000, pp.3247-3265.
[7] Olhager, J. and Ostlung, B., "An Integrated Push-Pull Manufacturing Strategy" European Journal of Operational Research, 45, 1990, pp. 135-142.
[8] Hodgson, T.J. and Wang, D., "Optimal Hybrid Push/Pull Control Strategies for Parallel Multistage System: Part II", International Journal of Production Research, 29 (7), 1991, pp. 1453-1460.
[9] Wang, H. and Xu, C., "Hybrid Push/Pull Production Control Strategy Simulation and its Applications", Production Planning and Control, 8, 1997, pp. 142-151.
[10] Beamon, B.M. and Bermund, J.M., "A hybrid push-pull Ccontrol algorithm for multi-stage, multi-line production systems", Production Planning & Control,11(4), 2000, pp. 349-356.
[11] Duri, C., Frein, Y., and Dimascolo, M., "Comparison among three pull control policies: Kanban, Base Stock and Generalized Kanban", Annals of Operations Research, 93(1), 2000, pp. 41-47.
[12] Hillier, F.S. and So, K. C., "The Effect of the Coefficient of Variation of Operation Times on the Allocation of Storage Space in Production Line System", IIE Transactions, (23), 1991, pp. 198-206.
[13] Hillier, F.S., So, K. C., and Boling, R. W., "Notes: Toward Characterizing the Optimal Allocation of Storage Space in Production Line Systems with Variable Processing Times", Management Sci. 39(1), 1993, pp. 126-133.
[14] Papadopoulos, H. T. and Heavey, C., "Queuing Theory in Manufacturing Systems Analysis and Design: A Classification of Models for Production and Transfer Lines", European Journal of Operational Research, (92), 1996, pp. 1-27.
[15] Papadopoulos, H. T., and Vouros, G. A., "A Model Management System (MMS) for the Design and Operation of Production Lines", Int. Journal of Production Research, 35(8), 1996, 2213-2236.
[16] Powel, S. G. and Pyke, D. F., "Allocation of buffers to serial production lines with bottlenecks" IIE Transactions, 28, 1996, pp.18-29.
[17] Vouros, G. A. and Papadopoulos, H.T., "Buffer Allocation in Unreliable Production Lines Using a Knowledge Based System", Computers & Operations Research, 25(12), 1996, pp. 1055-1067.
[18] Vouros, G. A., Vidalis, M. I., and Papadopoulos, H. T., "A Heuristic Algorithm for Buffer Allocation in Unreliable Production Lines", International Journal of Quantitative Methods, 6(1), 2000, pp. 23-43.
[19] Spinellis, D.D. and Papadopoulos, C.T., "Stochastic Algorithms for Buffer Allocation in Reliable Production Lines", Mathematical Problems in Engineering, 5, 2000a, pp. 441-458.
[20] Spinellis, D.D. and Papadopoulos, C.T., "A Simulated Annealing Approach for Buffer Allocation in Reliable Production Lines", Annals of Operations Research, 93, 2000b, pp. 373-384.
[21] Gershwin, S.B. and Schor, J.E., "Efficient algorithms for buffer space allocation", Annals of Operations Research, 93, 2000, pp. 117-144.
[22] Savsar, M. and Youssef, A. S., "An Integrated Simulation-Neural Network Meta Model Application in Designing Production Flow Lines", WSEAS Transactions on Electronics, 2 (1), 2004, pp. 366-371.
[23] Savsar, M. "Effects of Maintenance Policies on the Productivity of Flexible Manufacturing Cells", OMEGA, Vol. 34, 2006, pp. 274-282.

Mehmet Savsar is a professor and chairman of the Industrial & Management Systems Engineering Department at Kuwait University. Prof. Savsar received his B.Sc. degree from Black Sea Technical University in Turkey, his M.Sc. and PhD. Degrees from the Pennsylvania State University, PA, USA in the area of Industrial Engineering. He has been with Kuwait University since 1997. Prior to joining Kuwait University, he has worked as a faculty member at Anatolian University in Turkey and at King Saud University in Riyadh, Saudi Arabia and as a researcher at Pennsylvania State University, USA. Prof. Savsar has taught a variety of courses in the areas of Production Planning and Inventory Control, JIT Production Control, Quality Control, Maintenance and Reliability, Operations Research, Stochastic Processes, Computer Simulation, Plant Layout and Facilities Planning, Engineering Cost Analysis, and Manufacturing Systems.

His research interests include: Modeling and Analysis of Production Systems, Plant Layout, Reliability and Maintenance Management, Flexible Manufacturing Systems, Quality Control, and Just-In-Time (JIT) Production Control. He has over 100 publications in refereed international journals and international conference proceedings. He has completed several research projects. He serves in editorial boards of several international journals and is a referee to several journals, including EJOR, IJAMT, Simulation, Int. J. of Systems Science, Int. J. of Production Economics, Int. Journal of Prod. Research. E-Mail: msavsar@yahoo.com