Modeling and Analysis of Congestion in the Design of Facility Layouts

Saifallah Benjaafar
Division of Industrial Engineering, Department of Mechanical Engineering, University of Minnesota, Minneapolis, Minnesota 55455
saif@tc.umn.edu

Reducing manufacturing lead times and minimizing work-in-process (WIP) inventories are the cornerstones of popular manufacturing strategies such as Lean, Quick Response, and Just-in-Time Manufacturing. In this paper, we present a model that captures the relationship between facility layout and congestion-related measures of performance. We use the model to introduce a formulation of the facility layout design problem where the objective is to minimize work-in-process (WIP). In contrast to some recent research, we show that layouts obtained using a WIP-based formulation can be very different from those obtained using the conventional quadratic assignment problem (QAP) formulation. For example, we show that a *QAP-optimal* layout can be WIP-infeasible. Similarly, we show that two QAP-optimal layouts can have vastly different WIP values. In general, we show that WIP is not monotonic in material-handling travel distances. This leads to a number of surprising results. For instance, we show that it is possible to reduce overall distances between departments but increase WIP. Furthermore, we find that the relative desirability of a layout can be affected by changes in material-handling capacity even when travel distances remain the same. We examine the effect of various system parameters on the difference in WIP between QAP- and WIP-optimal layouts. We find that although there are conditions under which the difference in WIP is significant, there are those under which both layouts are WIP-equivalent.
(*Facility Layout; Queueing Networks; Quadratic Assignment Problem; Material Handling; Performance Evaluation*)

1. Introduction

Reducing manufacturing lead times and minimizing work-in-process (WIP) inventories are the cornerstones of popular manufacturing strategies such as Lean, Quick Response, and Just-in-Time Manufacturing (Hopp and Spearman 2000, Suri 1998, Womack and Jones 1996). Although various facets of the manufacturing function have been redesigned in recent years to support lead-time and inventory reduction, the physical organization and layout of manufacturing facilities have continued to be largely guided by more traditional concerns of efficiency and cost (Benjaafar et al. 2002). This approach to facility design tends to ignore the potential impact of facility layout and material handling on short-term operational effectiveness. It also does not account for the possibility of using facility design strategically to support a manufacturing firms' need for agility and responsiveness.

In this paper, we show that physical layout of facilities can impact operational performance in significant ways, as measured by lead time, WIP, or throughput. We also show that the traditional facility design criterion can be a poor indicator of this performance. More importantly, we identify characteristics of layouts that tend to reduce lead time and WIP and improve throughput. Because operational performance is driven by system congestion, which

0025-1909/02/4805/0679$5.00
1526-5501 electronic ISSN

is itself a function of system capacity and variability, we introduce a model that captures the relationship between layout and congestion. We use the model to examine the effect of layout on capacity and variability and derive insights that can be used to generate what might be termed as *agile* or *quick response* layouts.

In our analysis, we focus on WIP as our primary measure of congestion. However, by virtue of Little's law, this could also be related to both lead time and throughput. In our analysis, we assume that capacity decisions have already been made. Therefore, our objective is to determine a layout that minimizes WIP given fixed material-handling capacity. Our model can, however, be used to solve problems where both material-handling capacity and layout are decision variables, so that the objective is to minimize both investment and congestion costs. Alternatively, the model can be used to formulate problems where the objective is to minimize one type of cost subject to a constraint on the other.

Our work is in part motivated by two recent papers by Fu and Kaku (1997a, 1997b) in which they present a plant layout problem formulation for job-shop-like manufacturing systems where the objective is to minimize average work-in-process. In particular, they investigate conditions under which the familiar quadratic assignment problem (QAP) formulation, where the objective is to minimize average material-handling costs, also minimizes average work-in-process. By modeling the plant as an open queueing network, they show that under a set of assumptions the problem reduces to the quadratic assignment problem. Using a simulation of an example system, they found that the result apparently holds under much more general conditions than are assumed in the analytical model.

To obtain a closed-form expression of expected WIP, Fu and Kaku (1997b) make the following assumptions: (1) external part type arrival processes into the system are Poisson; (2) processing times at a department are i.i.d. exponential; (3) material handling is carried out via discrete material-handling devices, such as forklifts or automated guided vehicles (AGV), (4) travel times of the material-handling devices are exponentially distributed, (5) input and output buffer sizes at departments are sufficiently large so that blocking is negligible, and (6) service discipline is first come, first served (FCFS). In modeling travel times, they ignore empty travel by the material-handling devices and account only for full trips. These assumptions allow them to treat the network as a Jackson queueing network—i.e., a network of independent $M/M/n$ queues—for which a closed-form analytical expression of average WIP is available. They show that WIP accumulation at the processing departments—at both the input and output buffers—is always independent of the layout and that travel times are a linear transformation of the average distance traveled by the material-handling system when full. Because the measure of material-handling cost used in the QAP formulation is itself a linear function of the same average travel distance, they show that the queueing and QAP formulations are equivalent.

In this paper, we show that when some of the assumptions used by Fu and Kaku are relaxed, their key observation regarding the equivalence of the two formulations is not always valid. In fact, under general conditions, we show that layouts generated using the queueing-based model can be very different from those obtained using the conventional QAP formulation. More importantly, we show that the choice of layout does have a direct impact on WIP accumulation at both the material-handling system and at the individual departments, and that the behavior of expected WIP is not necessarily monotonic in the average distance traveled by the material-handling device. This leads to a number of surprising and counterintuitive results. In particular, we show that reducing overall distances between departments can increase WIP. We also show that the desirability of a layout can be affected by non-material-handling factors, such as department, utilization levels, variability in processing times at departments, and variability in product demands. In general, we find the objective function used in the QAP formulation to be a poor indicator of WIP. For example, we show that a *QAP-optimal* layout can be WIP-infeasible—i.e., it results in infinite WIP. Similarly, we show that two QAP-optimal layouts can have vastly different

WIP values. Furthermore, we find that the QAP formulation, by accounting only for full travel by the material-handling system, ignores the important role that empty travel plays. For example, we find that minimizing full travel, as the QAP formulation does, can cause empty travel to increase, which in turn can increase WIP. This leads to some additional counterintuitive results. For instance, we find that it can be highly desirable to place departments in neighboring locations even though there is no direct material flow between them. Likewise, we show that it can be beneficial to place departments with high intermaterial flows in distant locations from each other. On the other hand, we also show that there are conditions on flow distribution, layout geometry, and material-handling capacity under which both a QAP-optimal and a WIP-optimal layout are WIP-equivalent.

In our model, we relax several of the assumptions used by Fu and Kaku (1997a, 1997b). In particular, we let part interarrival times and processing times be generally distributed and determined by the number of product types and their routings. We also allow the distances traveled by the material-handling devices to be determined by the layout configuration. This allows us to characterize exactly the distribution of travel times and to capture both empty and full travel by the material-handling system. We show that relaxing these assumptions enables us to capture important interactions between the layout configuration, the distribution of travel times, and several operating characteristics of the processing departments. These interactions are absent in the Fu and Kaku (1997b) model, which in part explains the results we obtain. Our results are applicable to systems where a shared material-handling system consisting of discrete devices is used. This excludes systems with continuous conveyors and systems with dedicated material handling for each segment of the flow.

Although there is an extensive literature on design of facility layouts (see Meller and Gau 1996 for a recent review), the design criterion in the majority of this literature is material-handling cost, measured either directly as a function of material-handling distances or indirectly through an adjacency score (Tompkins 1996). Few papers consider operational performance measures, such as WIP, throughput, or cycle time, as design criteria or design constraints. Among those that do, we note the previously mentioned papers by Fu and Kaku (1997a, 1997b) and papers by Kouvelis and Kiran (1990, 1991). Kouvelis and Kiran introduce a modified formulation of the quadratic assignment problem where the objective is to minimize the sum of material handling and WIP-holding costs subject to a constraint on throughput. In modeling travel times, however, they ignore empty travel and consider only the mean of the full travel-time distribution. A similar approach is also used by Solberg and Nof (1980) in evaluating different layout configurations. Outside the layout literature, there is a related body of research on design of material-handling systems. Although some of this literature addresses the modeling of empty travel, especially as it pertains to the design of automated guided vehicles (AGV), it generally assumes a fixed layout (see, for example, Egebelu (1987) for a discussion of empty travel in AGV design under varying assumptions). A recent review of this literature can be found in Johnson and Brandeau (1996).

2. Model Formulation

We use the following assumptions and notation.

(i) The plant produces N products. Product demands are independently distributed random variables. Unit orders arrive according to a renewal process with rate D_i (average demand per unit time) and a squared coefficient of variation C_i^2 for $i = 1, 2, \dots, N$. The squared coefficient of variation denotes the ratio of the variance over the squared mean of unit order interarrival times.

(ii) Material handling is carried out by a set of discrete material-handling devices, or transporters. For ease of discussion, we first present the model for the single-transporter case. Extension to systems with multidevices is discussed in §3. In responding to a request, a material-handling device travels empty from the department location of its last delivery to the department location of the current request. Material transfer requests are serviced on a first-come-first-served (FCFS) basis. In the absence of any requests, the material-handling device remains at the location of its last delivery.

(iii) The travel time between any pair of locations k and l, t_{kl}, is assumed to be deterministic and is given by $t_{kl} = d_{kl}/v$, where d_{kl} is the distance between locations k and l and v is the speed of the material-handling transporter.

(iv) Products are released to the plant from a loading department and exit the plant through an unloading (or shipping) department. Departments are indexed from $i = 0$ to $M+1$, with the indices $i = 0$ and $M+1$ denoting, respectively, the loading and unloading departments.

(v) The plant consists of M processing departments, with each department consisting of a single server (e.g., a machine) with ample storage for work-in-process. Jobs in the queue are processed in first-come-first-served order. The amount of material flow, λ_{ij}, between a pair of departments i and j is determined from the product routing sequence and the product demand information. The total amount of workload at each department is given by:

$$\lambda_i = \sum_{k=0}^{M} \lambda_{ki} = \sum_{j=1}^{M+1} \lambda_{ij} \quad \text{for } i = 1, 2, \ldots, m, \tag{1}$$

$$\lambda_0 = \lambda_{M+1} = \sum_{i=1}^{N} D_i, \quad \text{and} \tag{2}$$

$$\lambda_t = \sum_{i=0}^{M} \sum_{j=1}^{M+1} \lambda_{ij}, \tag{3}$$

where λ_t is the workload for the material-handling system.

(vi) Processing times at each department are independent and identically distributed with an expected processing time $E(S_i)$ and a squared coefficient of variation $C_{s_i}^2$ for $i = 0, 1, \ldots, M+1$ (the processing time distribution is determined from the processing times of the individual products).

(vii) There are K locations to which departments can be assigned. A layout configuration corresponds to a unique assignment of departments to locations. We use the vector notation $\mathbf{x} = \{x_{ik}\}$, where $x_{ik} = 1$ if department i is assigned to location k and $x_{ik} = 0$ otherwise, to differentiate between different layout configurations. The number of locations is assumed to be greater than or equal to the number of departments.

We model the plant as an open network of GI/G/1 queues, with the material-handling system being a central server queue. Note that because parts are delivered to the departments by the material-handling system, the operating characteristics of the material-handling system, such as utilization and travel-time distribution, directly affect the interarrival time distribution of parts to the departments. Similarly, since the queue for the material-handling system consists of the department output buffers, the interarrival time distribution to this queue is determined by the departure process from the departments, which is in turn determined by the operating characteristics of the departments. Therefore, there is a close coupling between the inputs and outputs of the processing departments and the material-handling system. In our model, we explicitly capture this coupling and show that there exists a three-way interaction between the department operating characteristics, the operating characteristics of the material-handling system, and the layout configuration, and that this interaction has a direct effect on WIP accumulation.

In order to show this effect, let us first characterize the travel-time distribution. In responding to a material-transfer request, the material-handling device performs an empty trip from its current location (the location of its last delivery), at some department r, followed by a full trip from the origin of the current request, say department i, to the destination of the transfer request at a specified department j (see Figure 1). The probability distribution p_{rij} of an empty trip from r to i followed by a full trip from i to j is, therefore, given by:

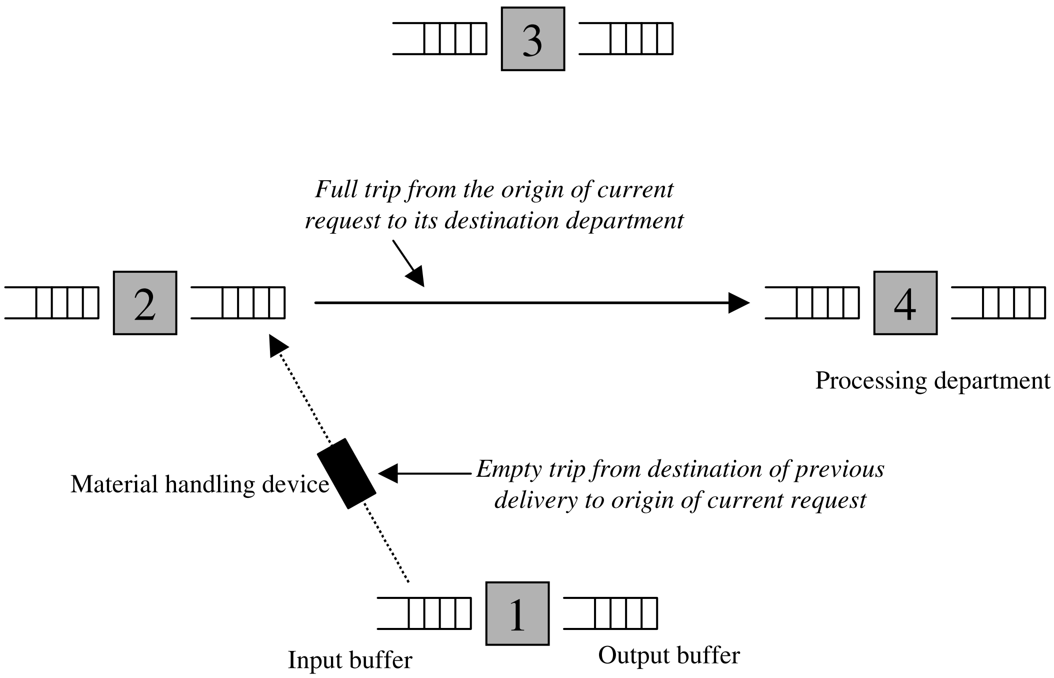
$$p_{rij} = \sum_{k=0}^{M} p_{kr} p_{ij}, \tag{4}$$

where $p_{ij}$is the probability of a full trip from department i to department j, which can be obtained as

$$p_{ij} = \frac{\lambda_{ij}}{\sum_{i=0}^{M} \sum_{j=1}^{M+1} \lambda_{ij}}. \tag{5}$$

We show in §3 that Expressions (4) and (5) are also valid for systems with multiple transporters. Given a layout configuration $\mathbf{x}$, the time to perform an empty trip from department r to department i followed by

Figure 1 **Empty and Full Travel in a System with Discrete Material Handling Devices**



a full trip to department j is given by $t_{rij}(\mathbf{x}) = t_{ri}(\mathbf{x}) + t_{ij}(\mathbf{x})$, where

$$t_{ij}(\mathbf{x}) = \sum_{k=1}^{K}\sum_{l=1}^{K} x_{ik}x_{jl}d_{kl}/v \tag{6}$$

and is the travel time from department i to department j. From (4)–(6), we can obtain the mean and variance of travel time as follows:

$$\begin{aligned} E(S_t) &= \sum_{r=1}^{M+1}\sum_{i=0}^{M}\sum_{j=1}^{M+1} p_{rij}t_{rij}(\mathbf{x}) \\ &= \sum_{r=1}^{M+1}\sum_{i=0}^{M}\sum_{j=1}^{M+1}\sum_{k=0}^{M}\left(\lambda_{kr}\lambda_{ij}/\lambda_t^2\right)t_{rij}(\mathbf{x}), \end{aligned} \tag{7}$$

and

$$\mathrm{Var}\ (S_t) = E\left(S_t^2\right) - E(S_t)^2, \tag{8}$$

where

$$\begin{aligned} E\left(S_t^2\right) &= \sum_{r=1}^{M+1}\sum_{i=0}^{M}\sum_{j=1}^{M+1} p_{rij}\left(t_{rij}(\mathbf{x})\right)^2 \\ &= \sum_{r=1}^{M+1}\sum_{i=0}^{M}\sum_{j=1}^{M+1}\sum_{k=0}^{M}\left(\lambda_{kr}\lambda_{ij}/\lambda_t^2\right)\left(t_{rij}(\mathbf{x})\right)^2, \end{aligned} \tag{9}$$

and

$$\begin{aligned} t_{rij}(\mathbf{x}) &= \sum_{k}\sum_{l}\sum_{s} x_{rk}x_{il}x_{js}(d_{kl}+d_{ls})/v \\ &= \sum_{k}\sum_{l} x_{rk}x_{il}d_{kl}/v + \sum_{l}\sum_{s} x_{il}x_{js}d_{ls}/v. \end{aligned} \tag{10}$$

We can also obtain the average utilization of the material-handling system, ρ_t, as:

$$\rho_t = \lambda_t E(S_t) = \sum_{r=1}^{M+1}\sum_{i=0}^{M}\sum_{j=1}^{M+1}\sum_{k=0}^{M}(\lambda_{kr}\lambda_{ij}/\lambda_t)(t_{ri}(\mathbf{x}) + t_{ij}(\mathbf{x})), \tag{11}$$

which can be simplified as:

$$\rho_t = \sum_{r=1}^{M+1}\sum_{i=0}^{M}(\lambda_r\lambda_i/\lambda_t)t_{ri}(\mathbf{x}) + \sum_{i=0}^{M}\sum_{j=1}^{M+1}\lambda_{ij}t_{ij}(\mathbf{x}) \tag{12}$$

or equivalently,

$$\rho_t = \rho_t^e + \rho_t^f, \tag{13}$$

where

$$\rho_t^e = \sum_{r=1}^{M+1}\sum_{i=0}^{M}(\lambda_r\lambda_i/\lambda_t)t_{ri}(\mathbf{x}) \tag{14}$$

corresponds to the utilization of the material-handling system due to empty travel, and

$$\rho_t^f = \sum_{i=0}^{M} \sum_{j=1}^{M+1} \lambda_{ij} t_{ij}(\mathbf{x}) \tag{15}$$

is the utilization of the material-handling system due to full travel.

From the above expressions, we can see that the travel-time distribution is determined by the layout configuration and that this distribution is not necessarily exponential. As a result, the arrival process to the departments is not always Poisson distributed, even if external arrivals are Poisson and processing times are exponential. This means that our system cannot be treated, in general, as a network of M/M/1 queues. Unfortunately, exact analytical expressions of performance measures of interest, such as expected WIP, expected time in system, and time in queue, are difficult to obtain for queues with general interarrival and processing time distributions. Therefore, to estimate these measures of performance we resort to network decomposition and approximation techniques, where each department, as well as the material-handling system, is treated as being stochastically independent, with the arrival process to and the departure process from each department and the material-handling system being approximated by renewal processes. Furthermore, we assume that two parameters, mean and variance, of the job interarrival and processing time distributions are sufficient to estimate expected WIP at each department. The decomposition and approximation approach has been widely used to analyze queueing networks in a variety of contexts (Bitran and Dasu 1992, Bitran and Tirupati 1989, Buzacott and Shanthikumar 1993, Whitt 1983a). A number of good approximations have been proposed by several authors (see Bitran and Dasu 1992 for a recent review). In this paper, the approximations we use have been first proposed by Kraemer and Langenbach-Belz (1976) and later refined by Whitt (1983a, 1983b) and shown to perform well over a wide range of parameters (Buzacott and Shanthikumar 1993, Whitt 1983b). The approximations coincide with the exact analytical results obtained by Fu and Kaku for the special case of Poisson arrival and exponential processing/travel times. Because in layout design our objective is primarily to obtain a ranked ordering of different layout alternatives, approximations are sufficient as long as they guarantee accuracy in the ordering of these alternatives. Approximations are also adequate when we are primarily interested, as we are in this paper, in the qualitative behavior of the performance measures. Comparisons of our analytical results with results obtained using simulation are discussed in §4.

Under a given layout, expected WIP at each department i $(i = 0, 1, \ldots, M+1)$ is approximated as follows:

$$E(\mathrm{WIP}_i) = \frac{\rho_i^2 \left(C_{a_i}^2 + C_{s_i}^2\right) g_i}{2(1-\rho_i)} + \rho_i, \tag{16}$$

where $\rho_i = \lambda_i E(S_i)$ is the average utilization of department i, $C_{a_i}^2$ and $C_{s_i}^2$ are, respectively, the squared coefficients of variation of job interarrival and processing times, and

$$g_i \equiv g_i\left(C_{a_i}^2, C_{s_i}^2, \rho_i\right) = \begin{cases} \exp\left[\frac{-2(1-\rho_i)(1-C_{a_i}^2)^2}{3\rho_i(C_{a_i}^2 + C_{s_i}^2)}\right] & \text{if } C_{a_i}^2 < 1, \\ 1 & \text{if } C_{a_i}^2 \geq 1. \end{cases} \tag{17}$$

Similarly, expected WIP at the material-handling system is approximated by:

$$E(\mathrm{WIP}_t) = \frac{\rho_t^2 \left(C_{a_t}^2 + C_{s_t}^2\right) g_t}{2(1-\rho_t)} + \rho_t. \tag{18}$$

Note that ρ_t and ρ_i must be less than one for expected work-in-process to be finite. The squared coefficients of variation can be approximated as follows (Buzacott and Shanthikumar 1993, Whitt 1983a):

$$C_{a_i}^2 = \sum_{j \neq i} \frac{\lambda_j \hat{p}_{ji}}{\lambda_i} \left(\hat{p}_{ji} C_{d_i}^2 + (1 - p_{ji})\right) + \frac{\lambda_0 \gamma_i}{\lambda_i} \left(\gamma_i C_{a_0}^2 + (1-\gamma_i)\right), \quad \text{and} \tag{19}$$

$$C_{d_i}^2 = \rho_i^2 C_{s_i}^2 + (1-\rho_i^2) C_{a_i}^2, \tag{20}$$

where $C_{d_i}^2$ is the squared coefficient of interdeparture time from department i, $\hat{p}_{ij}$ is the routing probability from node i to node j (nodes include departments and the material handling device), γ_i is the fraction of external arrivals that enter the network through node i, and $1/\lambda_0$ and $C^2_{a_0}$ are, respectively, the mean and squared coefficient of variation of the external job interarrival times. In our case, $\gamma_0 = 1$ and $\gamma_i = 0$ for all others since all jobs enter the cell at the loading department. The routing probability from departments $i = 0$ through M to the material-handling system is always one, that from the material-handling system to departments $j = 1$ through $M+1$ is

$$\hat{p}_{tj} = \frac{\sum_{i=0}^{M+1} \lambda_{ij}}{\sum_{i=0}^{M} \sum_{j=0}^{M+1} \lambda_{ij}}, \tag{21}$$

and to the loading department ($j = 0$) is zero. Parts exit the cell from department $M+1$ (unloading department) so that all the routing probabilities from that department are zero. Substituting these probabilities in the above expression, we obtain:

$$C^2_{a_0} = \sum_{i=1}^{N} \left(D_i \Big/ \sum_{i=1}^{N} D_i \right) C_i^2, \tag{22}$$

$$C^2_{a_t} = \sum_{i=0}^{M} (\lambda_i/\lambda_t) C^2_{d_i} = \sum_{i=0}^{M} \pi_i C^2_{d_i}, \quad \text{and} \tag{23}$$

$$C^2_{a_i} = \pi_i C^2_{d_t} + 1 - \pi_i \quad \text{for } i = 1, 2, \dots, M+1, \tag{24}$$

where $\pi_i = \lambda_i/\lambda_t$. Equalities (22)–(24), along with (20), can be simultaneously solved to yield:

$$C^2_{a_i} = \pi_i \left(\rho_t^2 C^2_{s_t} + (1-\rho_t^2) C^2_{a_t} \right) + 1 - \pi_i, \quad \text{for } i = 1, 2, \dots, M+1, \quad \text{and} \tag{25}$$

$$C^2_{a_t} = \left(\sum_{i=0}^{M} \pi_i \rho_i^2 C^2_{s_i} + \sum_{i=1}^{M} \pi_i (1-\rho_i^2)(1-\pi_i) + \sum_{i=1}^{M} \pi_i^2 (1-\rho_i^2) \rho_t^2 C^2_{s_t} + \pi_0 (1-\rho_0^2) C^2_{a_0} \right) \Bigg/ \left(1 - \sum_{i=1}^{M} \pi_i^2 (1-\rho_i^2)(1-\rho_t^2) \right). \tag{26}$$

From the expression of expected WIP, we can obtain additional measures of performance. For example, by virtue of Little's law, expected flow time through department i ($i = 0$), is simply $E(F_i) = E(\text{WIP}_i)/\lambda_i$ and expected total flow time in system is $E(F) = E(\text{WIP})/(D_1 + \cdots + D_N)$, where

$$E(\text{WIP}) = \sum_{i=0}^{M+1} E(\text{WIP}_i) + E(\text{WIP}_t)$$

is total expected WIP in the system. We can also obtain expected flow time in the system for a specific product j as:

$$E(F^{(j)}) = \sum_{i=0}^{M+1} n_{ij} E(F_i),$$

where n_{ij} is the number of times product j visits department i.

Any of the above performance measures could be used as a criterion in layout design. In the remainder of this article, we limit ourselves to expected total WIP. However, the analysis can be extended to other measures. The layout design problem can be formulated as:

$$\text{Minimize} \quad E(\text{WIP}) = \sum_{i=0}^{M+1} E(\text{WIP}_i) + E(\text{WIP}_t) \tag{27}$$

$$\text{subject to:} \quad \sum_{k=1}^{K} x_{ik} = 1 \qquad i = 0, 2, \dots, M+1, \tag{28}$$

$$\sum_{i=0}^{M+1} x_{ik} = 1 \qquad k = 1, 2, \dots, K, \tag{29}$$

$$\rho_t \le 1, \tag{30}$$

$$x_{ik} = 0, 1 \qquad i = 0, 2, \dots, M+1; \quad k = 1, 2, \dots, K. \tag{31}$$

The above formulation shares the same constraints, Constraints (28), (29), and (31), as the QAP formulation. Constraints (28) and (29) ensure, respectively, that each department is assigned to one location and each location is assigned to one department. We require an additional constraint, Constraint (30), to ensure that a selected layout is feasible and will not result in infinite work-in-process. As in the QAP formulation, we assume $K = M+2$. The case where $K > M+2$ can be handled by introducing dummy departments with zero input and output flows. The objective function is, however, different from that of the QAP. In the conventional QAP, the objective function is a positive linear transformation of the expected full travel time and is of the form:

$$\text{Minimize } z = \sum_i \sum_j \sum_k \sum_l x_{ik} x_{jl} \lambda_{ij} d_{kl}. \tag{32}$$

Therefore, a solution that minimizes average full travel time between departments is optimal. Because expected WIP is not, in general, a linear function of average full travel time, the solutions obtained by the two formulations, as we show in the next section, can be different. However, a special case where the two formulations lead to the same solution is the one considered by Fu and Kaku, where all interarrival, processing, and transportation times are assumed to be exponentially distributed and empty travel time is negligible. In this case, we have $C^2_{a_i} = C^2_{s_i} = 1$, for $i = 0, 1, \ldots, M+1$, which when substituted in the expression of expected WIP, while ignoring empty travel, leads to:

$$E(\text{WIP}) = \sum_{i=0}^{M+1} \frac{\rho_i}{1-\rho_i} + \frac{\rho_t}{1-\rho_t}, \tag{33}$$

with

$$\rho_t = \sum_i \sum_j \sum_k \sum_l \lambda_{ij} x_{ik} x_{jl} d_{kl}/v. \tag{34}$$

Since only $E(\text{WIP}_t)$ is a function of the layout (given the exponential assumption, the arrival process to the departments is always Poisson regardless of the layout configuration), and since $E(\text{WIP}_t)$ is strictly increasing in ρ_t, any solution that minimizes ρ_t also minimizes the overall WIP. Noting that ρ_t is minimized by minimizing $z = \sum_i \sum_j \sum_k \sum_l \lambda_{ij} x_{ik} x_{jl} d_{kl}$, we can see that minimizing z also minimizes expected WIP. In the next section, we show that when we either (1) account for empty travel or (2) relax the exponential assumption regarding interarrival, processing, or travel times (as we do in our model), the equivalence between the QAP and the queueing-based model does not hold any longer.

The quadratic assignment problem has been shown to be NP-hard (Pardalos and Wolkowicz 1994). Since the objective function in (27) is a nonlinear transformation of that of the QAP, the formulation in (27)–(31) also leads to an NP-hard problem. Although for relatively small problems implicit enumeration (e.g., branch and bound) can be used to solve the problem to optimality (Pardalos and Wolkowicz 1994), for most problems we must resort to a heuristic solution approach. Several heuristics have been proposed for solving the QAP (see Pardalos and Wolkowicz 1994 for a recent review) and any of these could be used to solve our model as well. In a software implementation of the formulation in (27)–(31), Yang and Benjaafar (2001) used both implicit enumeration and a modified 2-opt heuristic, similar to the one proposed by Fu and Kaku (1997a), to solve the problem. In this paper, we limit our discussion mostly to layouts where the QAP-optimal layout is easily identified.

3. Systems with Multiple Transporters

For a system with multiple transporters, the travel-time distribution is affected by the dispatching policy used to select a transporter whenever two or more are available to carry out the current material-handling request. Analysis of most dispatching policies is difficult. In this section, we treat the mathematically tractable case of randomly selecting a device when two or more are idle. Although not optimal, this policy does yield a balanced workload allocation among the different devices. Assuming transfer requests are processed on a first-come-first-served basis, this policy also ensures an assignment of transporters to departments proportional to the departments' workloads. As in the single-transporter case, we assume that vehicles remain at the location of their last delivery if there are no pending requests.

In order to characterize the probability distribution of travel time in a system with n_t transporters ($n_t > 1$), we need to first obtain the probability p_{rij} of an empty trip from department r followed by a full trip from department i to j. The probability of a full trip from i to j is still given by (5). The probability of an empty trip from r can be written as follows:

$$\begin{aligned}&\text{Prob(empty trip from } r)\\ &\quad = \sum_{n_s=1}^{n_t} \sum_{n_r=1}^{n_s} \{\text{Prob (selecting transporter at } r \mid n_r \\ &\qquad\qquad \text{and } n_s) \text{ Prob } (n_r \mid n_s) \text{ Prob } (n_s)\}\end{aligned} \tag{35}$$

where Prob(selecting transporter at $r \mid n_r$ and n_s) refers to the probability of selecting one of the idle vehicles at department r given that there are n_r idle vehicles at

r and n_s total idle vehicles in the system, Prob $(n_r \mid n_s)$ is the probability of having n_r idle vehicles at department r where $n_r = 1, 2, \ldots, n_s$, given that there are n_s idle vehicles in the system, and $\text{Prob}(n_s)$ is the probability of having n_s idle vehicles in the system, where $n_s = 1, 2, \ldots, n_t$. It is straightforward to show that

$$\text{Prob}(\text{selecting transporter at } r \mid n_r \text{ and } n_s) = n_r/n_s, \quad \text{and} \tag{36}$$

$$\text{Prob}(n_r \mid n_s) = \binom{n_s}{n_r} p_r^{n_r}(1-p_r)^{n_s-n_r}, \tag{37}$$

where p_r is the probability of an idle vehicle being at department r which is given by:

$$p_r = \sum_{i=0}^{M} p_{ir} = \sum_{i=0}^{M} \lambda_{ir} \Big/ \sum_{i=0}^{M} \sum_{j=1}^{M+1} \lambda_{ij}. \tag{38}$$

We can now write the probability p_{rij} as

$$p_{rij} = \left\{ \sum_{n_s=1}^{n_t} \sum_{n_r=1}^{n_s} \frac{n_r}{n_s} \binom{n_s}{n_r} p_r^{n_r}(1-p_r)^{n_s-n_r} \text{Prob}(n_s) \right\} p_{ij} \tag{39}$$

or equivalently as

$$p_{rij} = \left\{ \sum_{n_s=1}^{n_t} (1/n_s)\text{Prob}(n_s) \sum_{n_r=1}^{n_s} n_r \binom{n_s}{n_r} p_r^{n_r}(1-p_r)^{n_s-n_r} \right\} p_{ij}. \tag{40}$$

Noting that

$$\sum_{n_r=1}^{n_s} n_r \binom{n_s}{n_r} p_r^{n_r}(1-p_r)^{n_s-n_r} = n_s p_r, \quad \text{and} \tag{41}$$

$$\sum_{n_s=1}^{n_t} \text{Prob}(n_s) = 1,$$

yields to

$$p_{rij} = \left\{ \sum_{n_s=1}^{n_t} \text{Prob}(n_s) p_r \right\} p_{ij} = p_r p_{ij}, \tag{42}$$

which is the same as in the single-transporter case (a result due to the random nature of the selection rule). One could also have argued directly that given the probabilistic routing and the random selection rule for the material-handling system, the probability of an empty trip from r to i followed by a full trip from i to j would depend only on the workloads assigned to each department and not on the number of transporters.

The mean and variance of travel time can now be obtained as in (7) and (8). Expected WIP due to the transporters can be obtained using approximations for a $GI/G/n_t$ queue. Similarly, the departure process from the transporters can be approximated as a departure process from a $GI/G/n_t$ queue, which can then be used to characterize the arrival process to the departments and the transporters as in (25)–(26). A detailed analysis and software implementation of this approach can be found in Yang and Benjaafar (2001). In §5, we examine the effect of the number of transporters on layout performance.

4. Model Analysis and Insights

In this section, we show that layouts obtained using a WIP-based formulation can be very different from those using the QAP formulation. We trace these differences to two major factors: *empty travel* and *travel-time variability*.

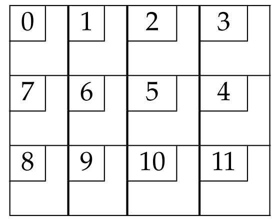
4.1. The Effect of Empty Travel

In the following set of observations, we examine the impact of empty travel. We show that a layout that minimizes ρ_f does not necessarily minimize ρ_e, and consequently, a layout that minimizes ρ_f does not necessarily minimize WIP. In fact, we show that a QAP-optimal layout (i.e., a layout that minimizes ρ_f) is not even guaranteed to be feasible. More generally, we show that two QAP-optimal layouts can result in different WIP values. Furthermore, under certain conditions we find that WIP is reduced more effectively by reducing empty travel, even if this increases full travel. This means that sometimes it can be desirable to place departments in neighboring locations even though there is no direct material flow between them. This also means that it can be beneficial to place departments with high intermaterial flows in distant locations from each other.

Observation 1. A layout that minimizes full travel does not necessarily minimize WIP.

The result follows from noting that reducing ρ_f can increase ρ_e. If the increase in ρ_e is sufficiently large,

Figure 2 Data for Example Layout



(a) Available department locations

From/To	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
0	0	1	2	3	4	3	2	1	2	3	4	5
1	1	0	1	2	3	2	1	2	3	2	3	4
2	2	1	0	1	2	1	2	3	4	3	2	3
3	3	2	1	0	1	2	3	4	5	4	3	2
4	4	3	2	1	0	1	2	3	4	3	2	1
5	3	2	1	2	1	0	1	2	3	2	1	2
6	2	1	2	3	2	1	0	1	2	1	2	3
7	1	2	3	4	3	2	1	0	1	2	3	4
8	2	3	4	5	4	3	2	1	0	1	2	3
9	3	2	3	4	3	2	1	2	1	0	1	2
10	4	3	2	3	2	1	2	3	2	1	0	1
11	5	4	3	2	1	2	3	4	3	2	1	0

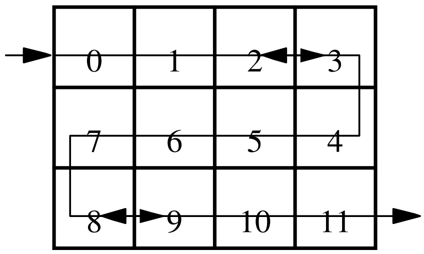
(b) Distances between department locations

Departments	Average processing time
0	18
1	18
2	6
3	6
4	18
5	18
6	18
7	18
8	6
9	6
10	18
11	18

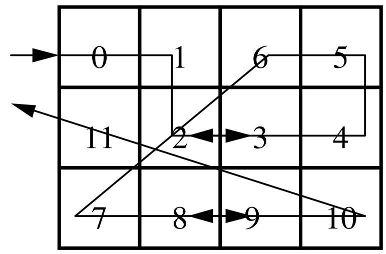
(c) Department average processing times

an increase in expected WIP can then follow. We illustrate this result using the following example. Consider a system consisting of 12 locations and 12 departments arranged in a 3×4 grid as shown in Figure 2(a). Departments are always visited by all products in the following sequence: $0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 9 \rightarrow 8 \rightarrow 9 \rightarrow 8 \rightarrow 9 \rightarrow 10 \rightarrow 11$. The distance matrix between locations

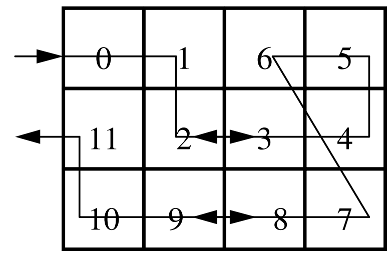
Figure 3 **Example Layouts**



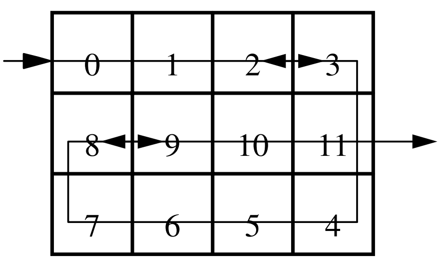
(a) Layout $\mathbf{x_1}$
$\rho_t = 0.990$, $\rho_f = 0.311$, $\rho_e = 0.679$,
$E(\text{WIP}_t) = 58.37$, $E(\text{WIP}) = 69.53$



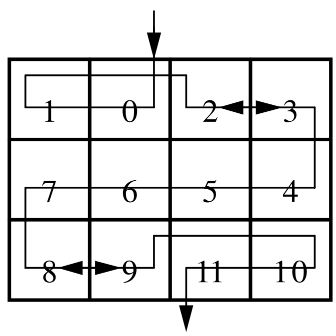
(b) Layout $\mathbf{x_2}$
$\rho_t = 0.951$, $\rho_f = 0.409$, $\rho_e = 0.542$,
$E(\text{WIP}_t) = 11.76$, $E(\text{WIP}) = 22.95$



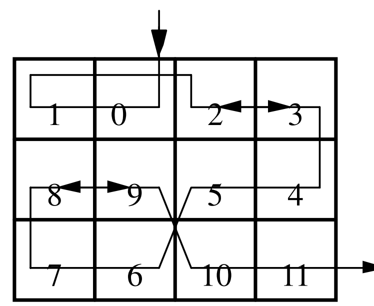
(c) Layout $\mathbf{x_3}$
$\rho_t = 0.885$, $\rho_f = 0.344$, $\rho_e = 0.542$,
$E(\text{WIP}_t) = 4.70$, $E(\text{WIP}) = 15.90$



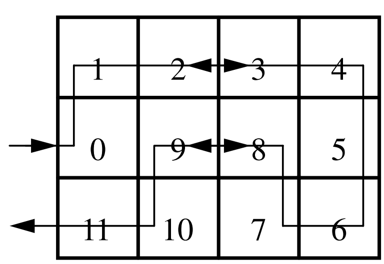
(d) Layout $\mathbf{x_4}$
$\rho_t = 0.961$, $\rho_f = 0.327$, $\rho_e = 0.634$,
$E(\text{WIP}_t) = 14.09$, $E(\text{WIP}) = 25.25$



(e) Layout $\mathbf{x_5}$
$\rho_t = 1.04$, $\rho_f = 0.344$, $\rho_e = 0.695$,
$E(\text{WIP}_t) = \infty$, $E(\text{WIP}) = \infty$



(f) Layout $\mathbf{x_6}$
$\rho_t = 0.992$, $\rho_f = 0.360$, $\rho_e = 0.632$,
$E(\text{WIP}_t) = 69.20$, $E(\text{WIP}) = 80.36$



(g) Layout $\mathbf{x_7}$
$\rho_t = 0.853$, $\rho_f = 0.311$, $\rho_e = 0.542$,
$E(\text{WIP}_t) = 5.56$, $E(\text{WIP}) = 14.77$

is shown in Figure 2(b)—we assume rectilinear distances with unit distance separating adjacent locations. Average processing times at departments are shown in Figure 2(c). We consider a system with a single material-handling device with speed of 1.65 (units of distance per unit of time) and overall demand rate of 0.027 (unit loads per unit time). Let us consider the two layouts shown in Figures 3(a) and 3(b), denoted respectively by $\mathbf{x_1}$ and $\mathbf{x_2}$ (the arrows are used to indicate the direction of material flow). It is easy to verify that layout $\mathbf{x_1}$ is *QAP-optimal* and minimizes full travel. In contrast, layout $\mathbf{x_2}$ is not QAP-optimal and, in fact, appears to be quite inefficient. Expected material-handling system WIP for layout $\mathbf{x_1}$ and $\mathbf{x_2}$, as well as the corresponding full and empty material-handling system utilizations, are shown in Figures 3(a) and 3(b). We can see that although layout $\mathbf{x}_2$ does not minimize full travel, it results in significantly less empty travel, which is sufficient to cause an overall reduction in material-handling system utilization. As a consequence, expected WIP for layout $\mathbf{x}_2$ is smaller than that of $\mathbf{x}_1$. In fact, material-handling system WIP is reduced by nearly 80% (from 58.37 to 11.76) when layout $\mathbf{x}_2$ is chosen over $\mathbf{x}_1$!

This surprising result stems from the fact that the frequency with which a device makes empty trips to a particular department is proportional to the volume of outflow from that department. The likelihood of the material-handling device being in a particular department is similarly proportional to the volume of inflow to that department. Therefore, if two departments are highly loaded, the number of empty trips between them would be large even if no direct flow exists between these departments. In our example, Departments 2, 3, 9, and 8 have three times the workload of any other department in the factory. Therefore, the likelihood of an empty trip between any two of the four departments is three times higher than between any other two departments. In layout $\mathbf{x}_2$, by placing these four departments in neighboring locations, empty travel is significantly reduced. Note that this is realized despite the fact that there is no direct material flow between the department pairs 2–3 and 8–9.

The above result also leads us to the following more general observation, which further highlights the fact that full travel is a poor indicator of WIP.

Observation 2. Expected WIP is affected by both ρ_f and ρ_e, which are not correlated and whose effect on expected WIP is not monotonic.

Observation 2 follows from the fact that an increase in $\rho_f(\rho_e)$ can result in either an increase or a decrease in $\rho_e(\rho_f)$. Depending on how $\rho_e(\rho_f)$ is affected, expected WIP may either increase or decrease. We illustrate this behavior by considering a series of layout configurations based on our previous example. The layouts, denoted $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_7$, are shown in Figure 3. The behavior of ρ_f, ρ_e, ρ_t, and $E(\text{WIP}_t)$ is graphically depicted in Figure 4. It is easy to see that ρ_f can behave quite differently from ρ_e and ρ_t. It is also easy to see that an increase or a decrease in ρ_f does not always have predictable consequences on expected WIP.

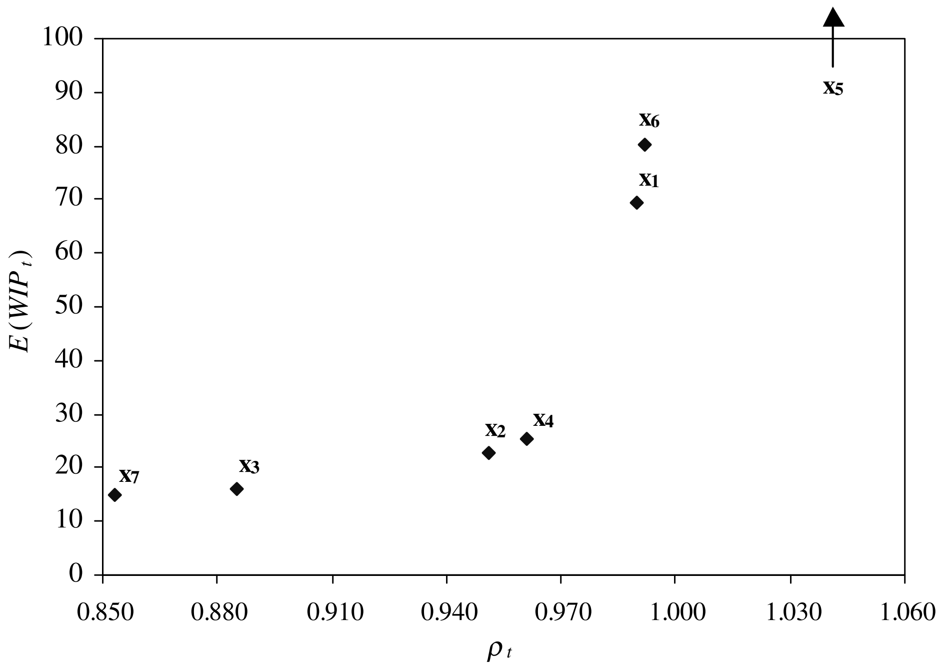
The fact that ρ_f can behave differently from ρ_e means that it is possible to have layouts with similar values of ρ_f but different values of ρ_e. This also means that layouts could have the same value of ρ_f but different values of expected WIP. In fact, it is possible to have two QAP-optimal layouts with very different WIP values. It is also possible for a layout to be QAP-optimal (i.e., it has the smallest value of ρ_f) and be WIP-infeasible.

Observation 3. Two QAP-optimal layouts can have different WIP values. Given a fixed material-handling capacity, a QAP-optimal layout can be WIP-infeasible for a system where there are one or more WIP-feasible layouts.

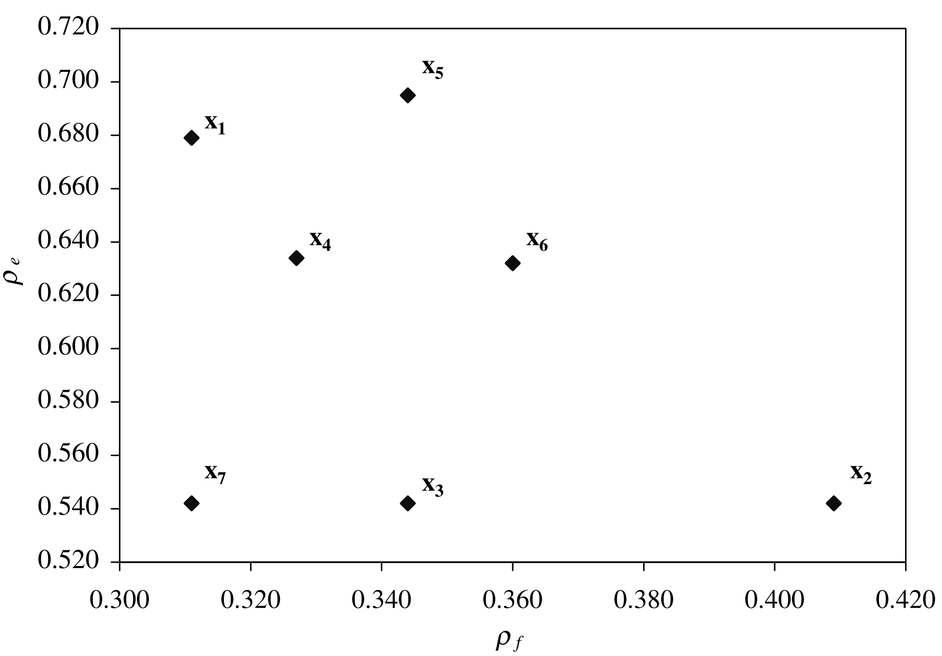
The first part of the result follows from noting that two layouts can have the same ρ_f but different values of ρ_e. For example, consider the two layouts, $\mathbf{x}_1$ and $\mathbf{x}_7$, shown in Figure 3. Both layouts are QAP-optimal. However, $E(\text{WIP}_t|\mathbf{x}_1) = 58.37$ and $E(\text{WIP}_t|\mathbf{x}_7) = 5.56$! The above result shows that QAP-optimality can be a poor indicator of WIP performance. The second part of the result is due to the fact that, even though ρ_f might be minimal, the corresponding ρ_e can be sufficiently large to make ρ_t greater than 1. We illustrate this result using the following example. Consider the same system description we used for the previous three observations except that material-handling system speed is 1.6 instead of 1.65. Now consider the performance of the layout configurations $\mathbf{x}_1$ and $\mathbf{x}_3$ shown in Figure 3. We have $E(\text{WIP}_t|\mathbf{x}_1) = \infty$ while $E(\text{WIP}_t|\mathbf{x}_3) = 10.5$. Thus, although layout $\mathbf{x}_1$ is QAP-optimal, it is infeasible. Layout $\mathbf{x}_3$ is not QAP-optimal but produces a relatively small WIP. Clearly, QAP-optimality does not guarantee feasibility. In systems where material-handling capacity is not a constraint, these results mean that a implementing a QAP-optimal layout would require a greater investment in material-handling capacity.

The previous three observations show that the QAP objective function can be a poor predictor of WIP. Therefore, there is a need to explicitly evaluate WIP if our objective is to design layouts that minimize it. In fact, regardless of the objective function, there is

Figure 4 **The Effect of Layout Configuration on Utilization and WIP**



(a) The effect of layout configuration on WIP



(b) The effect of layout configuration on material handling utilization always a need to at least evaluate both empty and full travel by the material-handling system since we must always generate feasible layouts. The fact that empty travel can be a significant portion of material-handling system utilization also means that we need to design layouts that minimize it. This may sometimes result in going counter to the common practice of favoring the placement of departments with large intermaterial flows in neighboring locations. As we saw in the previous examples, reducing WIP could lead to departments being placed in adjacent locations although there is no direct material flow between them (e.g., the department pairs 2–3 and 9–8). Because empty travel is more frequent from and to departments that are popular destinations (i.e., departments with high flow rates), placing these departments in neighboring locations can significantly reduce empty travel even when there is no direct flow between these departments. Therefore, the need to reduce full travel by placing departments with large intermaterial flows in neighboring locations must be balanced by the need to reduce empty travel by placing departments that are popular destinations in close proximity. In short, there is a need to always account for both full and empty travel since together they affect the utilization of the material-handling system, which in turn affects WIP accumulation.

4.2. The Effect of Variability

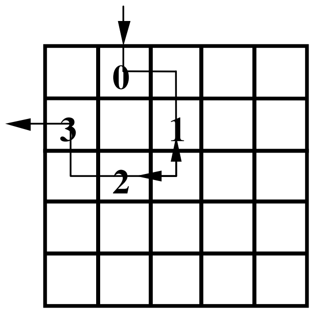
Examining the expression of expected WIP, we can see that in addition to utilization of the material-handling system, WIP accumulation is determined by (1) the variability in the arrival process, (2) the variability in the processing/transportation times, and (3) the utilization of the departments. We can also see that because the material-handling system provides input to all the processing departments, variability in transportation time, as well as the material-handling system utilization, directly affect the variability in the arrival process to all the departments. In turn, this variability, along with the variability of the department processing times and the department utilizations, determine the input variability to the material-handling system. Because of this close coupling, the variability of any resource and its utilization affect the WIP at all other resources. This effect is not captured by the exponential model of Fu and Kaku and can lead to very different results with regard to layout WIP performance.

From the examples of the previous section, expected WIP, although not monotonic in full travel utilization, appears to be monotonic in overall material-handling system utilization. We show that this is not always true. In fact, we show that reducing average travel time (i.e., reducing ρ_t) can increase WIP. As a result, increasing the average distance between departments could, in fact, reduce WIP. Moreover, we show that the relative desirability of a layout can be highly sensitive to changes in material-handling capacity even when travel distances are the same. We also find that WIP accumulation at the material-handling system can be affected by non-material-handling factors, such as the utilization of the processing departments or variability in the department processing time, which means that the relative desirability of two layouts could be affected by these factors.

Observation 4. A smaller average travel time (full + empty) does not always lead to a smaller expected WIP.

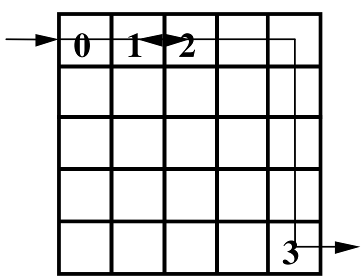
The proof of Observation 4 follows by noting that the expression of expected WIP is a function of both ρ_t and $C_{s_t}^2$. Since $C_{s_t}^2$ is not necessarily decreasing in ρ_t, a reduction in ρ_t may indeed cause an increase in $C_{s_t}^2$, which could be sufficient to either increase material-handling WIP or increase the arrival variability at the processing departments, which in turn could increase their WIP. We illustrate this behavior using the following example. Consider a facility with four departments ($i = 0, 1, 2,$ and 3). Products in the facility are always manufactured in the following sequence: $0 \to 1 \to 2 \to 1 \to 2 \to 1 \to 2 \to 3$. Other relevant data is as follows: $D_1 = 0.027$; $E(S_0) = E(S_3) = 30$ and $E(S_1) = E(S_2) = 10$, $C_{a_0}^2 = 1.0$, $C_{s_i}^2 = 0.5$ for $i = 0, 1, \ldots, 3$, $n_t = 1$ and $\nu = 0.68$. We consider two layout scenarios, $\mathbf{x}_8$ and $\mathbf{x}_9$. The distances between departments are as follows, layout $\mathbf{x}_8 : d_{01}(\mathbf{x}_8) = d_{02}(\mathbf{x}_8) = d_{03}(\mathbf{x}_8) = d_{12}(\mathbf{x}_8) = d_{13}(\mathbf{x}_8) = d_{23}(\mathbf{x}_8) = 2$; and layout $2 : d_{01}(\mathbf{x}_9) = 1$, $d_{02}(\mathbf{x}_9) = 2$, $d_{03}(\mathbf{x}_9) = 8$, $d_{12}(\mathbf{x}_9) = 1$, $d_{13}(\mathbf{x}_9) = 7$, and $d_{23}(\mathbf{x}_9) = 6$. The two layouts are graphically depicted in Figure 5. Since $E(\mathrm{WIP}(\mathbf{x}_8)) = 15.87 < E(\mathrm{WIP}(\mathbf{x}_9)) = 19.26$ although $\rho_t(\mathbf{x}_8) = 0.907 > \rho_t(\mathbf{x}_9) = 0.896$, our

Figure 5 **Example Layouts for Observation 4**



Layout $\mathbf{x_8}$

$\rho_t = 0.907$, $\rho_f = 0.555$, $\rho_e = 0.352$, $C^2_{s_t} = 0.087$,
$C^2_{a_t} = 0.567$, $C^2_{a_1} = C^2_{a_2} = 0.645$, $C^2_{a_3} = 0.882$,
$E(\text{WIP}_t) = 3.78$, $E(\text{WIP}) = 15.87$



Layout $\mathbf{x_9}$

$\rho_t = 0.896$, $\rho_f = 0.476$, $\rho_e = 0.420$, $C^2_{s_t} = 0.735$,
$C^2_{a_t} = 0.636$, $C^2_{a_1} = C^2_{a_2} = 0.878$, $C^2_{a_3} = 0.959$,
$E(\text{WIP}_t) = 6.16$, $E(\text{WIP}) = 19.26$

result is proven. As indicated in Figure 5, the difference in WIP between the two layouts is mostly due to the higher value of $C^2_{s_t}$ in the case of $\mathbf{x_9}$ ($C^2_{s_t}(\mathbf{x_9}) = 0.735$ vs. $C^2_{s_t}(x_8) = 0.087$). In turn, this leads to higher values of $C^2_{a_i}$ and $C^2_{a_t}$ which further contribute to the larger WIP in layout $\mathbf{x_9}$.

The above results show the important effect that variability in travel times can play in determining overall WIP. In each of the above examples, the smaller value of average travel time is associated with higher travel-time variability. This higher variability causes not only an increase in material-handling WIP, but also in department WIP (by increasing variability in the arrival process to the departments). These results point to the need for explicitly accounting for travel-time variance when selecting a layout. A layout that exhibits a small variance may, indeed, be more desirable than one with a smaller travel-time average. In practice, travel-time variance is often dictated by the material-handling system configuration. Therefore, special attention should be devoted to identifying configurations that minimize not only average travel time, but also its variance. For example, the star-layout configuration shown in Figure 6(a) has a significantly smaller variance than the loop layout of 5(b), which itself has a smaller variance than the linear layout of 5(c).

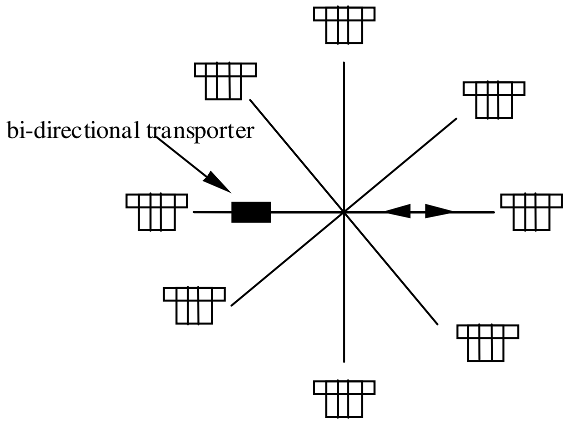
Although in the above examples the layout with the lower variance is more desirable, we should caution that this relative desirability can be sensitive to the available material-handling capacity. For example, from the stability condition ($\rho_t < 1$), we can see that the minimum feasible material-handling speed is higher for layout $\mathbf{x_8}$ than for layout $\mathbf{x_9}$. This means that for certain material-handling speeds layout $\mathbf{x_8}$ is, indeed, infeasible while layout $\mathbf{x_9}$ still results in finite WIP. More generally, as shown in Figure 7, the relative ranking of layouts can be affected by changes in material-handling capacity. For example, layout $\mathbf{x_8}$ is superior to layout $\mathbf{x_9}$ when material-handling speed is greater than 0.61, but it is clearly inferior for lower speeds. These results lead to the following important observation.

Observation 5. The relative ranking of layouts based on expected WIP can change with a change in material-handling capacity.

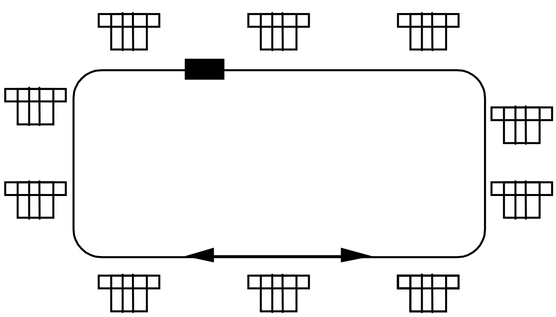
Observation 5 highlights the fact that material-handling capacity can have an unpredictable impact on layout desirability. It also points to the complex relationship between distribution of travel time, material-handling capacity, and WIP performance.

Travel distances and material-handling capacity are not, however, the only factors that affect the relative desirability of a layout. Non-material-handling factors such as department utilization levels, variability in department processing times, and variability in demand levels could determine whether one layout configuration is more desirable than another. For example, in the following observation we show that variability in processing times and demand can affect the relative ranking of a layout with regard to expected WIP.

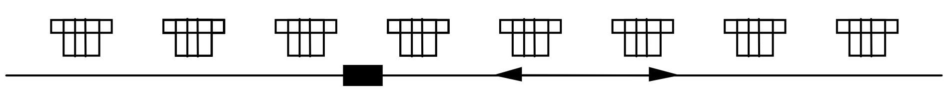
Figure 6 Star, Loop, and Linear Layouts



(a) Star layout



(a) Loop layout



(a) Linear layout

Observation 6. The relative ranking of a layout based on expected WIP can be affected by non-material-handling factors.

Since the arrival variability to the processing departments and the material-handling system is affected by the utilization of the processing departments, the processing time variability, and the variability in product demands, it is possible that changes in these parameters could affect the relative desirability of a particular layout. We illustrate this behavior by considering layouts—the two layouts shown in Figure 8—with similar parameters to those in the previous observation (in this case, we let material-handling speed be 0.7). In Table 1, we show the effect of processing time and demand variability on the performance of the two layouts. As we can see, the same layout can be superior under one set of parameters and inferior under another.

Since the results of the observations are based on approximations for both average WIP and the arrival processes to the various departments, we used computer simulation to confirm them. For each of the example layouts, we constructed a stochastic simulation model using the discrete event simulation language Arena (Kelton et al. 1998). The simulated models are identical to the analytical ones, except

Figure 7 The Effect of Material-Handling Capacity on WIP Performance

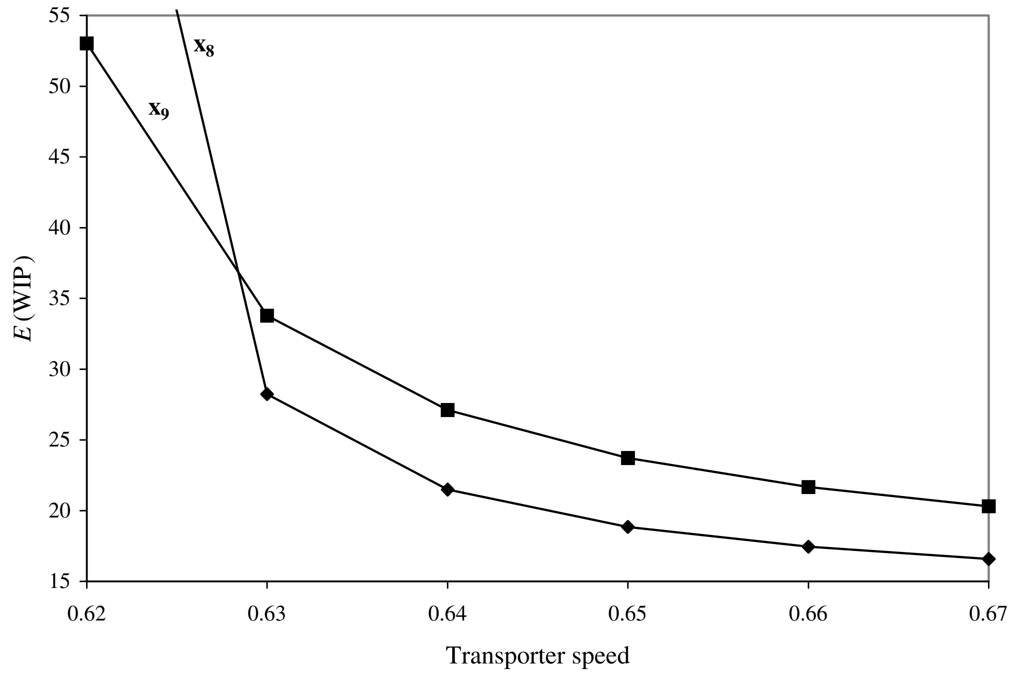
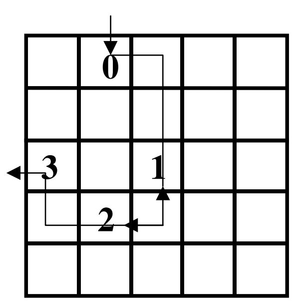
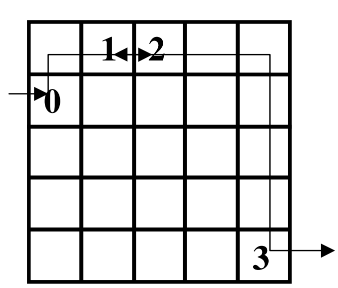


Figure 8 Example Layouts for Observation 7



Layout $\mathbf{x}_{10}$



Layout $\mathbf{x}_{11}$

that the travel time distribution is not prespecified. Instead, we provide the simulation model with the distances between departments, material-handling speed, and product routings. In contrast with the analytical approximations, the simulation model does away with the probabilistic routing assumption and captures dependencies between the length of consecutive trips that tend to occur in real systems (e.g., a long trip that takes the material-handling device to the outer edges of the layout tends to be followed by another long trip). For each case, we collected statistics on average WIP at the different processing departments and material-handling system. For each case, we also obtained a 95% confidence interval with a maximum half-width of 0.01. In addition to these

Table 1 The Effect of Variability on Layout Performance

Variability	$E(\mathrm{WIP}(\mathbf{x}_{10}))$	$E(\mathrm{WIP}(\mathbf{x}_{11}))$
$C_{a_0}^2 = 0.2$, $C_{s_i}^2 = 0.2$	14.96	17.54
$C_{a_0}^2 = 0.5$, $C_{s_i}^2 = 0.5$	19.96	21.66
$C_{a_0}^2 = 1.0$, $C_{s_i}^2 = 1.0$	28.3	28.54
$C_{a_0}^2 = 1.2$, $C_{s_i}^2 = 1.2$	31.63	31.29
$C_{a_0}^2 = 2.0$, $C_{s_i}^2 = 2.0$	44.98	42.29

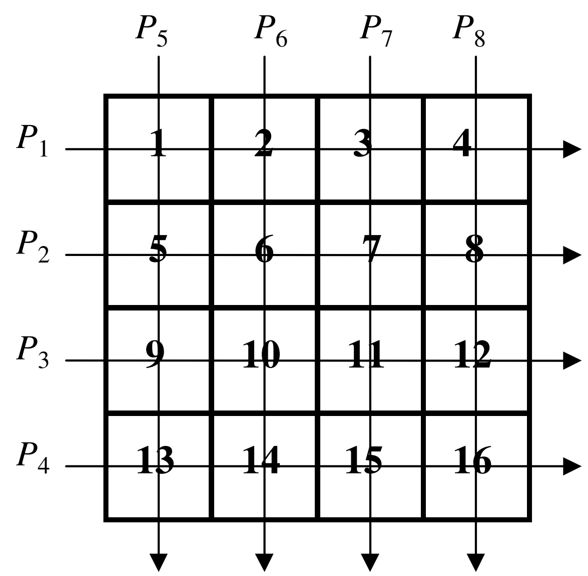
specific examples, we also simulated examples with randomly generated data sets and compared them to the approximation results. For brevity, the results are not included but are available from the author upon request.

Although specific values of the approximated average WIP are not always within the simulation 95% confidence interval, the simulated results confirm each of the observations (in each case, the relative ranking of the simulated layouts is consistent with the one obtained analytically; also, in each case, the differences between ranked layouts are found to be statistically significant). In general, we found the inaccuracy in estimating overall WIP to be mostly due to inaccuracies in estimating the variability in the arrival process to the departments and the material-handling system and variability in travel times. This is especially significant when both demand and processing time variability are small. In this case, variability is overestimated, which in turn results in higher estimates of WIP. This effect is due to the probabilistic approximation used in determining the origin of material-handling requests. This limitation can be addressed in part by extending the queuing network model to account for multiproduct deterministic job routings.

5. When Does Minimizing WIP Matter?

We have so far highlighted instances where the WIP formulation leads to a different layout from the one obtained using the QAP formulation. In this section, we examine factors that affect the degree to which the layouts obtained from the two formulations are different. In particular, we highlight conditions under which there is little difference in WIP between the two formulations or those under which the two formulations are actually *WIP-equivalent*. We consider six factors that we found through numerical experimentation to affect congestion the most. These include flow asymmetry, dimensional asymmetry, material-handling capacity, number of transporters, and variability in demand and processing time. To illustrate the effect of these factors, we carried out a

Figure 9 **The QAP-Optimal Layout**



full factorial design-of-experiments based on a system consisting of 16 departments, 16 locations and eight products. Routing sequences for each product are as follows: P_1: $1 \to 2 \to 3 \to 4$; P_2: $5 \to 6 \to 7 \to 8$; P_3: $9 \to 10 \to 11 \to 12$; P_4: $13 \to 14 \to 15 \to 16$; P_5: $1 \to 5 \to 9 \to 13$; P_6: $2 \to 6 \to 10 \to 14$; P_7: $3 \to 7 \to 11 \to 15$; and P_8: $4 \to 8 \to 12 \to 16$ (product flow is illustrated in Figure 9 for a QAP-optimal layout of a system with a 4×4 geometry). For each factor, we consider a set of values over a sufficiently wide range. Because of the nonlinear behavior of expected WIP in some factors, the number of levels considered varies per factor. However, for all factors, a minimum of three levels is evaluated. In total, we carried out over 2,000 experiments. For each experiment, we obtain both a QAP- and a WIP-optimized layout and the corresponding expected WIP values, which we denote respectively by $E(\text{WIP}|\text{QAP})$ and $E(\text{WIP}|\text{WIP})$. We use the ratio $\delta = E(\text{WIP}_t|\text{QAP})/E(\text{WIP}_t|\text{WIP})$ to measure the relative difference in material-handling WIP between the two layouts. The WIP due to the processing departments also varies. However, we found the effect of layout on this WIP, except for extreme cases, to be relatively small.

Table 2 Demand Scenarios for Example System

Scenario	1	2	3	4	5	6	7	8	9	10
D_1	40	40	58	58	76	112	148	202	256	292
D_2	40	40	22	22	4	4	4	4	4	4
D_3	40	40	22	22	4	4	4	4	4	4
D_4	40	40	58	58	76	112	148	94	40	4
D_5	40	58	58	76	76	40	4	4	4	4
D_6	40	22	22	4	4	4	4	4	4	4
D_7	40	22	22	4	4	4	4	4	4	4
D_8	40	58	58	76	76	40	4	4	4	4

The following describes how the value of each factor is varied.

Flow Asymmetry. Flow asymmetry refers to the imbalance in flow rates among departments. In an asymmetric system, some departments are more visited than others, leading to more empty trips ending and originating at these departments. We consider 10 demand scenarios corresponding to 10 different flow asymmetry scenarios. The demand scenarios and the associated flow rates for each department are shown in Tables 2 and 3, respectively. We use the standard deviation in the department flow rates, σ_f, to measure asymmetry under each scenario. To allow for a fair comparison between different scenarios, the capacity of a department is adjusted proportionally to its workload in order to maintain a constant utilization per department.

Dimensional Asymmetry. Dimensional asymmetry refers to asymmetry in the distances between different department locations (in a perfectly symmetric system, all department locations are equidistant from each other). We consider three levels of asymmetry corresponding to layouts with a 4×4, 2×8, and 1×16 geometry. Of the three geometries, the 1×16 is clearly the most asymmetric and the 4×4 is the least. In each case, we assume unit distance between adjacent locations and rectilinear travel between each pair of locations.

Table 3 Department Arrival Rates per Scenario for Example System

Scenario	1	2	3	4	5	6	7	8	9	10
λ_1	80	98	116	134	152	152	152	206	260	296
λ_2	80	62	80	62	80	116	152	206	260	296
λ_3	80	62	80	62	80	116	152	206	260	296
λ_4	80	98	116	134	152	152	152	206	260	296
λ_5	80	98	80	98	80	44	8	8	8	8
λ_6	80	62	44	26	8	8	8	8	8	8
λ_7	80	62	44	26	8	8	8	8	8	8
λ_8	80	98	80	98	80	44	8	8	8	8
λ_9	80	98	80	98	80	44	8	8	8	8
λ_{10}	80	62	44	26	8	8	8	8	8	8
λ_{11}	80	62	44	26	8	8	8	8	8	8
λ_{12}	80	98	80	98	80	44	8	8	8	8
λ_{13}	80	98	116	134	152	152	152	98	44	8
λ_{14}	80	62	80	62	80	116	152	98	44	8
λ_{15}	80	62	80	62	80	116	152	98	44	8
λ_{16}	80	98	116	134	152	152	152	98	44	8
λ_t	1280	1280	1280	1280	1280	1280	1280	1280	1280	1280
σ_f	0	19	26	42	53	59	74	84	108	129

Material-Handling Capacity. For a fixed number of transporters, material-handling capacity is determined by transporter speed. For each number of transporters, we consider six values of transporter speed that correspond to six different levels of transporter utilization (under a QAP-optimized layout) ranging from 0.6 to 0.99.

Number of Transporters. We consider systems with a number of transporters ranging from one to five. To allow for a fair comparison between systems with different numbers of transporters, we always maintain the same overall material-handling capacity by adjusting transporter speed proportionally to the number of transporters. This allows us also to distinguish between the effect of capacity and that of multiplicity of transporters.

Demand and Processing Time Variability. Six levels of demand variability are considered by varying the squared coefficients of variation of part external interarrival times from 0.33 to 2. Similarly, five levels of processing time variability are considered by varying the squared coefficients of variation of department processing times from 0.33 to 2.

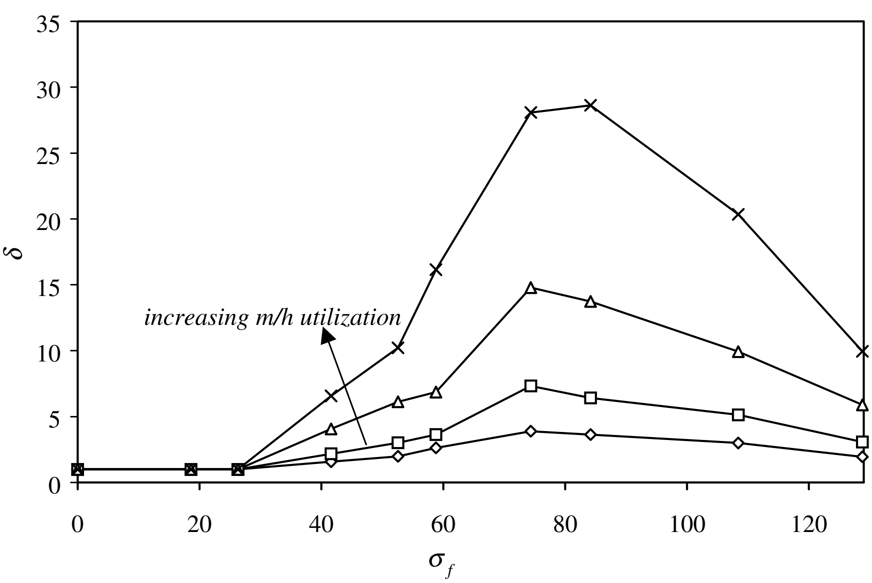
In the following sections, we summarize key results by describing the effect of each factor on the ratio δ. When appropriate, we also comment on interactions between different factors. For brevity, we show only a subset of the data we generated. The selected data are in all cases illustrative of the effects observed in the larger set.

5.1. The Effect of Flow Asymmetry

First, let us note that in a symmetric system where department flow rates are equal, empty travel is *layout independent* since there is equal likelihood of an empty trip originating at any department and ending at any other department. Consequently, the difference in expected WIP between a QAP-optimal and a WIP-optimal layout is always zero. However, a difference emerges, as we saw in previous sections, when some departments are visited more frequently than others. The effect of increasing flow asymmetry on this difference is illustrated in Figure 10.

From Figure 10, we see that while the WIP ratio δ is in the neighborhood of one when σ_f is relatively small, it can be significantly higher when σ_f is large. Surprisingly, the effect of σ_f is not monotonic. Although initial increases in σ_f do lead to a larger δ, additional increases invariably reduce its value. Thus, δ is maximum when σ_f is in the midrange and is significantly smaller in the extreme cases of either

Figure 10 **The Effect of Flow Asymmetry (4×4 Geometry; $n_t = 1$; $C_i = 1$ for $i = 1, \ldots, 8$; $\rho_j = 0.8$; $C_{s_j} = 1$ for $j = 1, \ldots, 16$; the Series Corresponds to ρ_t Values of 0.9, 0.95, 98, and 0.989)**



high or low asymmetry. A possible explanation for this nonmonotonic behavior is as follows. In a highly asymmetric system, the demand from one product dominates the demand from all others. Hence, the departments that are most visited are those that are visited by the product with the highest demand. Because these departments are already in neighboring locations under the QAP-optimal layout, the additional reduction in empty travel due to using the WIP criterion is limited. This is in contrast to situations where asymmetry is due to two or more products having relatively higher demands than the others. In that case, rearranging the layout so that the departments visited by these products are in neighboring locations does significantly reduce empty travel. The above results are summarized in the following observation.

Observation 7. The percentage difference in expected WIP between a QAP-optimal and a WIP-optimal layout is not monotonic in flow asymmetry. It is relatively small for either highly symmetric or asymmetric systems. However, it can be significant when flow asymmetry is in the midrange.

5.2. The Effect of Dimensional Asymmetry

In a perfectly dimension-symmetric system, all department locations are equidistant from each other. In this case, full and empty travel are always the same regardless of department placement. Hence, a QAP-optimal layout (or any other layout) is also WIP optimal. In other words, in a dimension-symmetric system a QAP-optimal and a WIP-optimal layout are WIP equivalent. Although it is difficult to predict in general the impact of an increase in dimensional asymmetry on the WIP ratio—this would largely depend on the specific geometry of the layout and the distribution of the flow among departments—large increases in asymmetry tend to increase the difference in WIP between a QAP-optimal layout and a WIP-optimal layout. This is supported by the results obtained for the three geometries we consider. A representative data set is shown in Figure 11.

From Figure 11, we see that the WIP ratio is largest for the most asymmetric system and smallest for the most symmetric one. The effect of dimensional asymmetry is sensitive to flow asymmetry. However, the effect of dimensional asymmetry is not always monotonic. Depending on the distribution of flow among departments, it is possible to see a reduction in the WIP ratio if an increase in dimensional asymmetry yields (unintentionally) a QAP solution where departments with the most flow are in neighboring locations. This tends to occur less frequently when both

Figure 11 The Effect of Layout Geometry ($n_t = 1$; $C_i = 1$ for $i = 1, \ldots, 8$; $\rho_j = 0.8$; $C_{s_j} = 1$ for $j = 1, \ldots, 16$)

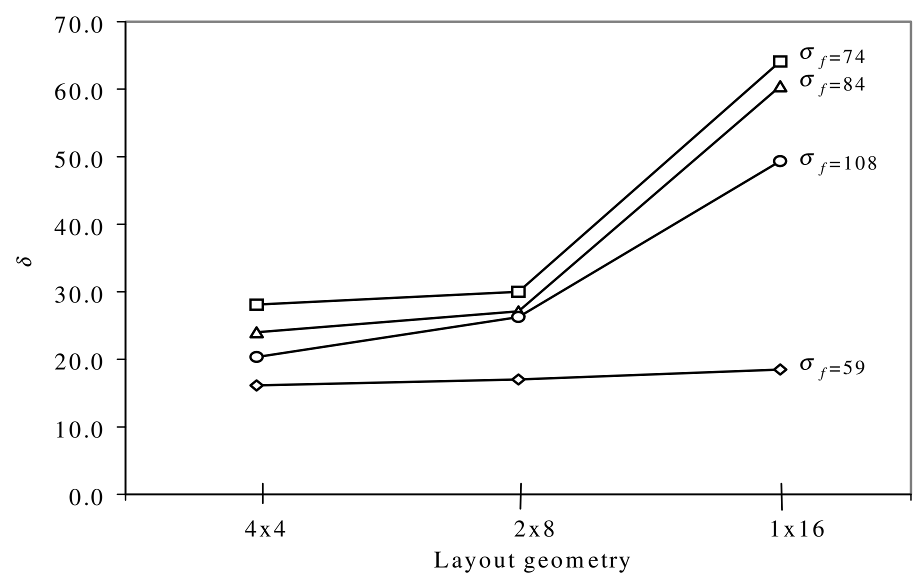
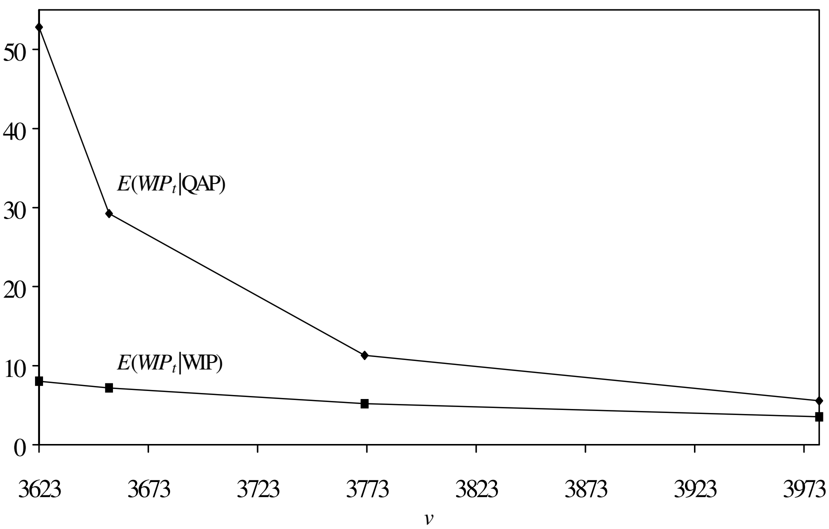


Figure 12 The Effect of Material Handling Capacity ($\sigma_f = 42$; 4×4 Geometry; $n_t = 1$; $C_i = 1$ for $i = 1, \ldots, 8$; $\rho_j = 0.8$; $C_{s_j} = 1$ for $j = 1, \ldots, 16$)



dimensional and flow asymmetry are high. In these cases, the QAP formulation does not usually favor placing departments in neighboring locations unless they have direct flows between them.

Observation 8. The percentage difference in expected WIP between a QAP-optimal and a WIP-optimal layout is generally increasing in dimensional asymmetry. In a dimension-symmetric system, a QAP-optimal and a WIP-optimal layout are WIP equivalent.

5.3. The Effect of Material-Handling Capacity

The effect of increasing transporter speed is illustrated in Figure 12. We see that while for small values of ν, the difference in WIP is significant, the two layouts are practically WIP-equivalent when ν is large. This result is observed to hold regardless of flow asymmetry or layout geometry. Noting that higher values of ν correspond to lower utilization of the material-handling system, these results are, however, not surprising. A WIP-optimal layout would generally result in a smaller fraction of utilization devoted to empty travel. This reduction in empty travel is of little consequence when there is excess material-handling capacity. However, it becomes crucial when material-handling capacity is tight. In fact, given that WIP grows exponentially in the utilization of the material-handling system, even small decreases in empty travel would have a dramatic impact on WIP accumulation when utilization is high (as utilization approaches 1, δ grows without bound).

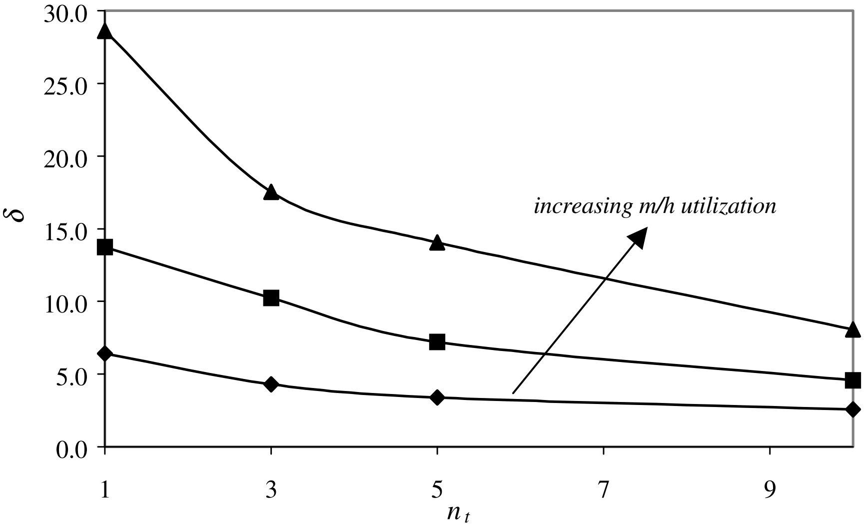
Observation 9. The difference in WIP between a QAP-optimal and a WIP-optimal layout is generally increasing in the utilization of the material-handling system.

Although the above result is generally true, there are instances when an increase in utilization could affect the variability of travel times sufficiently to cause a decrease in the WIP ratio. This does not occur often because the effect of utilization tends to dominate the effect of variability, particularly when material-handling system utilization is high.

5.4. The Effect of the Number of Transporters

The effect of increasing the number of transporters while maintaining the same overall material-handling capacity is illustrated in Figure 13. As we can see, the ratio δ is decreasing in n_t with δ approaching one when n_t is large. This means that the percentage difference between the two layouts is less significant in a system with many slow transporters than one with few fast ones (recall that ν decreases with an increase in n_t). This may seem surprising since expected WIP in both layouts is actually increasing in n_t (a single fast transporter is superior in terms of WIP to multiple slower ones). The fact that δ is decreasing in n_t

Figure 13 **The Effect of Number of Transporters ($\sigma_f = 84$; 4×4 Geometry; $C_i = 1$ for $i = 1, \ldots, 8$; $\rho_j = 0.8$; $C_{s_j} = 1$ for $j = 1, \ldots, 16$; the Series Corresponds to ρ_t Values of 0.95, 98, and 0.989)**



appears to be due mostly to how a change in utilization affects WIP for different values of n_t. For small values of n_t a drop in utilization (when it is initially high) can cause a greater decrease in WIP than the one seen when n_t is large. These results are in line with known queueing effects in multiserver systems—see, for example, Kleinrock (1976, pp. 279–285). We should note that although the ratio δ approaches 1 when n_t is large, the difference in WIP can remain significant. In fact, in many cases we observed the difference remains relatively constant in n_t.

5.5. The Effect of Demand and Processing Time Variability

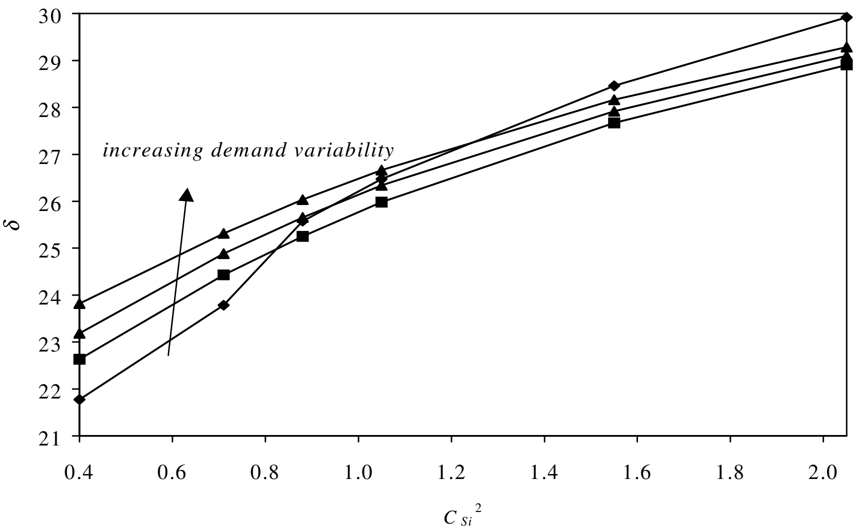
Varying the variability in either demand or processing times affects transporter WIP by affecting C_{a_t}. This effect can be gleaned from the expression of C_{a_t} in Equation (26). The value of C_{a_t} is linearly increasing in both C_{s_i} and C_{a_0}, with the rate of increase an (increasing) function of ρ_t. Hence, we should expect δ to be increasing in both C_{s_i} and C_{a_0}. This is supported by the numerical results, a sample of which is shown in Figure 14. Note that the effect of process variability is more pronounced than that of demand, since in our experiments process variability is increased uniformly for all the departments.

5.6. Managerial Implications

Table 4 provides a summary of our results and offers broad guidelines as to when using WIP as a design criterion is particularly valuable. The results suggest that a WIP-optimized layout is most beneficial when material-handling capacity is limited, dimensional asymmetry is high, there is asymmetry in the flows, the number of transporters is small, or variability in either demand or processing times is high. These results also point to strategies that managers and facility planners could pursue to increase the robustness of layouts with respect to WIP performance—i.e., investing in excess material-handling capacity, adopting layout geometries that reduce travel distance variance, and ensuring that the most visited processes are centrally located.

Furthermore, these results draw attention to the importance of indirect interactions that take place between different areas of a facility. These indirect effects have implications for the way we should organize areas of a facility that may otherwise appear independent. For example, consider a system consisting of multiple cells that do not share any products or processes but are serviced by the same material-handling system. Our results suggest that among these cells those that manufacture the products with

Figure 14 The Effect of Demand and Process Variability ($\sigma_f = 84$; 4×4 Geometry; $n_t = 1$, $\nu = 3436$, $\rho_j = 0.8$ for $j = 1, \ldots, 16$; the Series Corresponds to C_i Values of 0.35, 1, 1.5, and 2)



the highest demand should be placed in neighboring locations (although they do not share any flows). Our results also show that organizing these cells into parallel production lines, a common practice in many facilities, may lead to greater congestion and longer lead times. Instead, adopting a configuration that minimizes distance asymmetries (e.g., using a layout where cells are configured into a U-shape and are arranged along a common corridor where most travel would take place) would maintain the efficient transfer of material within cells while freeing up additional material-handling capacity to service the entire facility.

Many companies are beginning to realize the importance of these indirect effects and are increasingly designing layouts that minimize dimensional asymmetries and reduce empty travel. For example, GM built its new Cadillac plant in the form of a T to maximize supplier access to the factory floor and reduce the distance between loading docks and production stocking points (Green 2000). Volvo designed its Kalmar plant as a collection of hexagon-shaped modules where material flows in concentric lines within each module (Tompkins et al. 1996). Motorola is experimenting with layouts where shared processors are centrally located in functional departments and are equidistant from multiple dedicated cells within the plant. Variations of the *spine* layout, where departments are placed along the sides of a common corridor, have been successfully implemented in industries ranging from electronic manufacturing to automotive assembly (Smith et al. 2000, Tanchoco 1994, Tompkins et al. 1996). Layout configurations that minimize dimensional asymmetries and reduce empty travel are also found in nonmanufacturing applications. For example, both the spine and star layouts are common configurations in airport designs.

Table 4 When Is Using the WIP Criterion Valuable?

	Low	Medium	High
Flow asymmetry	Less valuable	More valuable	Moderately valuable
Dimensional asymmetry	Less valuable	Moderately valuable	More valuable
Material-handling capacity	More valuable	Moderately valuable	Less valuable
Number of transporters	More valuable	Moderately valuable	Less valuable
Demand and process variability	Less valuable	Moderately valuable	More valuable

Spine and T-shaped layouts are also popular designs for freight and cross-docking terminals (Gue 1999).

6. Concluding Comments

In this paper, we showed that minimizing material-handling travel distances does not always reduce WIP. Therefore, the criterion used in the QAP formulation of the layout design problem cannot be used as a reliable predictor of WIP. Because the QAP formulation accounts only for full travel, an optimal solution to the QAP problem tends to favor placing departments that have large intermaterial flows in neighboring locations. In this paper, we showed that when we account for empty travel, this may not always be desirable. Indeed, it can be more beneficial if departments that have no direct material flow between them are placed in neighboring locations. In particular, we found that empty travel can be significantly reduced by placing the most frequently visited departments in neighboring locations regardless of the amount of flows between these departments. Because WIP is affected by both mean and variance of travel time, we found that reducing travel-time variance can be as important as reducing average travel time. Equally important, we found that the relative desirability of a layout can be affected by non-material-handling factors, such as department utilization levels, variability in department processing times, and variability in product demands. We also identified instances where the QAP-and WIP-based formulation are WIP equivalent. This includes systems with flow/dimensional symmetry or systems with low material-handling system utilization.

Several avenues for future research are possible. In this paper, the objective function was to minimize overall WIP in the system. In many applications, it is useful to differentiate between WIP at different departments and/or different stages of the production process. In fact, in most applications, the value of WIP tends to appreciate as more work is completed and more value is added to the product. Therefore, it is useful to assign different holding costs for WIP at different stages. This would lead to choosing layouts that reduce the most expensive WIP first (e.g., letting departments that participate in the last production steps be as centrally located as possible).

In addition to affecting WIP, the choice of layout determines production capacity. From the stability condition, $\rho_t < 1$, we can obtain the maximum feasible throughput rate:

$$\lambda_{\max}(\mathbf{x}) = 1 \Big/ \sum_i \sum_j \sum_k \sum_l \sum_s p_{rij}(x_{rk}x_{il}x_{js}(d_{kl}+d_{ls})/\nu). \tag{43}$$

Maximizing throughput by maximizing $\lambda_{\max}$ could be used as an alternative layout design criterion. In this case, layouts would be chosen so that the available material-handling capacity is maximized (i.e., ρ_t is minimized). The stability condition can also be used to determine the minimum required number of material-handling devices, $n_{\min}$, for a given material-handling workload, λ_t:

$$n_{\min} = \lambda_t \sum_i \sum_j \sum_k \sum_l \sum_s p_{rij}(x_{rk}x_{il}x_{js}(d_{kl}+d_{ls})/\nu). \tag{44}$$

Minimizing $n_{\min}$ can be used as yet another criterion in layout design. More generally, our modeling framework offers the possibility of integrating layout design with the design of the material-handling system. For example, we could simultaneously decide on material-handling capacity, such as number or speed of material-handling devices, and department placement, with the objective of minimizing both WIP holding cost and capital investment costs.

In certain applications, there may be a mix of material-handling technologies. For example, material transport between neighboring departments could be ensured by continuous conveyors, while material movement between more distant departments is carried out by a combination of forklift trucks and overhead cranes. In this case, material-handling capacity is clearly determined by the mix of technologies used. It would be useful to examine the relationship between layout and the deployment of different material-handling technology within the same facility.

In this paper, we assumed that travel-time variability is driven by differences in the distances between different departments. In some applications, there might be additional variability due to inherent variability in the travel distance associated with the same trip or in the speed of travel. Including either type of variability in computing Expressions (8) and (9) is relatively straightforward. The effect in both cases would be an increase in C_{s_t}. When this type of variability is high, we should expect the WIP difference between the QAP-and the WIP-optimal layouts to increase.

Finally, in our analysis we have assumed that there is always sufficient capacity for queueing. There are manufacturing environments where this capacity is limited. Limits on queue sizes can lead to the occasional *blocking* and *starvation* of the processing departments and the material-handling system, which in turn can reduce system throughput. The analysis of systems with limited queue capacity is considerably more complex and is worthy of future research. Although it is not entirely clear how layouts would affect throughput in systems where there is blocking, it is reasonable to conjecture that layouts that reduce congestion by reducing empty and full travel (along with variability) would also reduce the probability of blocking.

Acknowledgments
The author would like to thank Wally Hopp, the associate editor, and two anonymous reviewers for many useful comments on an earlier version of the paper. The author is grateful to Te Yang for help in generating computer code for the analytical models and carrying out the simulation experiments. This research is supported by the National Science Foundation under grant No. DMII-9908437.

References
Benjaafar, S., S. Heragu, S. Irani. 2002. Next generation factory layouts: Research challenges and recent progress. *Interfaces* Forthcoming.

Bitran, G. R., S. Dasu. 1992. A review of open queueing network models of manufacturing systems. *Queueing Systems* **12** 95–133.

——, D. Tirupati. 1989. Multi-product queueing networks with deterministic routing: Decomposition approach and the notion of interference. *Management Sci.* **35** 851–878.

Buzacott, J. A., J. G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems*. Prentice Hall, Englewood Cliffs, NJ.

Egebelu, P. J. 1987. The use of non-simulation approaches in estimating vehicle requirements in an automated guided vehicle based transport system. *Material Flow* **4** 17–32.

Fu, M. C., B. K. Kaku. 1997a. Minimizing work-in-process and material handling in the facilities layout problem. *IIE Trans.* **29** 29–36.

——, ——. 1997b. On queueing network models and the quadratic assignment problem. Technical report, College of Business and Management, University of Maryland, College Park, MD.

Green, J. 2000. Why workers are lining up for jobs at this GM plant. *Business Week* (October 2) ⟨www.businessweek.com⟩.

Gue, K. 1999. The effect of trailer scheduling on the layout of freight terminals. *Transportation Sci.* **33** 419–428.

Hopp, W., M. L. Spearman. 2000. *Factory Physics*, 2nd ed. Irwin/McGraw-Hill, New York.

Johnson, M. E., M. L. Brandeau. 1996. Stochastic modeling for automated material handling system design and control. *Transportation Sci.* **30** 330–350.

Kraemer, W., M. Langenbach-Belz. 1976. Approximate formulae for the delay in the queueing system GI/G/1. *Congressbook, 8th Internat. Teletraffic Congress*, Melbourne, Australia, 235.1–235.8.

Kelton, D. W., R. P. Sadowski, D. A. Sadowski. 1998. *Simulation with Arena*. McGraw-Hill, Boston, MA.

Kleinrock, L. 1976. *Queueing Systems, Volume 2: Computer Applications*. John Wiley, New York.

Kouvelis, P., A. S. Kiran. 1990. The plant layout problem in automated manufacturing systems. *Ann. Oper. Res.* **26** 397–412.

——, ——. 1991. Single and multiple period layout models for automated manufacturing systems. *Eur. J. Oper. Res.* **52** 300–314.

Meller, R., K. Y. Gau. 1996. The facility layout problem: Recent and emerging trends and perspectives. *J. Manufacturing Systems* **15** 351–366.

Pardalos, P. M., H. Wolkowicz. 1994. *Quadratic Assignment and Related Problems*. DIMACS Series in Discrete Mathematics, American Mathematical Society, Providence, RI.

Smith, G., J. Wheatley, J. Green. 2000. Car power. *Business Week–International Edition* (October 23) ⟨www.businessweek.com⟩.

Solberg, J. J., S. Y. Nof. 1980. Analysis of flow control in alternative manufacturing configurations. *J. Dynamic System Measurement Control* **102** 141–147.

Suri, R. 1998. *Quick Response Manufacturing: A Comprehensive Approach to Reducing Lead Times*. Productivity Press, Portland, OR.

Tanchoco, J. M. A. 1994. *Material Flow Systems in Manufacturing*. Chapman and Hall, London, U.K.

Tompkins, J. A., J. A. White, Y. A. Bozer, E. H. Frazelle, J. M. A. Tanchoco, J. Trevino. 1996. *Facilities Planning*, 2nd ed. John Wiley, New York.

Whitt, W. 1983a. The queueing network analyzer. *Bell System Tech. J.* **62** 9.

Whitt W. 1983b. Performance of the queueing network analyzer. *Bell Systems Tech. J.* **63** 1911–1979.

Whitt, W. 1993. Approximations for the GI/G/m Queue. *Production Oper. Management* **2** 114–161.

Womack, J. P., D. T. Jones. 1996. *Lean Thinking*. Simon and Schuster, New York.

Yang, T., S. Benjaafar. 2001. FLQ: A software for facility layout with queueing. Technical report—0015, Division of Industrial Engineering, Department of Mechanical Engineering, University of Minnesota, Minneapolis, MN.

Accepted by Wallace J. Hopp; received June 8, 1999. This paper was with the author 17 months for 3 revisions.