Report No. FAA-RD-71-55

ANALYTICAL TOOLS FOR THE STUDY OF AIRPORT CONGESTION

Bernard O. Koopman

ARTHUR D. LITTLE, INC.

Cambridge, Massachusetts

September 1971

Interim Report

October 1970 to September 1971

Availability is unlimited. Document may be released to the Clearinghouse for Federal Scientific and Technical Information, Springfield, Virginia 22151, for sale to the public.

Prepared for

FEDERAL AVIATION ADMINISTRATION

Systems Research and Development Service

Washington, D.C. 20590

1. Report No. FAA-RD-71-55

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle

ANALYTICAL TOOLS FOR THE STUDY OF AIRPORT CONGESTION

5. Report Date

September 1970

6. Performing Organization Code

71860

7. Author(s)

B. O. Koopman

8. Performing Organization Report No.

9. Performing Organization Name and Address

Arthur D. Little, Inc.
Cambridge, Mass. 02140

10. Work Unit No.

11. Contract or Grant No.

FA-70WA-2141

13. Type of Report and Period Covered

Interim Report
(October 1970 to September 1971)

12. Sponsoring Agency Name and Address

Systems Research and Development Service
Federal Aviation Agency
Dept. of Transportation, Washington, D.C.

14. Sponsoring Agency Code

15. Supplementary Notes

None

16. Abstract

In the analysis of air terminal congestion, two distinct but interlocking systems have to be considered: the air traffic control system, with its sensors, communication links, computers and displays; and the mechanical system composed of the runways, the ground and air space surrounding them, and the aircraft in their various states of takeoff and waiting for takeoff, of landing, joining the landing pattern, in stacks and so on. The tools of analysis developed in the present report are concerned with the second of these systems, in which the phenomena of congestion are in fact the phenomena of waiting lines and their interactions.

The final object aimed at in this report is a set of tools for calculating the quantitative effects on the performance of the air terminal of various assumed conditions, operational options, and suggested technical improvements. The effects on the performance are to be expressed in terms of throughput rates, delay reductions, and diminished numbers of aircraft not accommodated. ( )

The first step is to develop various queue models under strongly time-dependent conditions – no such treatments apparently being [illegible]ilable. No stationary solution of these time-dependent queues can exist; but if the external conditions have a [illegible]rnal periodicity, so does just one solution, every other one approaching it exponentially. Three models are consid[illegible] in parallel: random (Poisson) arrival and random service; random arrival and fixed time of service; non-random arrival and service – i.e., deterministic flow. The first model was communicated in our 1969-1970 Interim Report No. FAA-RD-70-70. The second is quite new; it represents the opposite extreme of the Poisson service model; surprisingly enough, under corresponding conditions it gives results that are very close to the former. The deterministic flow model, on the other hand, leads to quite different results in many cases. All these properties are exhibited graphically.

The second step considers two interacting queues, such as the takeoff and landing aircraft using the same airstrip. First, an averaging process is applied to remove irrelevant detail, and yield a crosscut model of a Markovian form. Second, by a marginal summation, the double queue is reduced to two single queues, interacting only through certain coefficients, as in mutually perturbing physical systems. Methods of approximation are given. No numerical results are communicated.

Certain examples of practical application are given, calculating the benefits of hypothetical increased landing rates and flow control.

17. Key Words

- Air Traffic Control
- Congestion
- Queue Models
- Flow Control

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
None	None	62	$3.00 (hard copy) .65 (microfiche)

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

1. SUMMARY

Because demand for air transportation is already taxing the capacity of our present system, and because it is projected to increase greatly in the years ahead, every consideration is being given to technical methods of maximizing the capacity of the system. Many of the technical developments will lead to great financial expenditures in development, manufacture, installation, and maintenance. It is necessary to seek every method to optimize the choice of technical system as well as its mode of operation. For this purpose, the method so often used in engineering developments of actually installing the system and then finding how well it works is not only very expensive, but is likely to give inconclusive and ambiguous results – unless enough precise theoretical knowledge can be brought to bear on the situation to clarify the nature of its interactions and the relationships of cause and effect. In this respect, the situation has all the well-known features of problems which can be clarified by operations research.

Of the two categories of system that are involved in an air terminal – the ATC system, and the actual aircraft, air-space, and runways – the latter is the object of the present examination. By a quantitative study of airport congestion, with its notorious tendency to lead to lengthening waiting lines, aircraft stacked in holding patterns, and the consequent delays and diverted flights, we have been able to reveal causal relations and to develop methods of prediction. But this has been possible only because we have been able to develop appropriate analytical tools.

Already in our previous Interim Report, issued in October 1970 (Report No. FAA-RD-70-70), we have set forth the reasons for rejecting standard queue methodology as inapplicable to the study of air terminals (its inability to deal with strongly time-dependent conditions, its ignoring of queue length limitations, the inflexibility of its formulas). In Chapter 6 of that report we established a method for dealing with single queues under strongly time-dependent conditions, limited total length and random service times (e.g., Poisson statistics of landing).

The present report analyzes the whole question of the effect of various service statistics, and shows by a set of computations under rather wide ranges of conditions that the results are decidedly insensitive to the assumptions concerning these statistics: the extreme of random service and the extreme of regularity lead to nearly the same numerical values. This is an entirely new result, that had never been suspected. It is of great usefulness, since it makes it unnecessary to go to great lengths and complexities of formula in the attempt to come as close as possible to realism.

In the present report, the rather simple traffic model that ignores random and concerns itself with the mean motions only, has been compared in its quantitative results with the queue models taking random into account. This simpler model, while giving results rather close to the others in cases of gross overloading of the terminal, does not agree at all well with them in cases of more moderate utilization. It tends to be over-optimistic – a fact that may suggest the possibility of approximating to its assumed lack of random, as a measure of improvement by flow control.

Another development of a tool of analysis is a systematic method of setting up in a rational form the equations of evolution of a double queue, by the concept of the "cross-cut averaging." Once set up, such a system of equations can be reduced, by a simple method we develop, to two equations of single queues, where each influences the other as a perturbation of its coefficients, rather than by the details of its condition. Various methods of approximate solution are set forth. We have not carried out numerical computations by the method at the time of writing; but are prepared to do so as soon as a problem of practical evaluation of suggested equipment makes this appropriate.

The success of the methods reported on herein is connected, we believe, with their attempting an optimum mix of mathematical reasoning and formulation, with modern computing equipment. The relative limitations in mathematical treatments before modern computers were available, and the extreme expense in time, money, and inconclusive results, that so often occur when computers are not subordinated to mathematical reasoning, show the desirability of the middle ground.

2. INTRODUCTION

2.1 BACKGROUND

In examining the performance of the air transportation system, both in the present and in future forecasts, and in estimating the benefits of various proposed equipment and operational improvements, it is convenient to think of the whole situation as being made up of two components: On the one hand, there is the air traffic control system (ATC), composed of the various sensing devices, communication networks, displays, computerized data processing, and including all the feedback loops and the man-and-machine teams in the control tower and the cockpits. On the other hand, there is the physical system composed of the runways, air space above them, and the aircraft arriving, landing, and taking off. The ATC system is being examined and forms the subject of other interim reports under the present contract. The present interim report is addressed to the analysis of the behavior under various conditions of the aircraft-air-terminal system.

A major problem facing our air transportation system is that of capacity – the ability of the system to accommodate the demands increasingly placed upon it without unacceptable degradation of service, such as passengers turned away, delayed or cancelled flights, or decreased safety. Both parts of the system face the problem of capacity: the ATC can be overloaded; and even when it is able to handle its burden, the facilities for landing, takeoff, and air or ground holding can be exceeded. It is to this second category of capacity bottlenecks that this report is devoted. In principle, the subject comes under that of waiting lines or queues. But since practically all of the treatments of queues found in the literature are based on assumptions that are quite invalid in the case of air terminals, it could be seriously misleading to say that we are applying "queue theory" to the air transportation problems. Actually, we have had to create a new branch of this methodology to deal with these problems. The explanation and application of these new analytical tools are the subject of the present report.

2.2 THE NEED OF NEW TOOLS OF ANALYSIS

The inapplicability of standard queue methods is due primarily to the fact that they are based on the assumption of fixed external conditions, whereas, as should be clear to anyone who has traveled by air at various times of day, and as will be shown by a wealth of airport statistics, the arrival rates at the terminal, for both landing and takeoff, are strongly varying functions of the time of day: very few during the early morning hours, and very many at the morning and afternoon rush hours. The standard mathematical treatments, with their fixed external conditions, are led to transitions or transition probabilities that are independent of the time. Then and only then will there exist a stationary state solution; and most of the attention is usually confined to this special and simple solution. But even when the standard treatments get away from the stationary solution and examine "transients," i.e., time dependent solutions of time independent equations, these are still not applicable to the study of airport congestion, except when the examination is confined to such short intervals of time that the external conditions do not change appreciably during their course. Transients in a system with constant transition probabilities are far simpler and less realistic than the time-varying solutions in the case of time-dependent transition probabilities.

As noted above, when the external conditions (and thus the transition probabilities) vary with the time, no stationary solutions are possible. On the other hand, when the conditions exhibit a diurnal periodicity of T = 24 hours, one and only one solution having this period will exist; all other solutions approach it exponentially as time increases. This fact provides a "relaxation method" for calculating the periodic solution numerically. This periodic solution takes the place, in the present methodology, of the stationary solutions of the older methods.

A secondary reason for the limited applicability of standard queue theory to air terminals is the assumption usually made that the number in the queue is unlimited. Taken literally, this is absurd, since it not only assumes infinitely many available aircraft, but infinitely many spaces for air stacking or ground waiting lines. If this infinite assumption were able to simplify the quantitative discussion, it could be used as a convenient fiction – as similar assumptions are permitted in so many applications of mathematics to physics. Actually, the reverse is the case: the queue equations are infinite in number when the queue is allowed to be unlimited, and this greatly complicates the theory that is needed. With the assumption that will be made throughout this report, a number is set as the limit of queue size. Beyond its greater realism, this assumption has two obvious advantages: it simplifies the theory and it facilitates the numerical computation. But the assumption of queue limitation has the even more important advantage of making possible the quantitative study of the expected number of aircraft "turned away" – not allowed into the waiting line and either held at the port of origin or diverted. This is an important element in gauging the capacity of the system.

A method increasingly used in operational problems of the present sort is "simulation," or more explicitly "computer simulation" or "Monte Carlo simulation on a computer." In this method, a modern electronic computer has one part of its capacity programmed to record any one of a set of states, intended to correspond in a one-to-one way with the possible states of the actual system studied (here, the air terminal). The second part of its capacity is made to contain a rule of transition, whereby, when at a particular time the computer records a particular state, the next state or states are selected. The selection may be deterministic, the next state being determined by the record of the time and state that the computer is in; or it may be random, employing a selection according to a given probability distribution, implemented by a table of random numbers. Since the present conditions are strongly time-dependent, there will be a different rule of statistical selection for each different time.

Clearly, if suitably accurate knowledge of states and state transitions is inserted into the computer, and it is allowed to make one run corresponding to a twenty-four hour period of real time, it will give information no more complete than that obtained by observing the actual operation of the airport during one day. This would not be regarded as a sufficient basis for important decisions because of its inability to separate reproducible from random effects: it is necessary to make a large number of runs to obtain a given level of reliability.

There exists at Arthur D. Little an extensive experience in the use of this method of computer simulation. Its application to the air terminal problems in this contract has been given detailed consideration, but it has been discarded in favor of a conventional mathematical treatment, with computers used for computation only – that is, "computer-assisted analysis." The reasons for our decision are three-fold: first, the cost-effectiveness of obtaining the required results with the required degree of reliability is unfavorable – by several orders of magnitude – with the use of simulation as compared with computer-assisted analysis; second, the simulation method, while avoiding no difficulty intrinsic to the problem, makes the exact assumptions and reasoning less visible and the relations of cause and effect less clear than the analytical method we have adopted; finally, simulation is incapable of establishing such general properties as the existence and uniqueness of the periodic state and the exponential approach of all other solutions to this one – it could at most suggest such a situation.

2.3 SCOPE

The necessity of developing new methods for the study of time-dependent queues in the study of airport capacity was first set forth in the September 1969 to August 1970 Interim Report submitted under this contract (Report No. FAA-RD-70-70, Chapter 6).[1] The single, time-dependent queue was examined on the basis of purely random (Poisson) arrival and service times (landing or takeoff), with strongly time-dependent parameters (expected numbers per hour). Data taken from 1968 activity statistics at J.F. Kennedy and LaGuardia were used as input parameter values in a number of calculations of different important indicators of capacity, such as expected length of waiting lines, their fluctuations, the delays, and expected numbers turned away at various times of day. The results were exhibited as graphs. The double-queue problem was given mathematical study, but this was only preliminary and was not carried through to numerical results.

The present report concerns material that forms a deepening and amplification of that methodology presented in the Interim Report. Its first step is to examine the range of possible assumptions regarding the service statistics, all the way from the purely random (Poisson) to the purely deterministic (fixed interval of service) in the case of the single queue. A full model based on the latter assumption is set up, and its results compared with those of the former. From the practically unimportant differences, it is concluded that there is little sensitivity to this category of assumption. The arrival assumption of pure random has a high degree of plausibility and observational confirmation. Nevertheless, it has often been replaced by an assumption of perfect regularity, usually to avoid the mathematics required to deal with the random features. This assumption, coupled with that of perfect regularity of service, makes the problem coincide mathematically with the model of the flow of a liquid into and out of a reservoir. This has been discussed in Chapter 3 of the 1969-70 Interim Report.[1] We make calculations based on this deterministic flow model and compare them with the results of the random models, using the same inputs. The flow results are found to differ considerably from the latter, so that the regular arrival assumption cannot be regarded as a reliable basis for a treatment of airport congestion. All this material on the degree of sensitivity of results to assumptions seems to be entirely new; it forms the subject of Chapters 3 and 4.

The next problem is that of the double queue: a single-runway facility shared by two categories of aircraft – those wishing to land and those waiting to take off. More generally, there is the problem of the multiple queue, which has to be considered when two or more terminals are examined together as a system complex. Before any quantitative treatment of this class of problem can be started, some rather deep issues have to be faced regarding the degree of detail in which the *state* of the system must be specified before any sensible law of transition can be formulated. Both computer simulation and computer-assisted analysis are forced to act as if the state of the system, together with the time of day, determines the transition probabilities into the immediately succeeding states. In technical language, this means that a Markov assumption is required. (Even when the record of a few preceding states is used in the selection of the succeeding state, a simple mathematical reformulation shows that the Markov assumption is still being used.) The justification of this assumption – essentially a redefinition of the relationship between the idealized model and the day-to-day occurrences – is the first object of Chapter 5.

The second object of Chapter 5 is to formulate the equations of evolution of the multiple queue, in the light of the considerations adduced up to this point. The third object is to reduce the resulting equations by a method of marginal sums, to the equations of the separate single queues, acting upon one another by a perturbation of their individual transition probabilities. This method of simplification, inspired by the treatment of coupled systems in mathematical physics, may be made the basis of a sequential algorithm of successive approximations, whereby a crude first approximation is improved step by step. Each step is taken by a straightforward computer program.

In Chapter 6, an example of the use of these methods is given, answering the practical question: what are the benefits of an improvement, even slight, of handling rates? Since one of the possible contributions of advanced electronic systems, such as ARTS-III, could be to handling rates, this example is relevant to their practical evaluation.

As a further practical application, we compare the results for our random arrival queues with those of the deterministic flow model. The latter gives decidedly more optimistic values of such indices of capacity as aircraft turned away, etc. While this shows the misleading nature of the deterministic assumption when applied to a random situation, it also indicates quantitatively how the diminution of the randomness of the arrivals by flow control could improve the capacity of the terminal.

Finally, by means of the methods set forth herein, precise numerical calculations can be made concerning the *metering loss* in the following sense: When the positions of aircraft approaching the terminal and then the runway are not known precisely (either by tower or pilot), safety requires an increase in spacing over what would be possible with much more precise knowledge. The effects of this are two-fold: a decrease in the mean rate of landing and a greater degree of random (diminished flow control). As indicated above, the calculation of the impact of these effects on capacity is a simple exercise in the methods of this report. Work is continuing on this problem.

3. ASSUMPTIONS OF ARRIVAL AND SERVICE: THE SINGLE QUEUE

3.1 POSSIBLE SERVICE ASSUMPTIONS

"Service assumptions" in the present context concern such variables as, in the case of takeoff, the total interval of time taken by the aircraft from the moment when it is given the all-clear signal and order to take off to the actual moment when it takes off, leaving the runway safe for the next user (to take off or to land). Similarly, in the case of landing, the service time means the length of time between the instant when the clear-to-land signal is received by the aircraft in the approach pattern (the first in turn to land) and the moment when it has not only landed but has cleared the runway so that the latter can be used by the next aircraft.

The importance of the service time assumption is seen in the case of several aircraft waiting to use the same facility. Suppose that, as in the highly regimented takeoff of fighter aircraft from a carrier on a combat mission, the use of the runway takes an exact interval of time (c units). Then the instants of takeoff would form the regular array, starting with the first time t_0; that is, $t_0, t_0 + c, t_0 + 2c, t_0 + 3c$, etc. The number of fighters waiting to go would change only at these instants. This law of service is called that of "fixed-service time." It is possible, even with such a regimentation, that the interval c might vary in a predictable way with such conditions as degree of darkness; but this would not be a random but a slow "secular change."

In contrast to this regular case, there are many situations in which various causes of irregularity operate, as for example, individual differences of aircraft or of their pilots, special conditions at the airport (physical or personal) including wind and weather. These are but examples: it is of the essence of random that a list of its origins can never be conclusive. All that can be said, in general, is that the length of service time has (at a given time of day, etc.) a predictable expected value (mean, found by averaging), and that the distribution of service times about it is some more or less complicated curve – which may be peaked about the mean, or spread more or less widely away from it. The more it is peaked, the closer the situation resembles the case of perfect regularity of length c. What, mathematically, is the opposite extreme?

The above question asks, in effect, for the "most random" distribution having a given mean $\bar{t}$. The question, while using "random" in a rather intuitive sense, can be made precise enough to answer by adopting the measure of random provided by modern information theory. If p(t) is the probability density of the length of service t ($t \geq 0$), the quantity of information regarding the value of t provided by this given density is the integral

$$I_p = \int_0^\infty p(t) \log p(t)\, dt$$

PRECEDING PAGE BLANK

If the *bit* is the unit of information, "log" must be the logarithm to the base 2; but this is immaterial to the following reasoning. Our question then takes the following form: Of all the distributions p(t) on the positive time axis having a given mean

$$\int_0^\infty p(t)\, t\, dt = \bar{t}$$

which one contains a minimum information I_p (or, equivalently, a maximum "entropy" $-I_p$).

A simple exercise in calculus of variations (using the convexity of the function x log x) gives the answer that it is the Poisson density $p(t) = \mu e^{-\mu t}$, where $\mu = 1/\bar{t}$. This, then, will be taken as the extreme case of the perfectly random service time. Fortunately for our subsequent calculations, it is very simple.

The point of view to which our examination of the air terminal problem has led us is that the reality lies in an unknown position between the absolute regularity of service time c and the complete random of the Poisson service with the same mean ($c = \bar{t} = 1/\mu$). Our basic methodological strategy is to carry through formulations and calculations of quantities of importance to the practical issues on the basis of each of these extremes: when the results, graphed as functions of time, are curves lying close together, the "reality" will be regarded as the *band* between the two curves. This is quite different, and we submit, far more realistic than the point of view which regards one of the two curves as the correct one, the other as a more or less good approximation of the correct one.

The striking fact to be illustrated below is that the two curves calculated on the basis of the two extreme assumptions lie close together in the ranges of practical importance. This shows that either one of the curves – and, hence, the assumptions which led to it – is a satisfactory approximation to reality, as represented by the band between them. Strengthened by this observation, the Poisson assumption of service time will be used as the basis of the treatment of the multiple queue problems of Chapter 4.

3.2 THE ARRIVAL ASSUMPTION

Even under the most extreme degree of military regimentation, fighters returning to base after a routine drill arrive at irregular intervals; and this irregularity in the times at which aircraft, destined for a particular terminal, become candidates for admission to its landing pattern – or holding stacks – is a matter of general observation. Put precisely, the probability that a new arrival – or application for admission – occurs between an epoch (time of day) t and an immediately subsequent epoch $t + \Delta t$ (Δt small: e.g., a minute or two) is a quantity having the following two properties: first, to quantities of higher order in Δt, it is proportional to Δt, and may thus be written as $\lambda \Delta t$; second, it is the same, for a given number of aircraft ahead of it, without regard of the exact times at which these had entered the queue. Of course, the coefficient λ will vary with the time of day, but if its amount of change is negligible during the period from t to $t + \Delta t$, then the usual elementary reasoning shows that the probability of exactly k arrivals during this period is given by the Poisson expression $(\lambda\Delta t)^k e^{-\lambda\Delta t}/k!$.

3.3 THE EQUATIONS OF EVOLUTION: POISSON LAW OF SERVICE

In dealing with a single queue which, in the present applications, may be the aircraft in stacks and in the landing patterns awaiting clearance to land, or else the line of aircraft waiting on the ground for takeoff, the crucial quantity is the number of aircraft at time t. Since this is a random number it must be handled by the methods of probability. We introduce the function $P_n(t)$ which is the probability that at the time t there be n aircraft in the queue. Clearly, n must be zero or a positive integer. Furthermore, since, as explained in 2.2, we are assuming a given upper limit m for the number of aircraft admitted to the queue, we must have $n = 0,1,2,..., m$. For any other values of the subscript, $P_n(t) = 0$.

We first recall briefly the equations of evolution of our system when, in addition to assuming a Poisson arrival time, we assume a Poisson service time. To provide enough flexibility for the study of a multiplicity of situations, we shall assume that the Poisson parameters λ and μ may depend not only on the time of day t, but on the size of the queue n, and we shall write

$$\lambda = \lambda(t) = R_n(t), \quad \mu = \mu(t) = L_n(t)$$

The R quantities are "transition probabilities" of a motion to the right in the array of values of n, the L quantities corresponding to a leftward jump; they may be called the right- and the left-transition probability coefficients.

With these designations, the probabilities $P_n(t + \Delta t)$ may be computed algebraically in terms of the probabilities $P_{n'}(t)$ from which the number n could arise. The final result is the system of differential equations

$$P_n'(t) = R_{n-1}(t) \quad P_{n-1}(t) - [R_n(t) + L_n(t)]\, P_n(t) + L_{n+1}(t)\, P_{n+1}(t) \tag{3.1}$$

which is true for $n = 1, 2, ..., m - 1$; whereas for $n = 0$, we have

$$P_0'(t) = - R_0(t)\, P_0(t) + L_1(t)\, P_1(t), \tag{3.2}$$

and for $n = m$, the maximum number allowed in the queue,

$$P_m'(t) = R_{m-1}(t)\, P_{m-1}(t) - L_m(t)\, P_m(t). \tag{3.3}$$

Equations (3.1), (3.2), and (3.3) form a system of m + 1 ordinary, linear, homogeneous differential equations of the first order, of a type studied in mathematical physics for centuries and the general properties of which are thoroughly known. In particular, given any initial values $P_n(o)$, there will be one and only one solution. If the initial values are probabilities (non-negative and adding up to unity), the same will be true of the corresponding solution. Finally, if the coefficients are periodic in t, i.e., if

$$R_n(t+T) = R_n(t), \; L_n(t+T) = L_n(t), \tag{3.4}$$

and if all the transition probabilities go through an interval of positiveness, there will be one and only one set of probabilities $P_n(t)$ satisfying the equations and having this period T: $P_n(t+T) = P_n(t)$. This unique solution is approached exponentially, with increase in time, by every other system of probability solutions. It plays, therefore, the role of the steady-state solution, which exists only when all the transition probabilities $R_n(t)$ and $L_n(t)$ are independent of time.

It remains to obtain the solution numerically for various actual or representative values of the transition probabilities which, in general, (particularly λ) have to be given in tabular form. This requires computer routines of a standard type for solving differential equations of the present sort. Actually, the quantities of practical significance are the expected value $\bar{n}(t)$ of the number in the queue at the times t, its standard deviation, the cumulative number of aircraft "turned away" (not admitted because the stack is full: they may be diverted or held on the ground at their port of origin), and, finally, the expected waiting time of an aircraft which joins the queue at the epoch t. All these come out of the computer program and are discussed in Section 4 below.*

3.4 THE EQUATIONS OF EVOLUTION: FIXED SERVICE TIME

The treatment of this case is vastly simplified if we "atomize" the time axis, as follows: Starting with an arbitrarily fixed initial time $t = 0$, let us consider the moments or epochs $t_1 = c$, $t_2 = 2c$, ..., $t_s = sc$, etc. Suppose that an aircraft, at the head of the queue, is serviced (and hence, removed) at $t = 0$. As long as there are others awaiting their turns, aircraft will be removed at the regularly spaced epochs t_1, t_2, etc. Next, consider arrivals. To confine changes in queue size to the exact moments defined above, we shall assume that, if during the interval of values of t ($t_s < t \leq t_{s+1}$), k aircraft "apply for admission," they will be held off until any aircraft scheduled to be serviced at t_{s+1}, admitting them at this precise epoch. Any aircraft of these k for which no spaces exist at t_{s+1} are definitively turned away – i.e., their presence will have no influence on the numbers of applications for admission during the subsequent period $t_{s+1} < t \leq t_{s+2}$.

It follows from these assumptions that the changes in queue length can occur only at the discrete, equally spaced epochs t_s defined above. We submit that these conventions for

*For details, see Reference 1.

admission cannot make any material difference, nor will our subsequent assumption that c goes exactly into the T = 24-hour day.

We shall denote by p_n^s the probability that at the epoch $t_s = sc$ (equivalently, the epoch s) there be precisely n aircraft in the queue; that is, we set $p_n^s = P_n(sc)$ as defined earlier. Evidently p_n^s will be zero, except for the m + 1 values of the subscript $n = 0, 1, \ldots, m$.

Since the Poisson assumption of arrival is being made, the probability of exactly k applications for admission occur during $t_s < t \leqslant t_{s+1}$ is given by the formula

$$\frac{a^k}{k!} e^{-a}, \quad a = \lambda c.$$

Since λ is time-dependent, a must bear an index referring to the time of day it is considered, e.g., $a = a_s = c\lambda(t_s)$. For complete generality, we should, of course, also affix it with a second index, indicating the number in the queue at the time considered, and write $a = cR_n(t_s) = a_{s,n}$. However, this will not be done, for the following reason: the principal objective in this part of the study is the comparison of the results of the Poisson service assumption with those of the fixed service time assumption. While this comparison should be made under reasonably broad conditions, it need not be under the widest conceivable ones, such as arrival rates depending on number in queue, which do not apply to simple, single queues, and enter only in the study of complexes. Accordingly, we are confining this section to the study of arrival rates independent of queue length.

On the basis of these assumptions the equations of evolution of the system during successive epochs s take the form of m + 1 recurrence equations, coming under the calculus of finite differences. They determine p_n^{s+1} in terms of the m + 1 values of p_n^{s+1}. They are obtained by the usual probability reasoning, as follows:

Consider first the case of p_0^{s+1}. The event of no aircraft in the queue at s + 1 can occur in just two possible (and mutually exclusive) ways: by having none at s and none applying for admission between s and s + 1; or else by having one at s (which will be removed at s + 1), and none applying, as before. Compound and total probability show that

$$p_0^{s+1} = p_0^s\, e^{-a} + p_1^s\, e^{-a}$$

Next let n be any positive integer less than m. The event of probability p_n^{s+1} can occur in the following n + 2 different ways: no aircraft at s and n applications between s and s + 1; one aircraft at s and n applications; and, so on; finally, n + 1 aircraft at s and no application between s and s + 1. From this we obtain the first m equations of the system of m + 1 equations (3.5). The last equation is obtained by similar reasoning, but is modified by the "spill-over" possibility: to fill a queue at s + 1 when at s there were k members, any number of arrivals equal to, *or greater than,* m – k + 1 will work; the probability of this is the

u_{m-k+1} defined by summing the corresponding Poisson probabilities, as follows (with $m - k + 1$ replaced by r)

$$u_r = \sum_{i=r}^{\infty} \frac{a^i e^{-a}}{i!} = 1 - \sum_{i=0}^{r-1} \frac{a^i e^{-a}}{i!} .$$

These sums can be expressed as incomplete gamma functions.[2]

Thus we obtain the recurrence or "difference equations"

$$\begin{aligned}
p_0^{s+1} &= (p_0^s + p_1^s)\, e^{-a} \\
p_1^{s+1} &= (p_0^s + p_1^s)\, a e^{-a} + p_2^s\, e^{-a} \\
p_2^{s+1} &= (p_0^s + p_1^s)\, \frac{a^2}{2!} e^{-a} + p_2^s\, a e^{-a} + p_3^s\, e^{-a} \\
&\cdot \\
&\cdot \\
&\cdot \\
p_n^{s+1} &= (p_0^s + p_1^s)\, \frac{a^n}{n!} e^{-a} + p_2^s \frac{a^{n-1}}{(n-1)!} e^{-a} + \ldots + p_{n+1}^s\, e^{-a} \\
&\cdot \\
&\cdot \\
&\cdot \\
p_m^{s+1} &= (p_0^s + p_1^s)\, u_m + p_2^s\, u_{m-1} + \ldots + p_m^s\, u_1
\end{aligned} \tag{3.5}$$

We emphasized that the quantities a and u_i will, in general, depend on the epoch s, and for explicitness should be written as a_s and $u_{s,i}$.

Equations (3.5) form a system of recursive or "difference" equations, m + 1 in number, for obtaining p_n^{s+1} in terms of p_n^s, and thus successively in terms of any given initial p_n^0 (n = 0, 1, ..., m). Four facts stand out at once: *first,* if the given initial values are non-negative, then so are all those obtained by recurrence because the coefficients are non-negative; *second,* if the initial values add up to unity, so do all subsequent sets of values; to see this we have but to add all the equations (3.5), obtaining in virtue of (3.4) the relation

$$p_0^{s+1} + p_1^{s+1} + \ldots + p_m^{s+1} = p_0^s + p_1^s + \ldots + p_m^s \ .$$

Third, all values are uniquely determined – *and numerically computable* by multiplications and additions – in terms of the given initial values: if these are probabilities (non-negative, and adding up to unity), the same will be true of the computed values.

The *fourth* fact implied by (3.5) results from the vanishing of the determinant of the $(m+1)$ by $(m+1)$ matrix M of coefficients of the $m+1$ quantities p_n^s on the right in (3.5) – this determinant having its first two columns equal. By the elementary theorems of linear dependence, equations (3.5) will be *inconsistent,* unless the rank of the "augmented matrix" $\| p_n^{s+1}, M \|$ in (3.5) is the same as that of M; the augmented matrix is formed by adjoining the column of quantities on the left to M. Appendix B shows that the rank of M is m, and that the augmented matrix will have the same rank if and only if the quantities p_n^{s+1} satisfy the following equation (in addition to having a unit sum):

$$D_m \, p_0^{s+1} - D_{m-1} \, p_1^{s+1} + D_{m-2} \, p_2^{s+1} - \ldots + (-1)^{m-1} D_1 \, p_{m-1}^{s+1} + (-1)^m e^{-a} = 0 \tag{3.6}$$

$$D_m = \frac{a^{m-1}}{(m-1)!} - \frac{a^{m-2}}{(m-2)!} + \ldots + (-1)^m a + (-1)^{m-1}, \text{ etc.}$$

Note that the quantities a and D_m in these equations are functions of the epoch s or t_s as explained before.

The last result means that every set of probabilities p_n^{s+1} that evolves from initially given probabilities by the process expressed in (3.5) must satisfy (3.6), and that if the initial probabilities p_n^0 had not satisfied this relation, they could not have evolved from any previous ones through such a process. This is important in the case of periodic arrival rates $\lambda(t) = \lambda(t+T)$, and when we wish to find a periodic solution starting from selected initial values: a necessary (but not sufficient) condition for periodicity is that the initial values satisfy (3.6).

Both in numerical computation and in application of the general theory, it is better to replace (3.5) by a slightly modified system having m rather than $m+1$ unknowns and a non-singular matrix of coefficients. For this purpose we add the first equation in (3.3) to the second and then introduce the symbol

$$\overline{p_1^s} = p_0^s + p_1^s \quad (s = 0, 1, 2, \ldots), \tag{3.7}$$

whereupon we obtain m recurrence equations in the m unknowns $(\overline{p_1^s}, p_2^s, \ldots, p_m^s)$. Suppose that these m quantities are written as a single-column matrix p^s, and we let the square matrix of coefficients be

$$A = \left\| \begin{array}{cccc} (1+a)e^{-a} & e^{-a} & 0 & 0\ldots0 \\ \frac{a^2e^{-a}}{2!} & ae^{-a} & e^{-a} & 0\ldots0 \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \frac{a^{m-1}}{(m-1)!}e^{-a} & \frac{a^{m-2}}{(m-2)!}e^{-a} & & \ldots e^{-a} \\ u_m & u_{m-1} & & \ldots u_1 \end{array} \right\| = A(a) = A(a_s) = A_s \qquad (3.8)$$

Then our recurrence equations may be written in matrix form as $p^{s+1} = A_s p^s$. The solution by recurrence is then

$$p^s = A_{s-1}\, A_{s-2} \cdots A_0 p^0 \qquad (3.9)$$

The fact that A is non-singular and, indeed, that its determinant is given by the formula

$$\det A = \det A(a) = \frac{a^{m-1}e^{(m-1)a}}{(m-1)!} \qquad (3.10)$$

is proved in Appendix B. Therefore, the initial values, the elements of p^0, may be chosen without restriction, other than non-negativeness and unit sum. Consequently, there is a *fundamental system* of m independent solutions of (3.9), in terms of which all other solutions may be expressed as linear combinations. Such a fundamental system can be obtained by writing down the unit m-by-m matrix I (units in the principal diagonal; all other elements zero) and regarding each column in I as a set of initial probabilities p^0 – the k'th column corresponding to an initial queue length of just k aircraft. The resulting m independent solutions form the m columns in the matrix product

$$P^s = A_{s-1}\, A_{s-2} \ldots A_0 \qquad (3.11)$$

This is shown by the application of (3.9) to each of the columns p^0 of initial values in matrix I. Since the determinant of P^s is the product of those of the A matrices, which are not zero, the same is true of that of P^s. Therefore, it represents m independent solutions. Every other solution is a linear combination of the latter, with coefficients which are its initial probability values.

Once the values of $\bar{p}_1^s$, p_2^s, etc., are obtained, we find p_0^s and p_1^s from equations (3.6) and (3.7). It is usually sufficient to ignore p_0^s and find p_0^{s+1} in terms of $\bar{p}_0^s$ using (3.6) and the first equation in (3.5). Then $p_1^{s+1} = \bar{p}_1^{s+1} - p_0^{s+1}$.

We conclude by stating the theorem, the proof of which can be found in Appendix A.

THEOREM. If the arrival rate $\lambda(t)$ is periodic in t: $\lambda(t + T) = \lambda(t)$, one and only one solution p^s will be periodic and it will have T, as its period, $p^{s+S} = p^s$ ($S = T/c = [T/c]$). Every other solution will approach this one exponentially with an increase in t (i.e., of s).

Note that we have assumed that the service time c goes exactly into T (24 hours), but this represents no material limitation of the result.

3.5 REGULARIZATION OF ARRIVALS: DETERMINISTIC FLOW[3]

As explained in 3.2, the assumption of random arrivals obeying the Poisson Law corresponds to what would be expected and to what is observed in the ordinary situation at an air terminal, when no detailed scheme of regularizing them by "flow control" is used. If such regularization is done without changing the mean arrival rate $\lambda(t)$ – i.e., without rescheduling – it must take the form of making the aircraft arrive at regular intervals, as far as possible, which must therefore be of lengths $1/\lambda(t)$. Evidently such regular arrivals will tend to regularize the times of service, and therefore the appropriate service assumption will also imply that just one aircraft is removed from the waiting line during each interval of length $c = 1/\mu$. There remains the question of just how much improvement in service, if any, would such a regularization, if possible, produce.

To answer this question, it is necessary to examine the quantitative consequences of the assumptions of perfect regularity of input and output. Except for its "granularity," such an air terminal queue will behave mathematically precisely as in the case of a liquid flowing into an open reservoir of capacity m, the rate of inflow being $\lambda(t)$, and of outflow through an opening at the bottom at the rate μ. The finiteness of the air queue, with possible turning away of aircraft exceeding its capacity, has its counterpart in the openness of the reservoir with spilling out of excessive liquid. Accordingly, we shall call this the *deterministic flow model.*

Apart from its use in the study of possible benefits of flow control, this model has, in substance, been used because of its far greater simplicity than the models admitting the presence of random. An implicit application is made to scheduling by considering arrival, landing, takeoff *rates*, and balancing them so as to obtain consistent schedules: to reason in this way is, in fact, to apply the mathematics of the deterministic flow model. In justifying such applications to the actual cases that do, in fact, contain random, it is necessary to compare the results with those of the two models we have studied earlier in this chapter. This will be done in the next chapter and it will be seen that, as stated in 2.3, its results differ considerably – and optimistically – from those of the random models.

The mathematical treatment of the deterministic flow model requires the total input function

$$V(t) = \int_0^t \lambda(t)\, dt$$

which represents the total volume of liquid that has flowed into the reservoir from an initial time 0 to the arbitrary time t. This can be computed numerically from the table of values of $\lambda(t)$, given, for example, by airport statistics. With our regular rate of outflow $\mu = 1/c$, we do not need integration to find the amount of liquid leaving the reservoir through the lower opening, this being simply the product of μ by the time during which there is liquid in the reservoir.

The computation is based on the following law governing the rate of change, $dn(t)/dt$, of the volume $n(t)$ in the reservoir at time t (corresponding to the number of aircraft in the queue).

$$\frac{dn(t)}{dt} = \begin{cases} 0 & \text{if} \quad n(t) = m \quad \text{and} \quad \lambda(t) > \mu \\ 0 & \text{if} \quad n(t) = 0 \quad \text{and} \quad \lambda(t) < \mu \\ \lambda(t) - \mu & \text{if} \quad 0 < n(t) < m \end{cases}$$

To show how this works, suppose that at an arbitrarily chosen initial time t_0 the number $n(t_0)$ is positive but less than m, so that the third line applies. On integrating we obtain

$$n(t) = n(t_0) + V(t) - V(t_0) - \mu(t - t_0).$$

This formula applies between t_0 and the first time t_1 that $n(t)$ equals zero or m. Then the appropriate rate of change formula is applied and $n(t)$ is computed from t_1 until a first time t_2 of the next change of formula, etc. Numerical details are illustrated in Chapter 4. An easy use of the function $V(t)$ gives the amount of water spilled over the top (aircraft turned away), once the succession of times t_1, t_2, etc., has been found by the above process. Numerical examples are given in Chapter 4.

4. NUMERICAL COMPARISON AND SENSITIVITY

4.1 ASSUMPTIONS BASED ON ACTIVITY STATISTICS

The object of this chapter is to exhibit by means of graphs the numerical results of applying the three queue models of Chapter 3 to realistic situations. As in the 1969-1970 Interim Report (Chapter 6),[3] the starting point will be the statistics of activity (i.e., arrivals and departures) taken during one month of 1968 for J.F. Kennedy and LaGuardia.[4] These were supplied to us in a private communication from the FAA, as contained in an unpublished "FAA Staff Study of Airport Congestion in Major Metropolitan Areas." These data were in the form of graphs traced as follows: for each interval between consecutive hours (e.g., 10 and 11), the average number of events was found and plotted as the ordinate of a point, the abscissa of which was the corresponding midpoint of the interval (e.g., 10:30). The resulting 24 points were then joined by straight-line segments.

These activity data represent the most detailed and accurate ones bearing on arrivals and departures available to us. While rates of arrivals only are presumably about half as great as the total activity rates (on the average and at any particular time), the activity rates give a realistic idea of proportional degrees of variability during the course of the 24-hour day of the actual arrival rates. Therefore, in setting up the assumptions for our two "typical airports" we have taken as arrival rates $\lambda(t)$ for Airport A, the J.F. Kennedy activity rates, and for Airport B those of LaGuardia. Their graphs are shown in Figure 4.1.

The maximum number allowed in the queue is assumed in each case to be $m = 25$. It is obviously somewhat unrealistic to combine the large arrival rates introduced in the last paragraph with this rather limited landing facility. For example, a simple count shows some 60 spaces for stacking available to J.F. Kennedy. Nevertheless, the assumption is useful in bringing out most vividly the effects of congestion and the improvements obtainable by even moderate increases in service rates, as well as in flow control.

Three constant service rates are assumed at each terminal: $\mu = 45, 55$, and 70 aircraft per hour. These figures, as well as those of the preceding paragraphs, coincide with those used in the 1969-70 Interim Report. We are not, however, examining the fourth case considered there in which the service rate shows a sharp temporary dip: $\mu = 55$ from early morning to 1500; $\mu = 25$ between 1500 and 1700; $\mu = 55$ from then on. As set forth in the Interim Report cited above, the object was to show how such a temporary sharp degradation of service, produced for example by a sudden fog or similar deterioration of weather, could have all its effects on capacity, delays, etc., of the service of the terminal calculated quantitatively. The emphasis of the present report is rather on the comparison of effects of the various assumptions underlying our different models; this does not seem to require the examination of the case of the service dip.

4.2 THE GRAPHICAL EXAMPLES

Figures 4-1 through 4-13 are graphs of the most important functions, computed for each of the three models (random service, regular service, and deterministic flow) as applied to the data discussed above. In the case of Terminal A, the initial values (zero queue length at 4:00) was not exactly repeated 24 hours later: the solution is not yet periodic, but is getting close to it. In the case of Terminal B, the solution is, in fact, periodic: it was obtained by a very easy application of the relaxation method.

Figure 4-1 shows the two inputs: the arrival functions $\lambda(t)$. Figures 4-2 through 4-7 concern Terminal A and show for each of the three values of μ, the expected queue length, its standard deviation, and the probabilities of an empty and a full queue, at all times of day. Figures 4-8 through 4-13 give the corresponding results for Terminal B and the same three values of μ.

In all cases, the expected time from joining the queue at epoch t to the epoch of completion of service is the expected number in the queue at t, multiplied by $1/\mu$. Therefore, the graph of this expected waiting time as a function of t is obtained from the graphs for the expected queue length. It is to be observed, however, that an improvement in service time (i.e., an increase in μ) affects the waiting time in two ways: by decreasing the expected queue length, as well as in decreasing the factor $1/\mu$. However, the former effect far outweighs the latter in the important case when the conditions are varying rapidly, e.g., just before and after the rush hours. Moreover, this effect is highly complex – certainly non-linear – and it is a frequent fallacy to try to measure the diminution of delay by applying the same percentage improvement in waiting time as the percentage improvement in service time. This situation is examined in Chapter 6.

With regard to the expected number of aircraft turned away, this quantity was graphed on a cumulative basis in the 1969-70 Interim Report (Figures 6-4 and 6-9 of Chapter 6). In our present, purely comparative study, only the expected total number turned away during the 24-hour period is given, calculated as explained in Appendix C. It represents one "index of overloading" of the terminal to handle the traffic loads imposed.

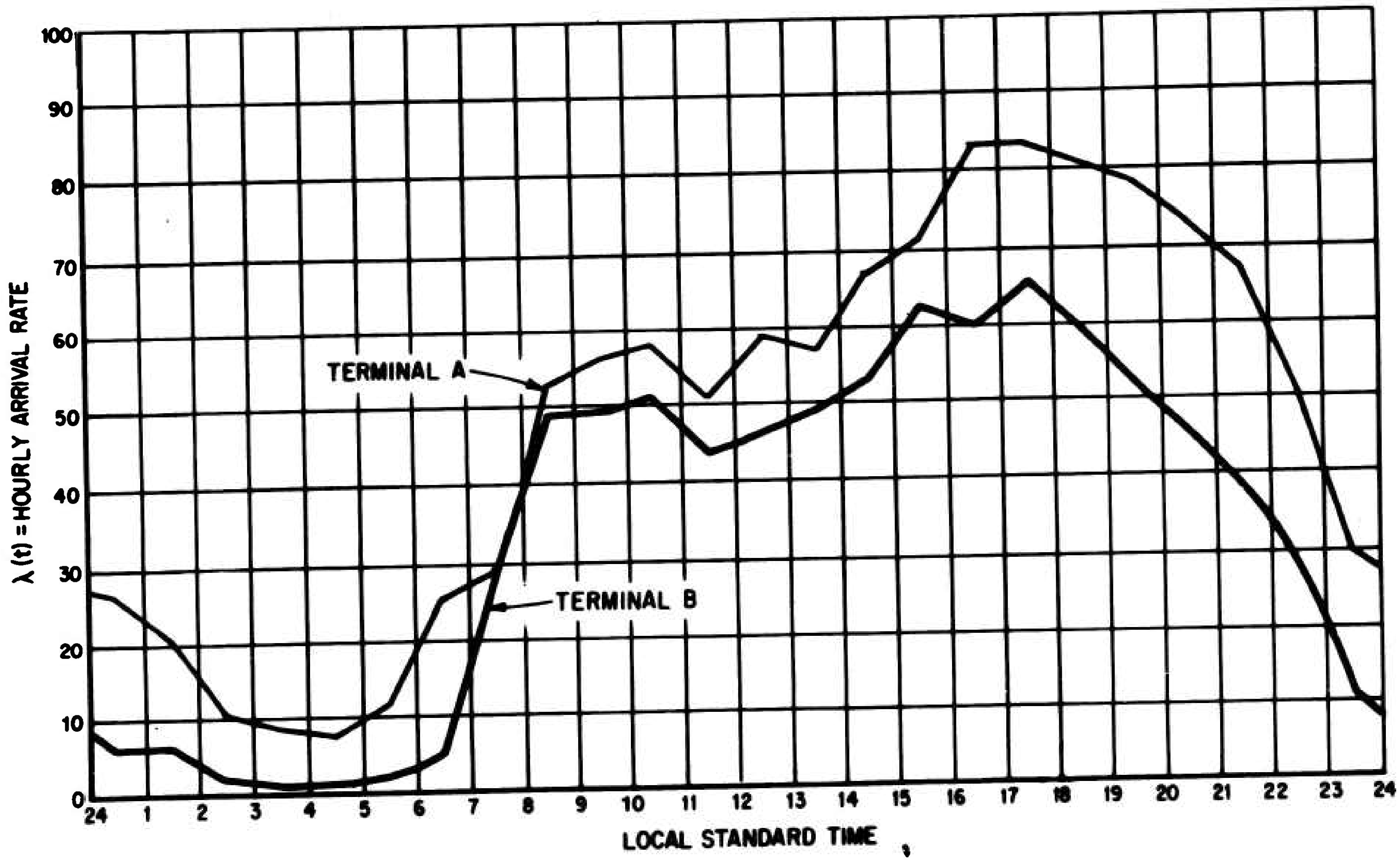


FIGURE 4–1 HOURLY ARRIVAL RATES - TERMINALS A AND B

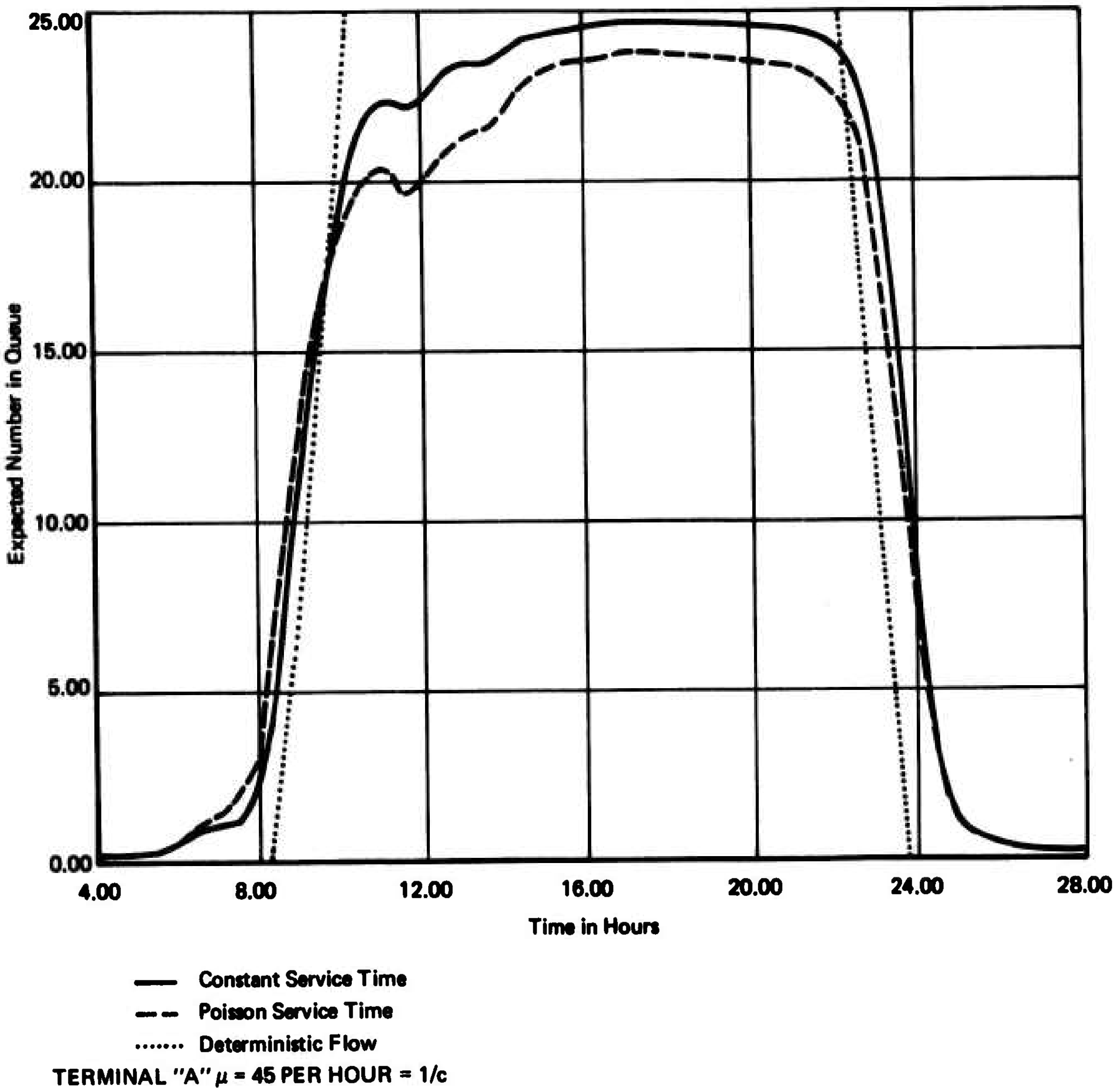


FIGURE 4-2 EXPECTED NUMBER IN QUEUE

FIGURE 4–3a STANDARD DEVIATION OF NUMBER IN QUEUE

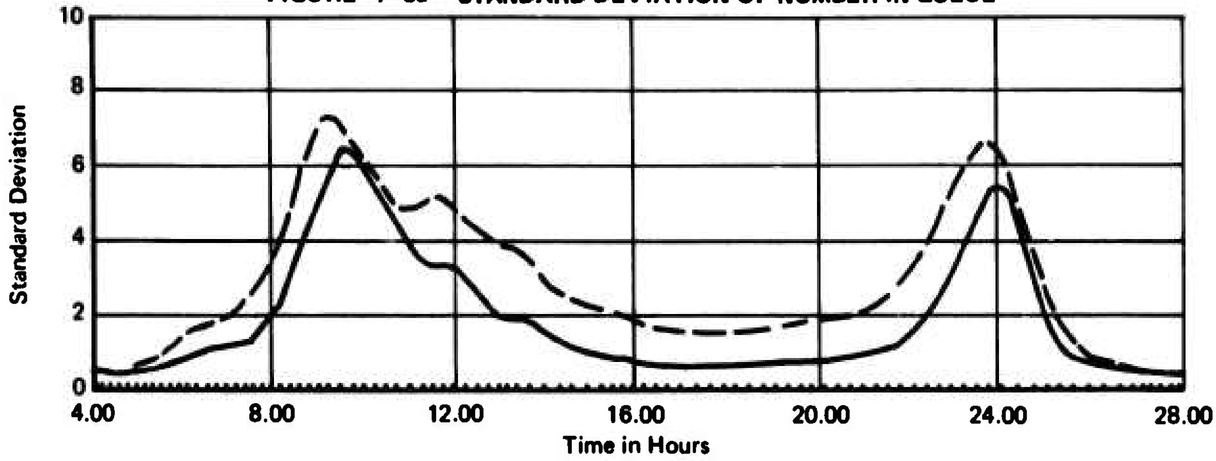


FIGURE 4–3b PROBABILITY OF EMPTY QUEUE

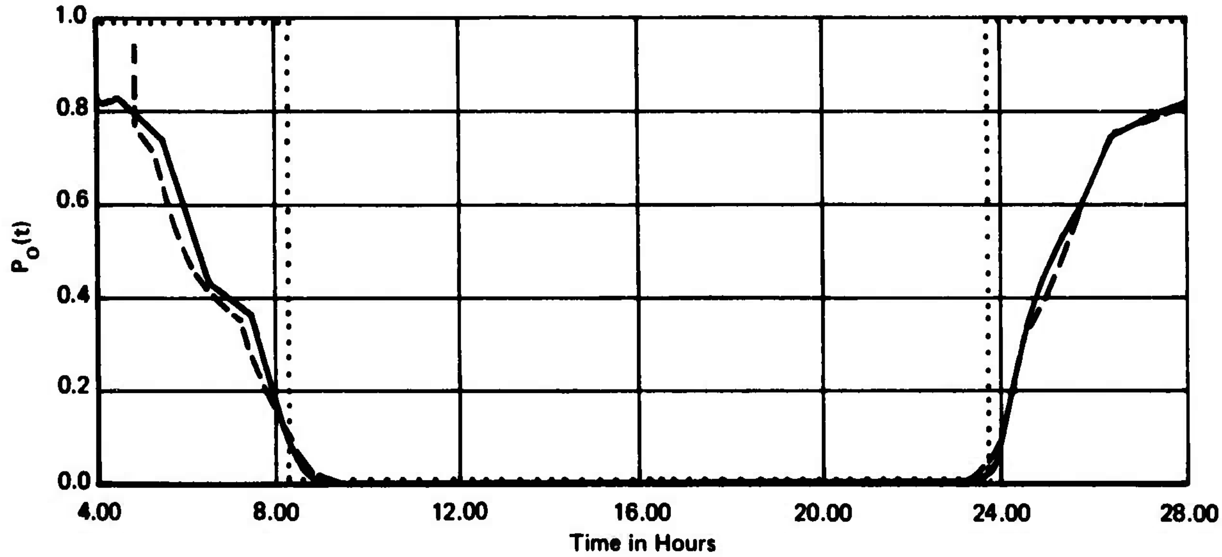
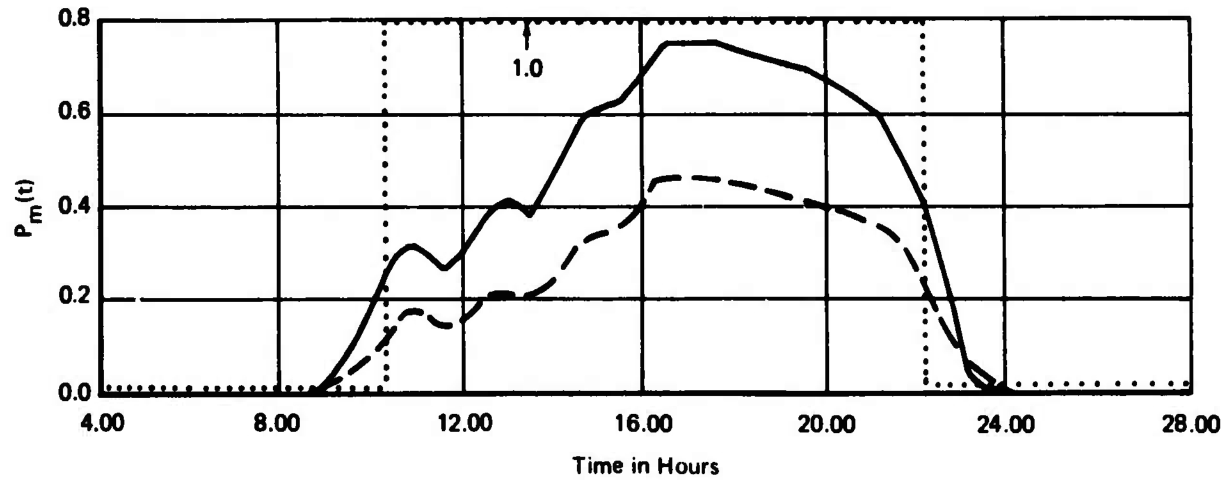
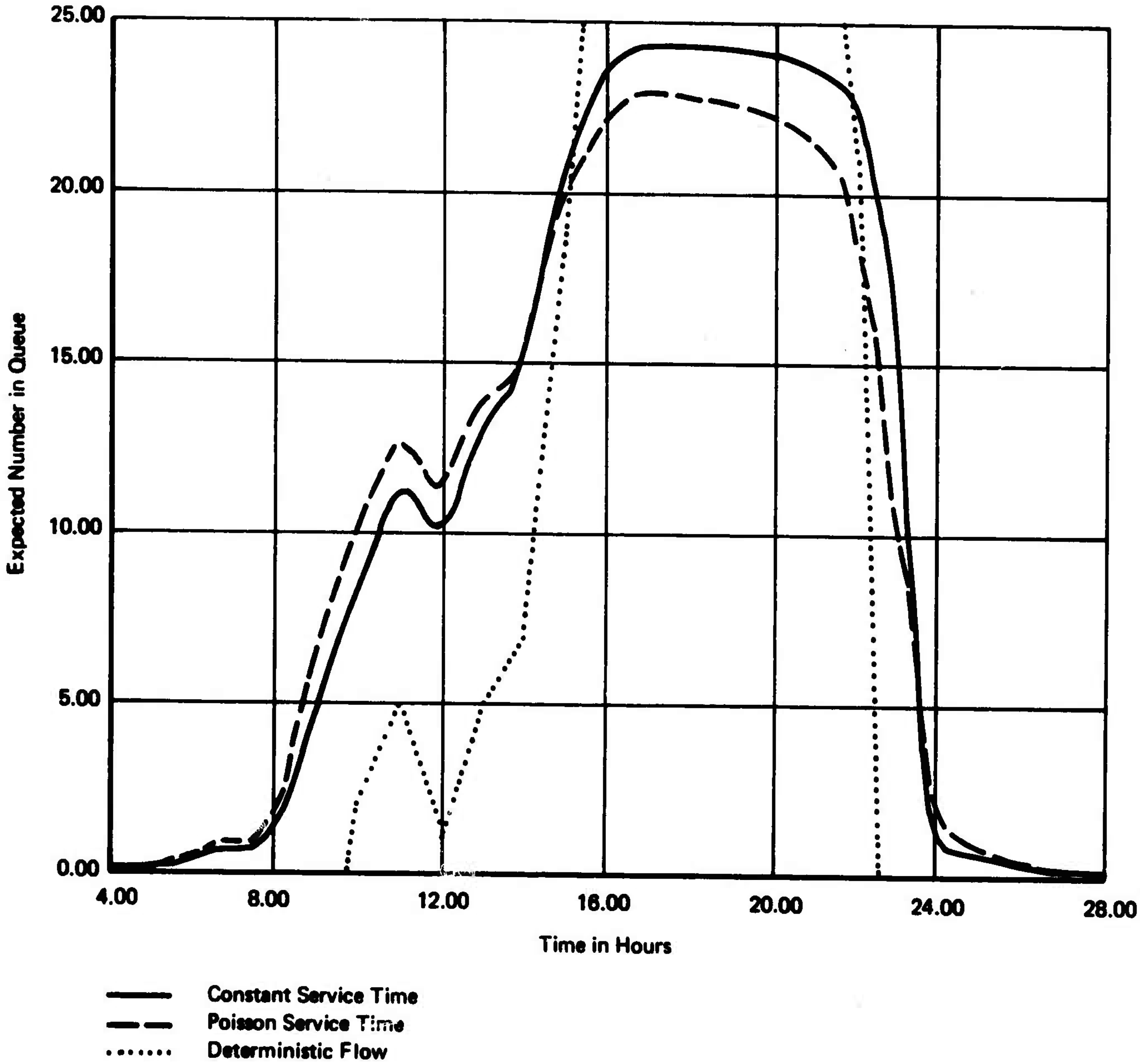


FIGURE 4–3c PROBABILITY OF FULL QUEUE



TERMINAL "A" μ = 45 PER HOUR = 1/c

—— Constant Service Time
– – Poisson Service Time
...... Deterministic Flow



TERMINAL "A" μ = 55 PER HOUR = 1/c

FIGURE 4–4 EXPECTED NUMBER IN QUEUE

FIGURE 4–5a STANDARD DEVIATION OF NUMBER IN QUEUE

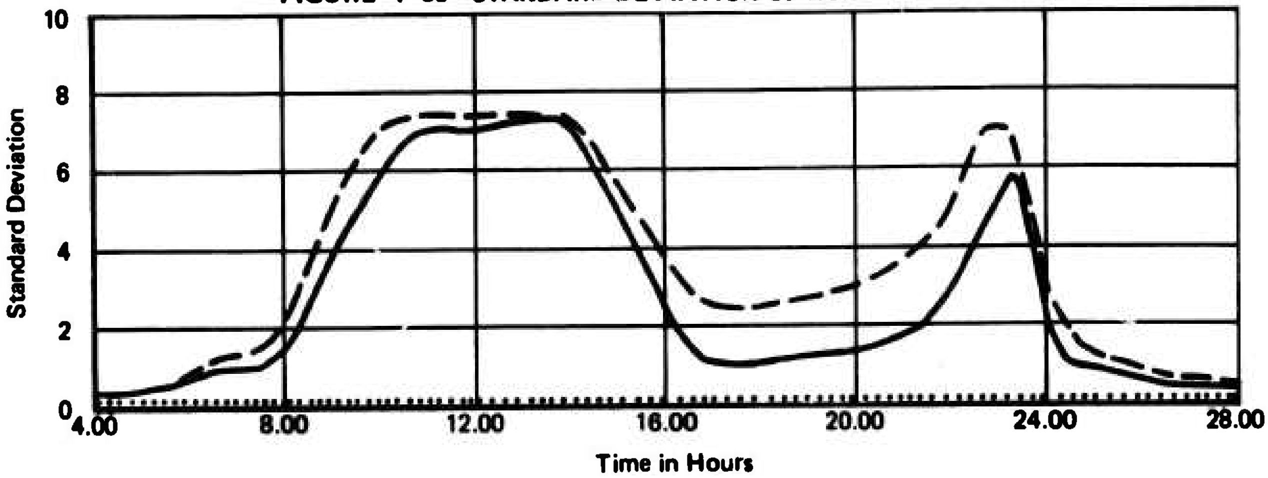


FIGURE 4–5b PROBABILITY OF EMPTY QUEUE

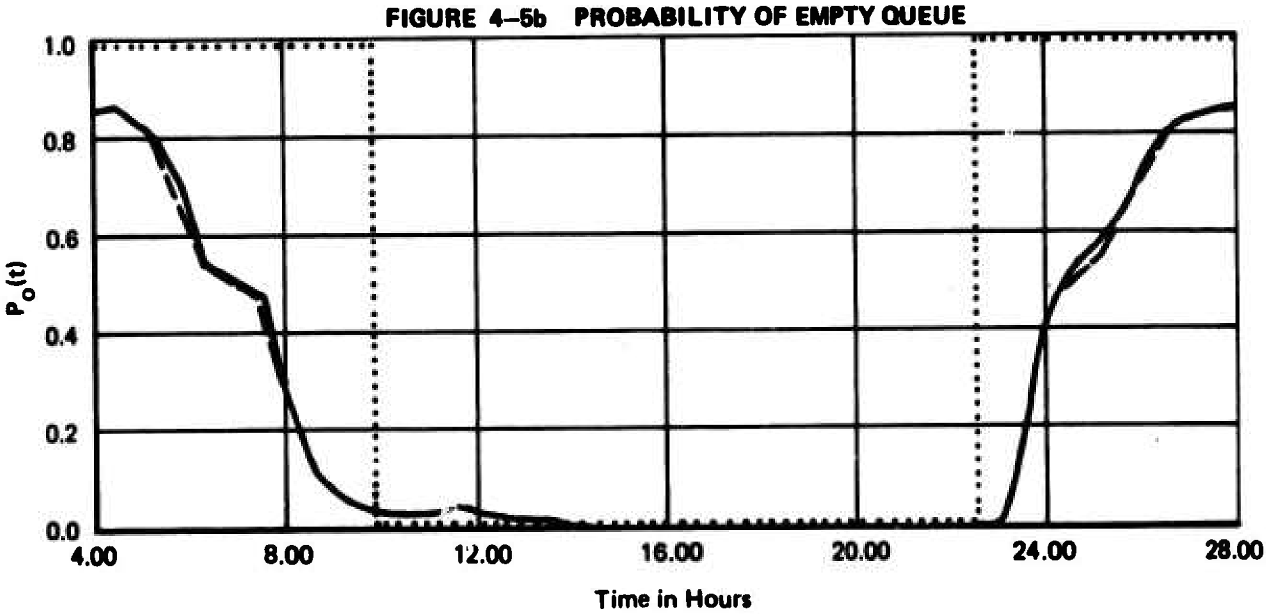
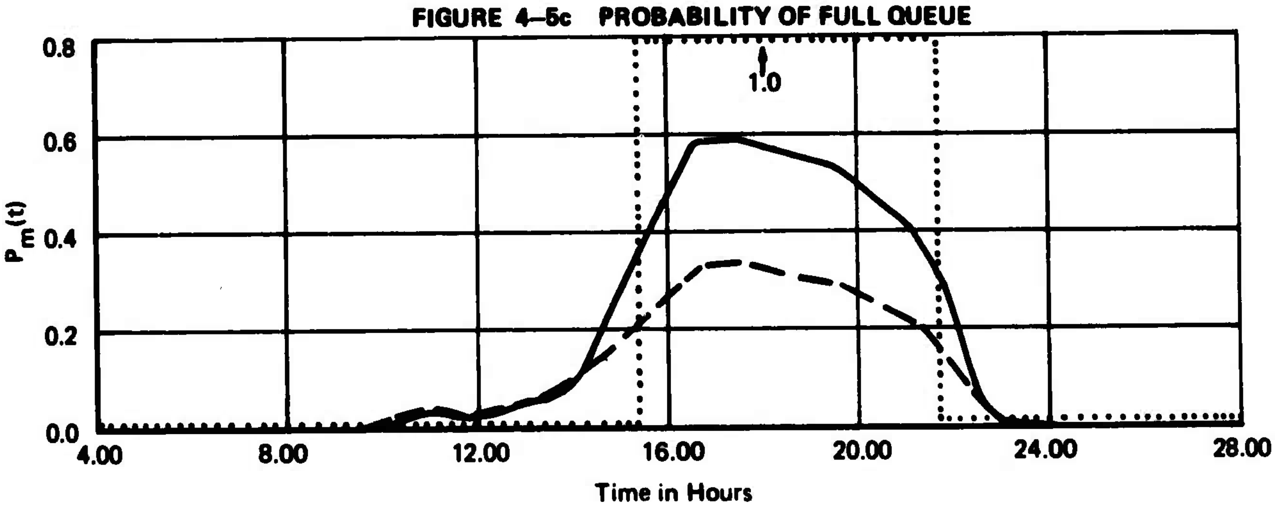


FIGURE 4–5c PROBABILITY OF FULL QUEUE



TERMINAL "A" μ = 55 PER HOUR = 1/c

———— Constant Service Time

– – – – Poisson Service Time

············ Deterministic Flow

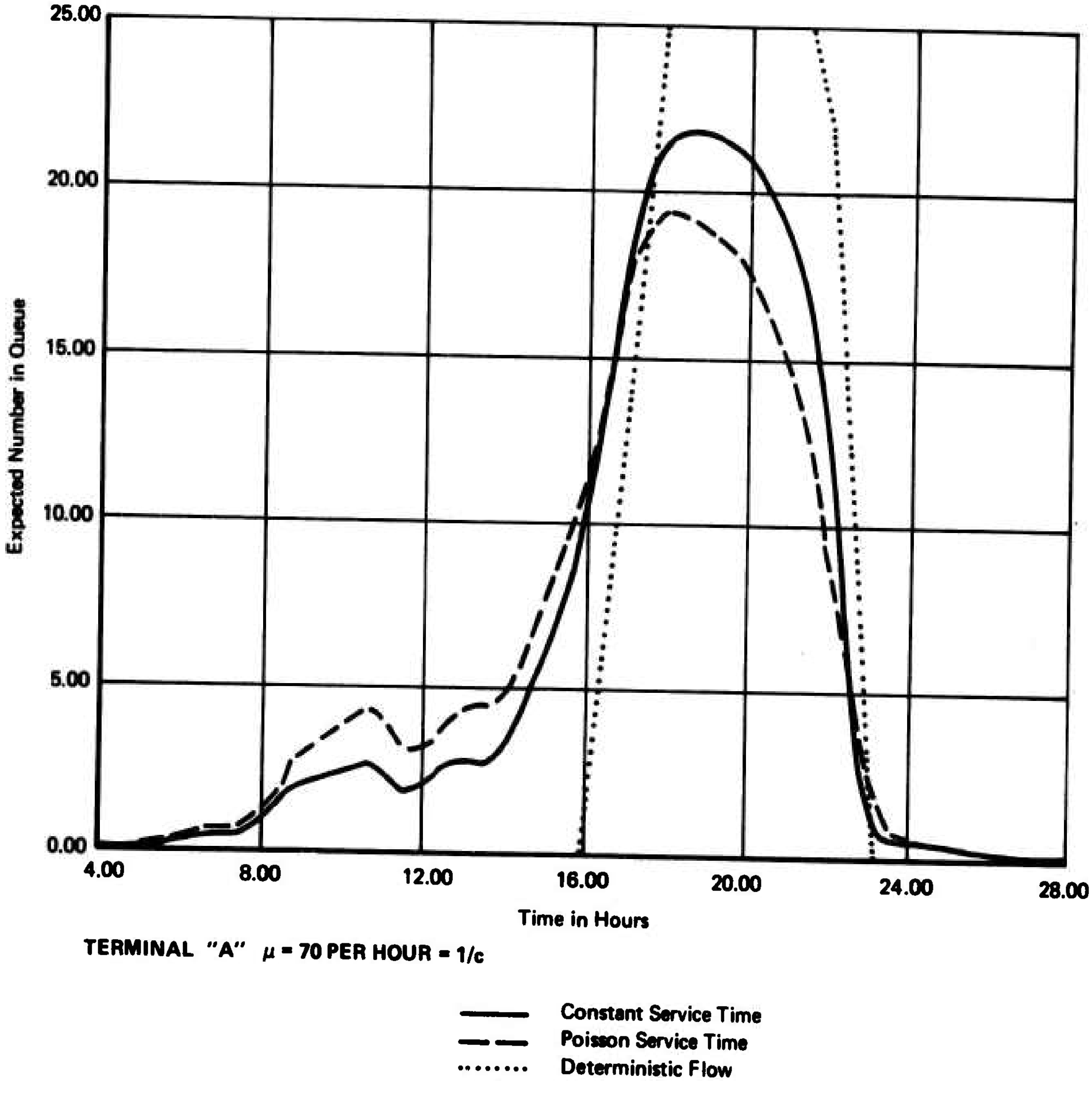
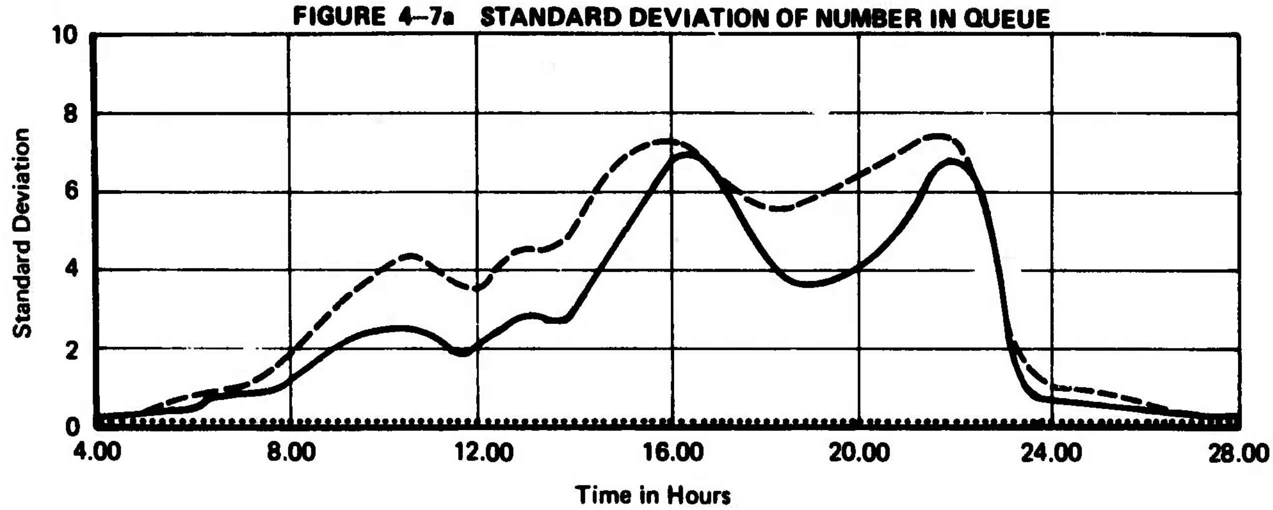
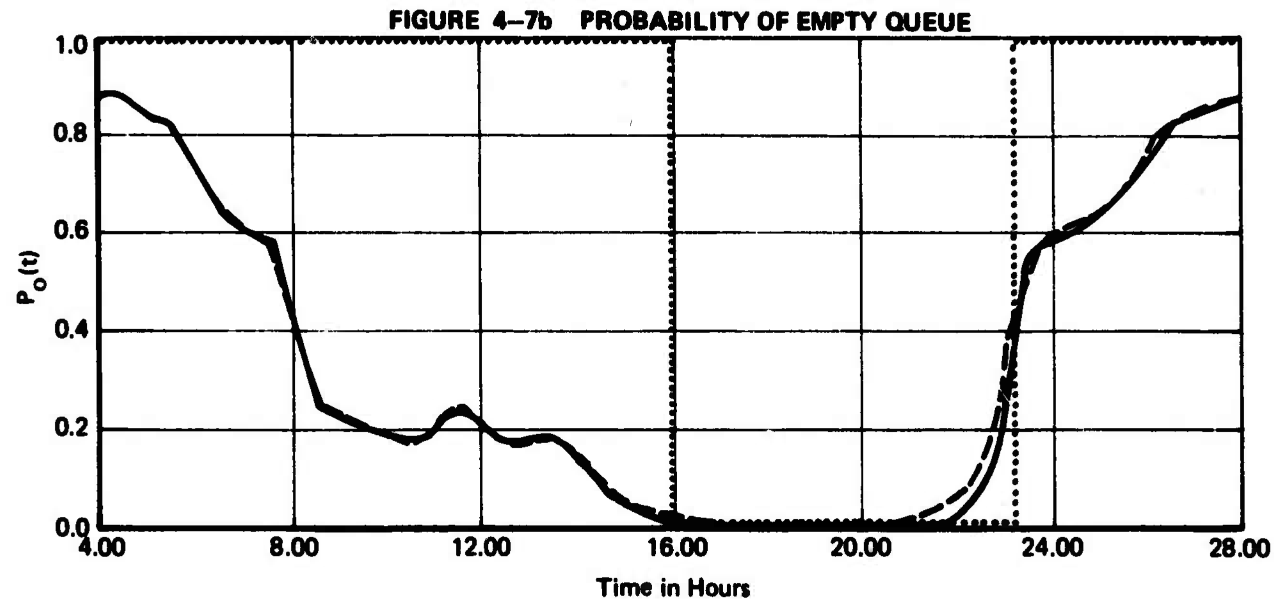
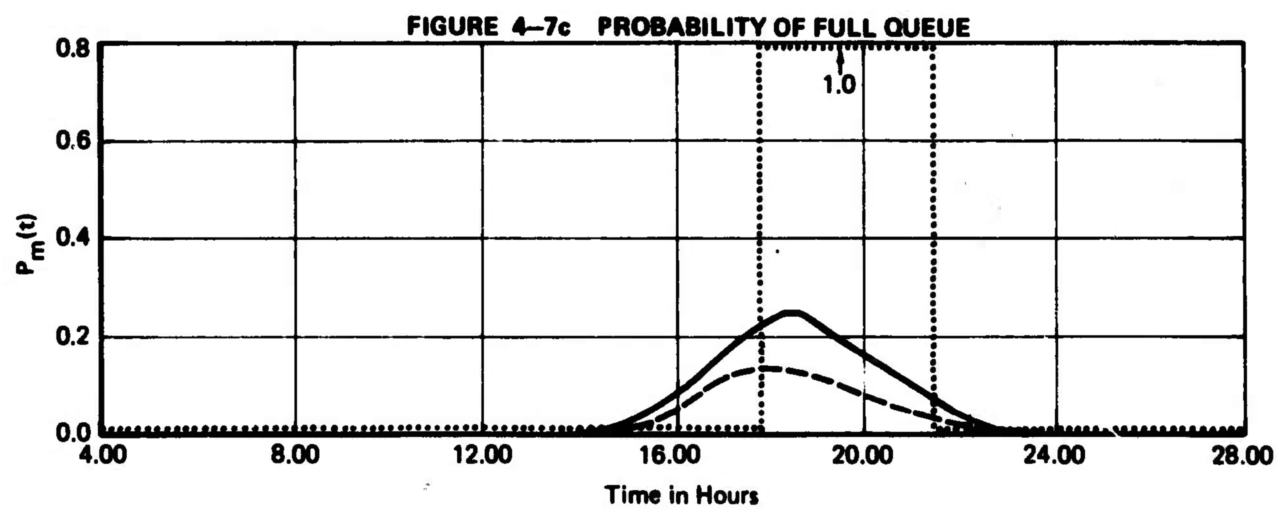


FIGURE 4–6 EXPECTED NUMBER IN QUEUE







TERMINAL "A" μ = 70 PER HOUR = 1/c

Constant Service Time
Poisson Service Time
Deterministic Flow

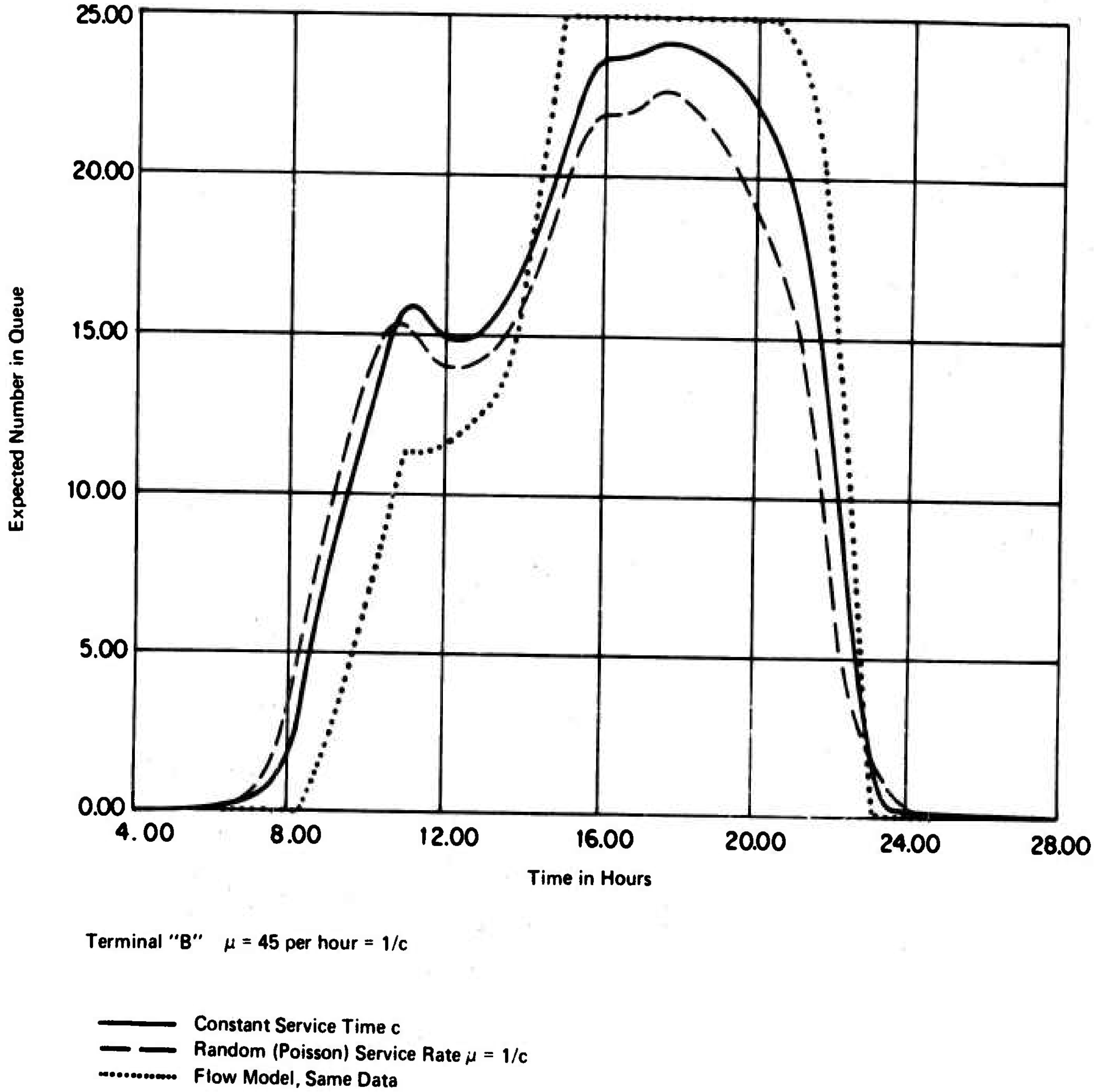


FIGURE 4–8 EXPECTED NUMBER IN QUEUE

FIGURE 4–9a STANDARD DEVIATION OF NUMBER IN QUEUE

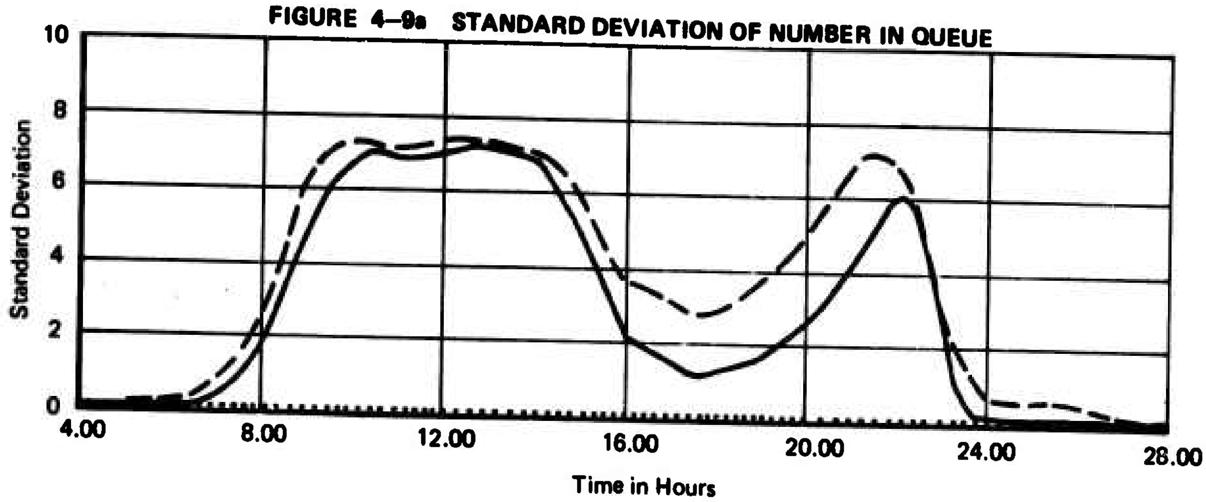


FIGURE 4–9b PROBABILITY OF EMPTY QUEUE

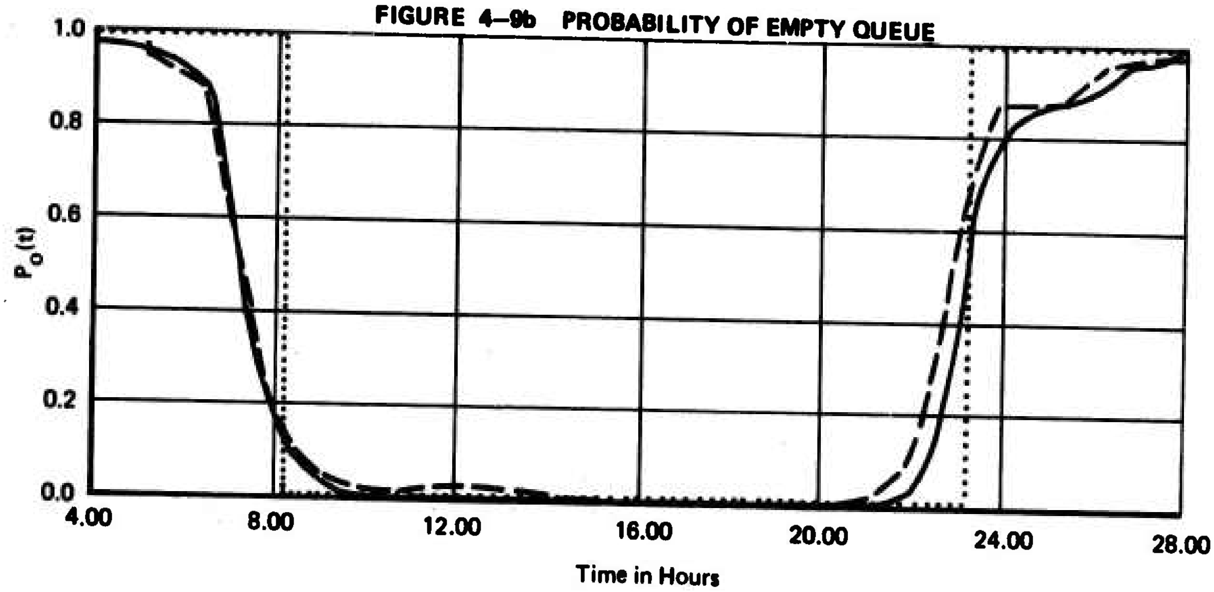
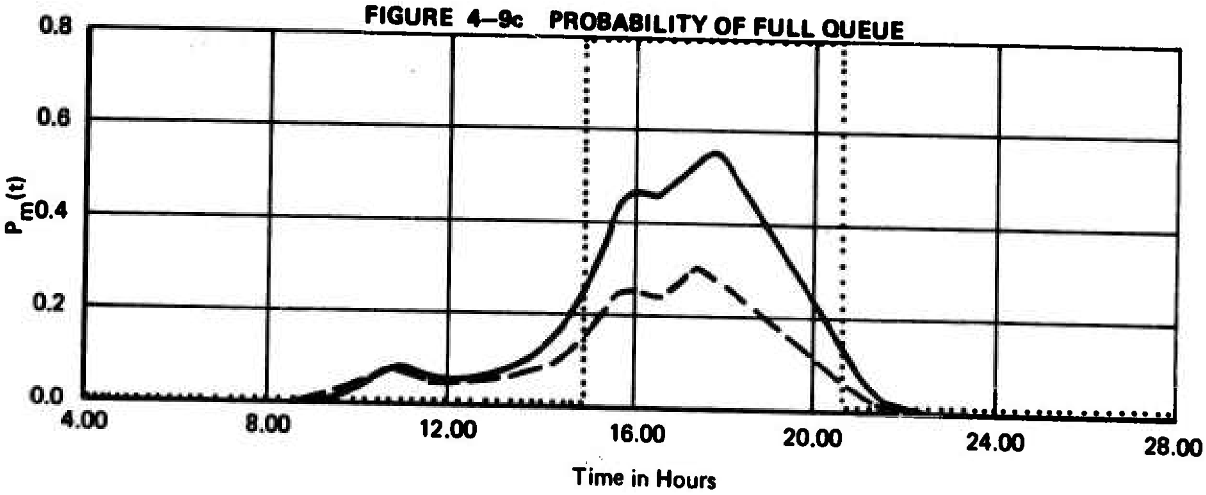


FIGURE 4–9c PROBABILITY OF FULL QUEUE



TERMINAL "B" μ = 45 PER HOUR = 1/c

——— Constant Service Time

– – – Poisson Service Time

.... Deterministic Flow

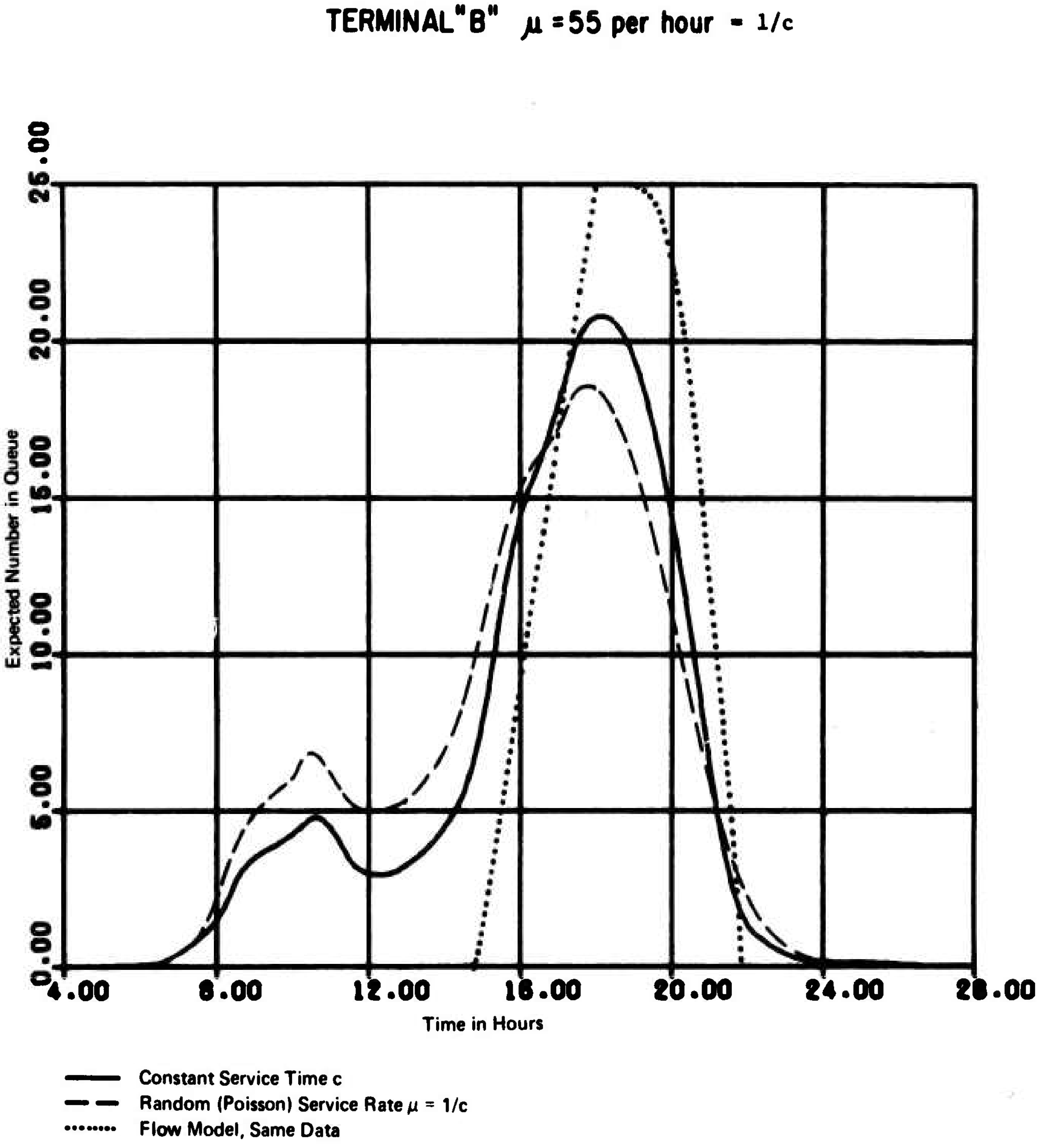


FIGURE 4-10 EXPECTED NUMBER IN QUEUE

FIGURE 11a STANDARD DEVIATION OF NUMBER IN QUEUE

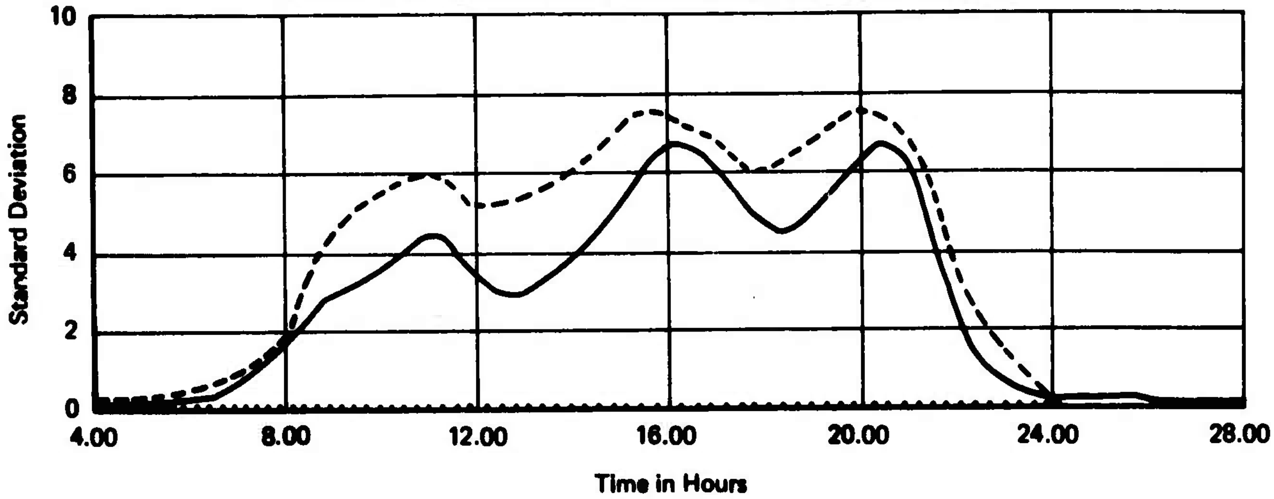


FIGURE 11b PROBABILITY OF EMPTY QUEUE

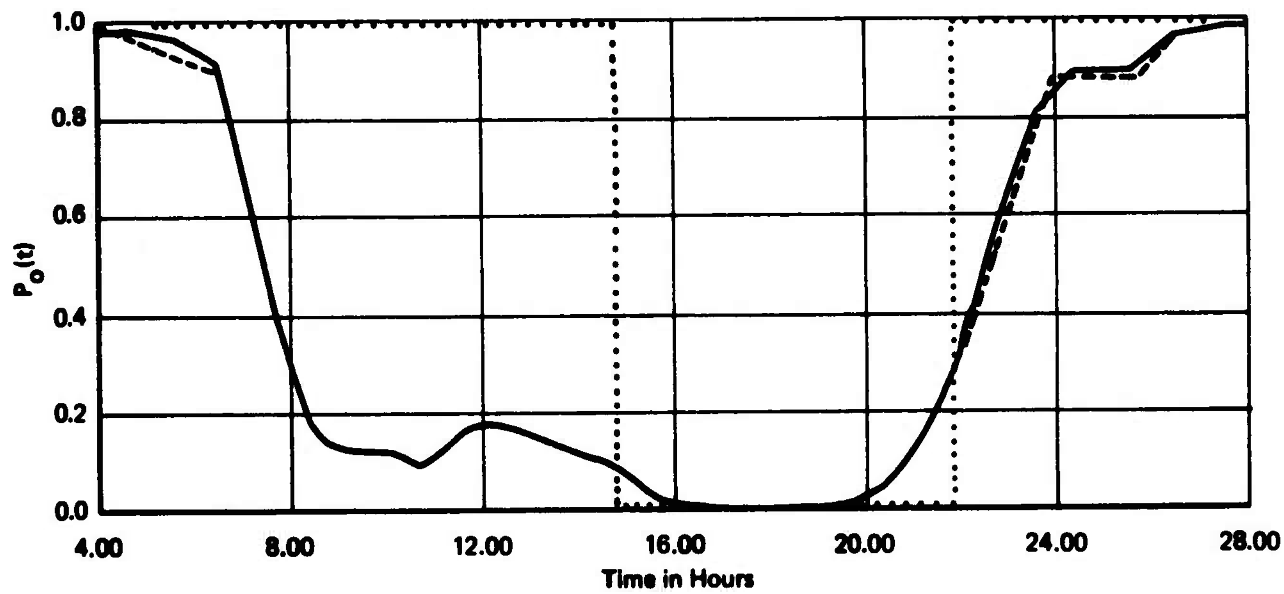
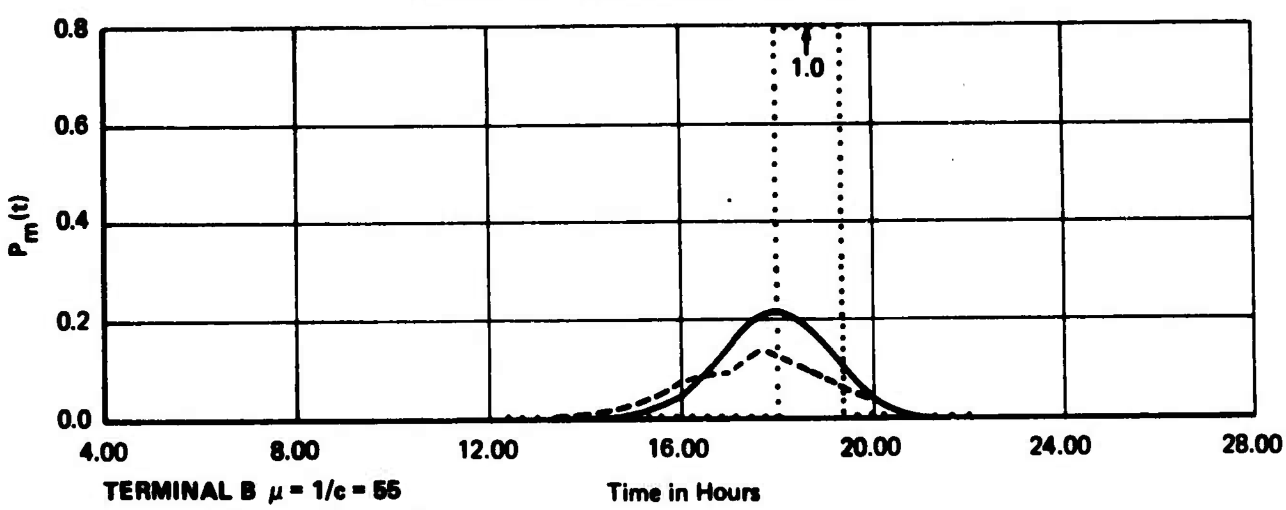


FIGURE 11c PROBABILITY OF FULL QUEUE

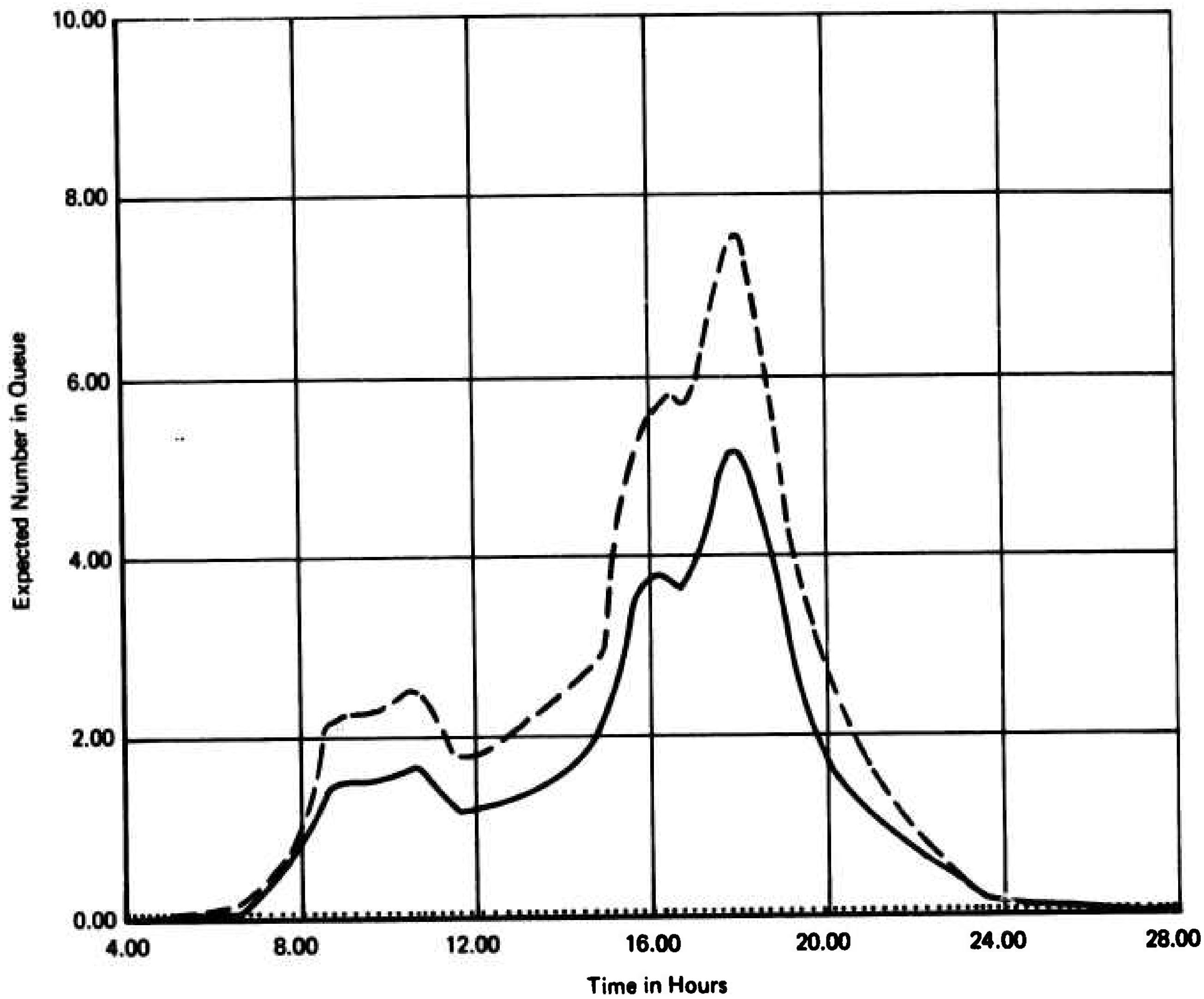


TERMINAL B μ = 1/c = 55

Constant Service Time

Poisson Service Time

Deterministic Flow



——— Constant Service Time c

— — Poisson Service Time

.... Deterministic Flow

TERMINAL "B" μ = 70 PER HOUR = 1/c

FIGURE 4–12 EXPECTED NUMBER IN QUEUE

FIGURE 4–13a STANDARD DEVIATION OF NUMBER IN QUEUE

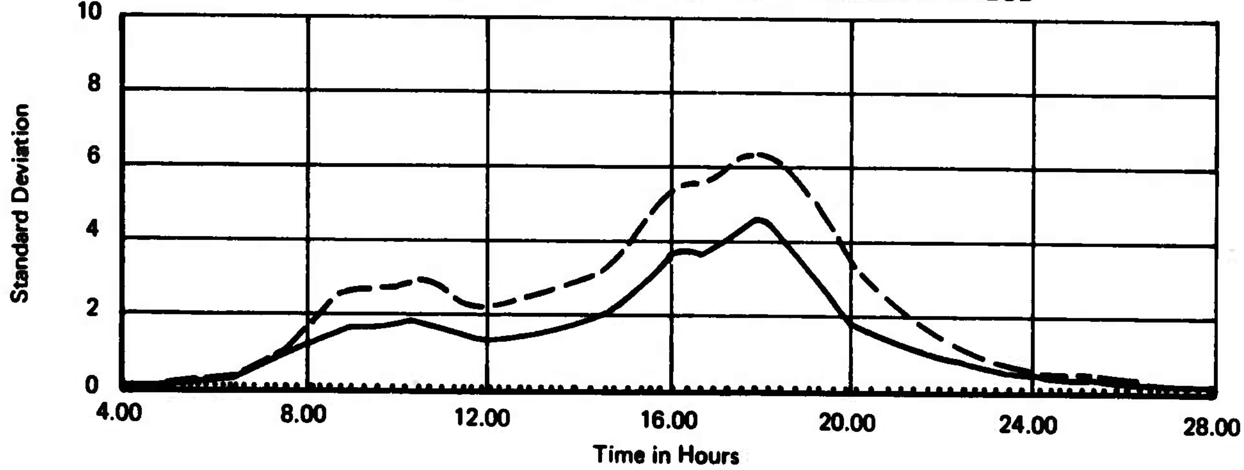


FIGURE 4–13b PROBABILITY OF EMPTY QUEUE

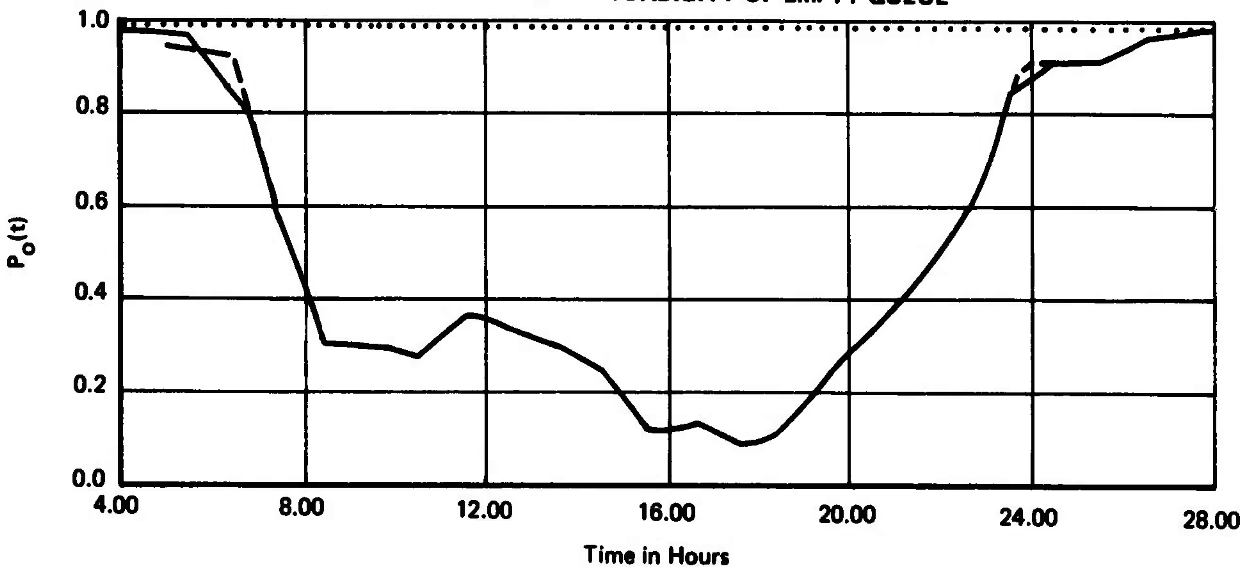
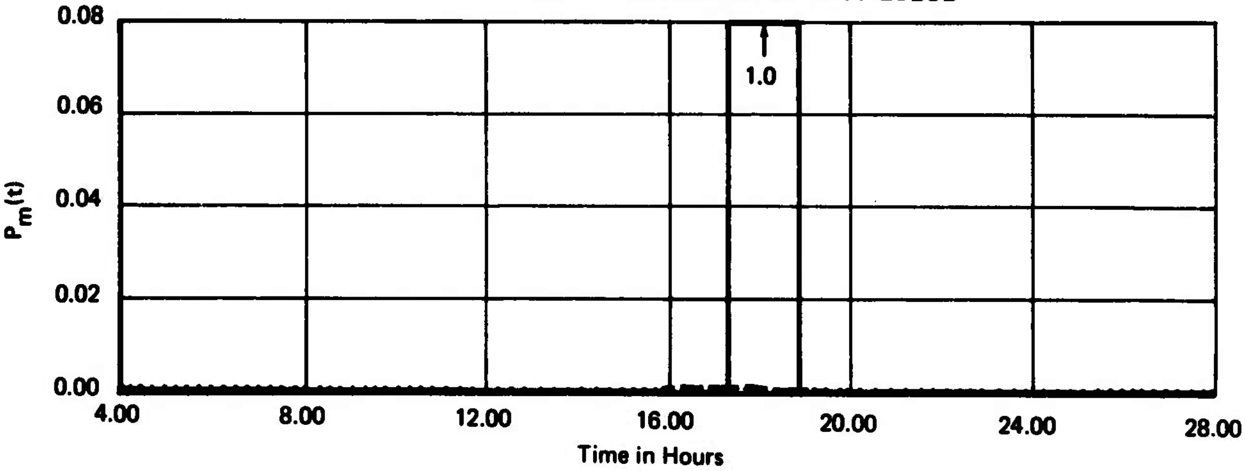


FIGURE 4–13c PROBABILITY OF FULL QUEUE



TERMINAL "B" μ = 70 PER HOUR = 1/c

—— Constant Service Time
– – Poisson Service Time
...... Deterministic Flow

5. MULTIPLE QUEUES

5.1. SIMULTANEOUS LANDING AND TAKEOFF USE OF RUNWAY

The situation to be studied is the very realistic one in which a single system composed of one runway or adjacent runways is serving for both takeoff and landing. In most air terminals at the present time the arrivals and departures are sufficiently infrequent so that there is no capacity bottleneck that cannot be easily removed by a little planning. In the highly important minority of "hub" terminals, on the other hand, very serious constraints on capacity – or, equivalently, serious delays – can be caused by the saturation of the runway system. The situation would be alleviated by having one runway system dedicated to landing and a different one to takeoff, the two being sufficiently separated and controlled to ensure no interference between any stage (air or ground) of these two operations. But this would imply either essentially distinct terminals, or a degree of control going well beyond what exists at the present time. It is necessary, therefore, to examine quantitatively the effects of having the air and ground queues complete in the use of the same runway system.

The practical value of such a quantitative examination is, of course, that the magnitude of the problem can be established under the various conditions of practice, and that the probable effects of improved procedures and equipment can be calculated in advance.

As in the case of the single waiting lines examined above, the time-dependent statistical law of arrivals has to be known. It will be assumed, as before, that the aircraft arrive for landing at the (Poisson) rate $\lambda = \lambda(t)$, and also that the aircraft arrive at the departure waiting line in the same way, but at the rate $\lambda' = \lambda'(t)$. The removals of aircraft from each of the air and ground queues by servicing (landing or takeoff) are, as discussed before, intermediate in their statistics between a fixed time and a random (Poisson) one. Fortunately, as we have shown in the preceding sections, the quantitative results are insensitive to the assumptions; accordingly, we shall choose the one that proves to be simpler – the random one. More precisely, we shall assume that the opportunities for runway use (clear for landing, or takeoff) occur at random rather than with perfect regularity, and that when such an opportunity opens up, and if an aircraft in the air is permitted to use it, it takes a random (Poisson) time to complete this operation, and similarly, when the permission is given for takeoff. The Poisson rates of service under the two assumptions will be denoted, respectively, by μ (landing) and μ' (takeoff).

Up to this point, the discussion shows a situation altogether similar to that of the single queue problem, solved completely above. There is, however, a new factor: the necessity of stating the "queue discipline," or rules for deciding which of the two classes of aircraft is to use the runway facility each time it becomes available, and which is to be delayed. It is a fact that the decision, which has to take account of a large number of conditions, is not amenable to embodiment in an absolutely precise set of rules, and reserves a large part to judgment. It is also a fact that, unless some reasonably realistic general assumption is

PRECEDING PAGE BLANK

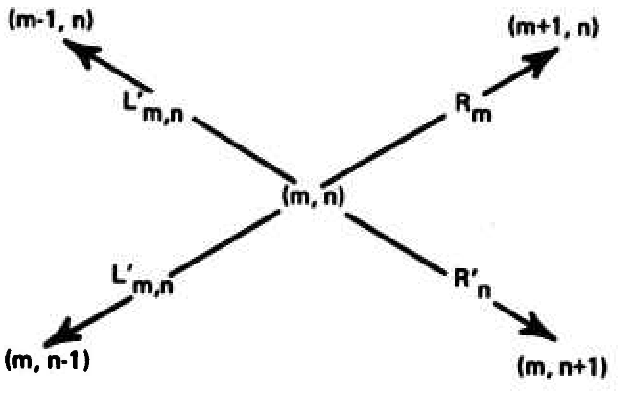
actually made, the behavior of the terminal cannot be calculated quantitatively by any mathematical technique, simulation, etc. This is a simple fact of logical indeterminacy.

To bring out the point concretely, suppose that an air terminal at some given time has 13 aircraft in the air stack and 17 in the ground waiting line, and that a space opens up on the landing strip. Will permission to use this space be given to a landing or to a departing aircraft? Clearly this can depend on other factors than the mere numbers (13 and 17) of aircraft in the two waiting lines. For example, if the policy is to take care of a group of planes together in one line first, and then to switch to the other, and if this is being done for the landing ones, the next few spaces may be given to the latter category; later a group of departing ones will have its turn; similarly for other factors affecting the situation. As a general matter, the mere numbers m and n in the air and ground queues do not in themselves determine the "state of the system" completely enough to allow the prediction of the next state into which it goes – or even its transition probabilities.

There is, however, another way of looking at the situation. Suppose that an air terminal is considered, not just on one occasion, but on a very large number of different occasions under similar conditions of weather, time of day, etc. Suppose that from this ensemble of samples a subensemble is drawn, for example, those cases in which the numbers in the two queues are $m = 13$ and $n = 17$. In this subensemble, a certain fraction would have the next open space assigned to the landing aircraft, while in the complementary fraction, to a departing one. In other words, this cross-sampling in the subensemble will give perfectly definite probabilities of evolution; and these are thus *determined* as functions of the numbers m and n (at the stated time t).

Such an idealized system with transition probabilities defined in this manner will be called a "cross-cut" model. It will not tell us very reliably what the probabilities of evolution of the two queues are on some particular day, in which a host of special items of information are at hand – information which can supplement what we know and alter the probabilities obtained by merely picking at random from our ensemble. But it can be highly useful as an *indicator* of the effects of various proposals for improving capacity and delay – on the average and in the long run.

From the point of view of the cross-cut model, the state of the air queues at time t is fully described by the numbers m and n in the air stack and ground waiting line, in the sense that the transition probabilities are all determined. We write for the probabilities of *decrease* of m to $m - 1$ during the time interval $(t, t + \Delta t)$ $L_{m,n}\Delta t$. Similarly, the probability of decrease of n to $n - 1$ is denoted by $L'_{m,n}\Delta t$. As explained earlier, the probabilities of *increase* of m or of n (arrivals to the landing or takeoff lines) are $R_m \Delta t$ and $R'_n \Delta t$ (given in terms of the arrival functions of the time $\lambda(t)$ and $\lambda'(t)$, as we shall see) – all this, to terms of higher order in Δt.



In all, we have the diagram of right-transitions R (increases) and left-transitions L (decreases). These are the only transitions during t, $t + \Delta t$ having probabilities of the order of Δt. All others would be multiple events and have transition probabilities of orders Δt^2 or higher. All these transition probabilities are functions of the time t, m, and n and vary strongly.

5.2 EQUATIONS OF EVOLUTION OF THE DOUBLE QUEUE

Inasmuch as all that can be told in general concerning the double-queue system is in terms of probabilities and averages, we introduce $P_{m,n}(t)$, the probability that at the time t there be m and n aircraft in the air and the ground queues. The differential equations determining the evolution of the system are obtained by the usual simple probability reasoning. For general values of m and n, we express the probability $P_{m,n}(t + \Delta t)$ in terms of those at time t: Neglecting quantities of higher order in Δt, we can say that if the system is in state (m, n) at $t + \Delta t$, this could only have occurred in the following five distinct ways:

1. It was in state (m, n) at t and no transition occurred during the period Δt. Probability:

$$P_{m,n}(t)\,[\,1 - (R_m + R'_n + L_{m,n} + L'_{m,n})\Delta t]$$

2. It was in (m−1,n) at t and gained a member to the air queue. Probability:

$$P_{m-1,n}\,(t)\,R_{m-1}\Delta t$$

3. It was in (m, n−1) at t and gained a member to the ground queue. Probability:

$$P_{m,n-1}(t)\,R'_{n-1}\Delta t$$

4. It was in (m + 1,n) at t and lost a member of the air queue. Probability:

$$P_{m+1,n}(t)\,L_{m+1,n}\Delta t$$

5. It was in (m, n + 1) at t and lost a member of the ground queue. Probability:

$$P_{m,n+1}\,(t)\,L'_{m,n+1}\Delta t.$$

On putting these facts together, an equation, correct to terms of order Δt, is obtained which, on slight rearrangement of terms, division by Δt, and then taking limits as $\Delta t \longrightarrow 0$, gives rise to the exact differential equation

$$
\begin{aligned}
dP_{m,n}(t)/dt = -(R_m + R'_n + L_{m,n} + L'_{m,n})\, P_{m,n}(t) \\
+ R_{m-1}\, P_{m-1,n}(t) \\
+ R'_{n-1}\, P_{m,n-1}(t) \\
+ L_{m+1,n}\, P_{m+1,n}(t) \\
+ L'_{m,n+1}\, P_{m,n+1}(t)
\end{aligned}
\qquad (5.1)
$$

Finally, we impose, as usual, the restriction to the numbers of aircraft in each queue, and write

$$0 \leqslant m \leqslant M, \; 0 \leqslant n \leqslant N \qquad (5.2)$$

In the differential equation (5.1) we must evidently set

$$
\begin{array}{lll}
R_M = 0 & \text{and } R'_N = 0 & \\
R_m = \lambda(t)\ (m < M) & \text{and } R'_n = \lambda'(t)\ (n < N) & \\
L_{0,n} = 0 & \text{and } L'_{m,0} = 0 & (5.3) \\
P_{m,n}(t) = 0 & & (5.4)
\end{array}
$$

whenever one or both m, n are negative, or exceed M, N (respectively).

5.3 REDUCTION BY MARGINAL SUMMATION

Equations (5.1) already represent a departure from – or rather, a condensation of – reality, inasmuch as the average evolution of the air terminal has replaced its detailed probabilities of evolution. As stated above, what this "cross-cut" model loses in detailed information it gains in general relevance. Although it represents a simplification, it is still too complicated for numerical computation, except when M and N are both small (for example, under 10). This is because the total number of unknown functions $P_{m,n}(t)$ is $(M + 1)(N + 1)$. In cases such as those considered in Chapter 3, with even moderately modest values such as M = N = 25, we should have 676 unknowns with almost five times as many coefficients in the equations, given as tabulated functions of the time.*

The form of the equations (5.1) is that of (m + 1) (n + 1) linear differential equations of the first order, homogeneous, and with coefficients given functions of the independent variable. Furthermore, they are stochastic equations: non-diagonal coefficients non-negative, with each column [i.e., for each (m,n)] adding up to zero. Therefore they possess a unique solution for each set of initial values; if the latter are probabilities (non-negative and adding up to unity), the same will be true of the former. Furthermore, if the coefficients are periodic functions of the time of a 24-hour period, the same will be true of just one of the solutions, all others approaching it with increase in time. These facts are all proved along the lines of the same ones in the cases of Chapter 3.

We now reduce equations (5.1) by the method of "marginal sums." On setting

$$P_m(t) = \sum_{n=0}^{N} P_{m,n}(t) \tag{5.5}$$

it is realized that $P_m(t)$ represents the probability that at time t there be exactly m aircraft in the air stack – nothing stated with regard to how many may be in the ground queue. If we had summed over m instead of n, we would have obtained the corresponding probability for n aircraft in the ground queue at t; and of course the methods which we shall illustrate in the case of the air stack will be applicable to the ground queue.

Thinking of equations (5.5) written for a fixed m but for all n = 0, 1, ..., N, and then added, we find, after using (5.3) and (5.4) and making certain rather evident formal manipulations the following equations for $P_m(t)$:

$$\frac{dP_m(t)}{dt} = R_{m-1} P_{m-1}(t) - (R_m + L_m) P_m(t) + L_{m+1} P_{m+1}(t) \tag{5.6}$$

where the new quantity L_m is defined as follows

$$L_m = \sum_{n=0}^{N'} L_{m,n} [P_{m,n}(t)/P_m(t)]. \tag{5.7}$$

*Any attempt to take the model, the postulated behavior of which is represented by (5.1), and to use it as the basis of a Monte Carlo simulation should be preceded by a cost-effectiveness comparison of such a use of machines with their use in solving (5.1) directly – at an acceptable level of reliability of the answers; i.e., a level suitable for decisions involving large expenditures for improving the air transportation system.

In this reduction, use is made of the fact that the pair of sums

$$-\sum_{n=0}^{N} R'_n P_{m,n}(t), \qquad \sum_{n=0}^{N} R'_{n-1} P_{m,n-1}(t)$$

cancel, since, in view of (5.3) and (5.4), the last term in the former and the first term in the latter are zero; the others are equal and opposite in sign. A similar cancellation takes place for the sums

$$-\sum_{n=0}^{N} L'_{m,n} P_{m,n}(t), \qquad \sum_{n=0}^{N} L'_{m,n+1} P_{m,n+1}(t).$$

It is now necessary to interpret the expression L_m given by (5.7). Clearly the ratio $P_{m,n}(t)/P_m(t)$ is a conditional probability: the probability that, at time t, there be n aircraft in the ground queue, given that the number in the air is m. Therefore L_m is a "weighted mean": strictly, the conditional expected value of landing rate, given that the number in the stack is m.

In formulas (to terms of order Δt),

$L_m \Delta t$ = probability of a landing during $(t, t + \Delta t)$, given that number awaiting landing is m but that nothing is given concerning the number waiting on the ground.

Evidently L_m depends not only on the time t, but on the whole solution $P_{m,n}(t)$ (for all m, n, t), and therefore, strictly speaking, (5.2) cannot be solved without fully solving the original equation (4.1). However, while L_m may have this degree of complication, it is not necessary to find it exactly, as will now be explained.

5.4 PERTURBATIONS AND SUCCESSIVE APPROXIMATIONS

As things have been left, the reduced equations (5.6) reflect the influence of the condition of the takeoff queue only through the quantity L_m: to use the language applied to loosely coupled physical systems, the latter queue *perturbs* the queue we are studying by causing L_m to vary. Simple reasoning indicates its order of values: during dull periods, there is little interference between the two uses of the landing strips: then $L_m = \mu$. During busy periods, when both λ and λ' are large, turns have to be taken and, on the average, the same number will take off as will land. Since the mean times taken for these operations are $1/\mu'$ and $1/\mu$, in every interval $1/\mu' + 1/\mu$ there will be one landing on the average at the rate L_m. Therefore L_m is the harmonic mean: $1/L_m = 1/\mu + 1/\mu'$. What this means is that if the operating procedure aims at avoiding a greater average accumulation of aircraft in one queue than in the other, the same number, on the average, will have to be serviced per hour: the expected rates will be the same as if they were serviced in strict alternation, giving the above harmonic mean for both L_m and for L'_m. This assumes that at each time of day, t, the arrival rates $\lambda(t)$ and $\lambda'(t)$ are approximately equal. In the contrary case, an easy argument shows that we must take the weighted harmonic mean, and use the formula

$$\frac{\lambda(t)}{L_m} = \frac{\lambda(t)}{\mu} + \frac{\lambda'(t)}{\mu'} = \frac{\lambda'(t)}{L'_n} \tag{5.8}$$

To implement a calculation based on these ideas, it is necessary to have a rule for deciding on the "dull periods" and the "busy" ones. One obvious method of doing this starts from the use of the notions in the deterministic flow model discussed in Chapter 3.5. Thus we can regard as a dull period that time during which the inflow $\lambda(t)$ (rate of arrival into the air queue) is less than the maximum possible outflow L_m, as determined by (5.8); i.e., when the available landing rate exceeds the arrival rate, even when the runway is being shared. Then we may use $L_m = \mu$. When, on the other hand, the arrival rate into the air queue equals or exceeds that of the available landing rate, L_m must be computed from (5.8). One can see at once from that equation that there will be a dull period or a busy one, depending on whether its middle member, $\lambda/\mu + \lambda'/\mu'$, is or is not less than unity.

The scheme of computation just described represents an approximation that may be considered quite crude to the actual solution of the exact equations (5.1). But, as is usual with such perturbation methods, it may be made the starting point of a process of successive approximations.

In the present case, we shall indicate merely the next step: Using the approximations to the quantities L_m and L'_n obtained above, solve the single-queue equations (5.6) for $P_m(t)$ and the corresponding ones for the ground-queue marginal sum probabilities $P'_n(t)$. Next, from the approximation $P_{m,n}(t) = P_m(t)\,P'_n(t)$, recalculate L_m from (5.7), which now simplifies, since the ratio of probabilities reduces to $P'_n(t)$. The last step is to solve (5.6) again, now using this improved L_m. Higher levels of approximation may be unable to accept the above product expression for $P_{m,n}(t)$. In view of the rather crude nature of the input data it is questionable whether such further search for numerical accuracy is relevant to the practical problems.

5.5 MULTI-TERMINAL SYSTEMS

Up to this point, our attention has been confined to congestion at a single air terminal, and the analytical techniques concerned with single or double and interacting queues at this terminal. All other air terminals had their effect on our terminal through the arrival rate $\lambda(t)$. However, in view of the fact that air terminals in a given region (e.g., the Northeast Region, the Golden Triangle, etc.) exert a more direct effect upon one another than merely as a source or withdrawal point of a stream of aircraft, it is necessary to take such effects into account in any more realistic study of the total capacity of the air transportation system.

One method, suggested by the results of the present chapter, would be to express the situation in terms of queues at the different terminals that are weakly interacting with one another. In this context, "weakly interacting" means that, whereas the detailed queue state at one terminal has no effect on the other, the degree of its burden (as expressed, i.e., by the expected value of its length, etc.) will have a decided effect on the queue under study. The mechanism of this effect is, of course, the "feedback" of the air traffic control, which will delay on the ground or divert aircraft destined for the terminal in question.

Flow models, refined by the stochastic ones of the sort examined in this report, can be used for the study of such phenomena. It is noted that, whereas in using the perturbation methods of the last section the service rate L_m was regarded as exhibiting the effect of the competing queue, in the multi-terminal studies, on the other hand, it is the arrival rate $\lambda(t)$, as manipulated by the distant terminal in response to growing saturation at the one under study, that mediates their interaction.

Because of the finite time of travel, there may be a delay of the effect (as in the case of "retarded" potentials in electromagnetism). Since such delays may be deleterious to the even flow of traffic between terminals, any method for predicting growing queues will, if used in the feedback loop, regularize the flow and hence improve capacity.

6. PRACTICAL APPLICATIONS

6.1 BASIC PRELIMINARIES

In concluding this report on analytical tools for the study of airport congestion, it is appropriate to illustrate their practical use by means of examples. To this end, we take two factors in the development of the congestion, capable of being improved by technical advances in air traffic control systems, such as ARTS-III in its present or later forms. These factors are the rate of delivering service μ (e.g., landing) and the regularizing possibilities of flow control, mentioned earlier.

In applying any analytical tool to the study of a system, there is always the question of what quantities should be examined to determine the worth of technical improvements: what are the valid criteria of desirable performance, and by what *indicators* can they be measured? These usually take the form of precise numbers or curves that not only can be computed by the analytical tools, but "tell a story." In the present case, the expected queue lengths at various times have practical meaning chiefly in indicating the probable lengths of time spent in the queue before landing. Another indicator of significance is the total number of flights during the 24-hour day that have to be refused admission to the terminal because its facilities are saturated. It is not necessary to suppose that they actually come within the physical limits of the terminal before being turned away: the refusal may take the form of cancelled or diverted flights or long waits at the ports of origin. Nevertheless, the "models" – the somewhat rigid and simplified fictitious situations – examined in Chapters 3 and 4 supply very useful indicators of quality of service before and after the technical improvements under consideration.

There is a rather precise *"complementarity"*: On the one hand, the *systems engineering* which, by considering the technical features of the ATC system, gives as output such parameters as the rates μ and the degree of regularity of arrivals. On the other hand, the *operations research* takes the outputs of the former as its own inputs, and by a quantitative study of the operation, delivers as output the answer in terms of the indicators of performance.

6.2 INCREASE OF SERVICE RATES

Suppose that, by a tighter degree of control of the aircraft arriving at a terminal, or by other means, the rate of landing μ is increased by some percentage, e.g., 10%; that is, μ is replaced by $h\mu$, where h is somewhat greater than unity (in the case suggested, $h = 1.1$). What effect will this improvement have on the expected time in the landing queue (as measured by its expected length), and what will its effect be on the total expected number of flights diverted?

Before going into the mathematical details, three facts can be stated regarding the effect of the technical improvement in question:

First, if the situation is well in hand before the improvement (short queues, virtually no refusals, etc.), there will be no practical benefit resulting from it. The situation assumed is one of *underuse* of the terminal facilities.

Second, if the situation at the terminal is badly saturated, a moderate improvement in service rate cannot effect any noticeable improvement of the conditions at the terminal. This is a case of *overuse.* Both this and the case of underuse are best relieved by adapting the schedules.

Third, there is the intermediate case in which the facilities can, under reasonably favorable conditions, just barely handle the air traffic, but in which if the conditions – in the course of the normal fluctuation of things – become less favorable, the load cannot be handled without delays and cancellations. Under these intermediate conditions the improvement produced by the greater service rate is great; but it is not a linear function of the latter, but of a far more complicated sort.

A simple exercise in algebra should dispel the too commonly held notion that improvements in capacity are proportional to – or even of the same order as – the improvement in such parameters as landing rate μ. To illustrate the point we shall oversimplify the situation by assuming that a single queue satisfying equations (3.1), (3.2), and (3.3) is considered during such a *short* time that coefficients can be regarded as constants: $R_n(t) = \lambda$, $L_n(t) = \mu$; and yet that the time is *long* enough for an approximately stationary state to have developed. We emphasize that under actual conditions such a situation will involve a contradiction, and that it is merely a device for making a point.

We have thus reduced our equations to a well known type that can be solved algebraically (the time derivatives all vanishing) and the solutions of which are expressed explicitly by means of formulas for geometric series. Let us take m = 25 as maximum queue length, as before, and denote by $\bar{n}$ its expected length. If the landing rate μ has a marked advantage over the arrival rate λ, as when $\lambda/\mu = 2/3$, the formulas show that there is practically no queue. If the reverse is true and $\mu/\lambda = 2/3$, they show that the mean queue $\bar{n}$ is within a couple of units of its maximum m. These are the cases in which a slight percentage increase in μ produces little change. The important case is the critical one in which, on setting $\mu/\lambda = 1 + v$, the quantity v is very small. Computations of queue lengths are in terms of v. If the increase in μ makes μ/λ go from 19/20 to 21/20 (i.e., a 10% increase: v goes from $-$ 1/20 to + 1/20), the expected queue length changes by 45 percent as compared with its median m/2 = 12.5. With the same change in μ/λ, the number diverted during any given time will fall by a *factor* of 7.9.

Returning to the more realistic model of time dependence, we have only to examine the graphs of Chapter 4 to see how the increases in service rates, $\mu = 45$ to 55 (22%), and from 55 to 70 (27%), change the relevant conditions. In particular, the expected length of time in the queue for an aircraft joining it at any given time of day is the queue length $\bar{n}$ divided by μ (the hour being the unit of time in the calculation). The expected total numbers turned away in a 24-hour day corresponding to the above three values of μ are

Terminal A	300, 170, 40	(Poisson service)
Terminal B	100, 25, 0	(Poisson service)

Clearly, the quantitative conditions assumed with Terminal A are, as we have noted all along, highly unrealistic. Those of Terminal B are much closer to reality, but represent conditions on the side of overloading. In any real application, the calculations would have to be repeated for the actual parametric values observed.

6.3 FLOW CONTROL

In each of the first two models examined in this report, the law of arrival was assumed to be purely random (Poisson); the only non-random part of it was the mean arrival rate, $\lambda(t)$. This assumption represents the actual situation at most air terminals at the present time. However, there have been suggestions that the service delivered by the terminal might be improved by a regulation of the arrivals, generally referred to as "flow control." Two things could be done: first, the actual mean rate $\lambda(t)$ could, by a change of scheduling or similar means, be changed to a function that varies less with the time, particularly with a flattening out of peaks. Second, without changing the function $\lambda(t)$, measures could be taken to diminish the random feature of the arrival law so that it approaches as much as possible the situation assumed in the third model examined in this report – the deterministic flow model. While the value of each of these two changes can easily be examined by the methods set forth in this report, it is to the latter, most properly called flow control, that we shall give our attention.

To show the greatest benefit that can be expected from such flow control, we shall compare the results calculated by the three models and presented graphically in Chapter 4. The two indicators are: (1) expected length of queue from which, as stated before, the expected waiting time is at once derived; and (2) the total number of aircraft in a day that are refused admission to the queue.

The first of these criteria is illustrated by the graphs of Chapter 4, which can be compared in the purely random and the perfectly regular cases. The second criterion, calculated by the methods of Appendix C, leads to the results of which the following are an illustrative sample:

Number of aircraft refused, Terminal B.

	$\mu = 45$	$\mu = 55$	$\mu = 70$
Random Arrivals and Service	100	55	0
Deterministic Case (flow control)	76	6	0

As noted at the end of Section 2.3, the methods set forth here allow the calculation of the effects of metering loss – and therefore the calculation of the results of its diminution – to be made in all precision. The effects of greater control of the distance between successive aircraft in a landing pattern are not merely to approach the state of affairs of the deterministic flow model, with the advantages explained in the last paragraph: more importantly, it is in the increase in the mean service rate μ, since a closer spacing, and therefore a more rapid feed-through, becomes an acceptably safe device with stricter metering.

APPENDIX A

PERIODIC PROPERTIES

For the case of a single queue with the random (Poisson) service statistics, governed by the system of differential equations (3.1), (3.2), (3.3), the proof of the following theorem was outlined in the 1969-1970 Interim Report FAA-RD-70-70, Chapter 6 (Ref. 1):

> If all the coefficients $R_n(t)$ and $L_n(t)$ are periodic in t with the period T (24 hours), and if none of these coefficients vanishes identically over a time interval of length T, then one and only one probability solution will be periodic of period T, while every other probability solution will approach the former exponentially as t increases.

In the case of the fixed service time c, the corresponding theorem, applying to the system of equations (3.5), runs as follows:

> If the fixed service time c goes into T exactly $N > m$ times ($T = Nc$), and if the quantities $a = a_s = c\lambda(t_s)$ are periodic with this same period ($a_{s+N} = a_s$), and if, finally, not every $a_s = 0$, then there exists one and only one probability solution p^s of period N, all other probability solutions approaching it exponentially as s increases.

It is evidently sufficient, in view of (3.6) and (3.7), to establish the conclusion for the modified equations in $p^s = (\bar{p}_1^s, p_2^s, ..., p_m^s)$ of matrix A_s. In view of our assumptions, this matrix is of period N, i.e., for every s, $A_{s+N} = A_s$. The same will then be true of products of these matrices. Further, the matrix form $p^{s+1} = A_s p^s$ of our equations shows that $p^N = Bp^0$, where $B = A_{N-1} A_{N-2} \cdots A_0$. If p^s has N as a period, the initial values p^0 will satisfy the equation $p^0 = Bp^0$. The converse is true, since this implies for every s

$$\begin{aligned} p^{s+N} &= A_{s-1+N} A_{s-2+N} \cdots A_{1+N} A_N A_{N-1} \cdots A_0 p^0 \\ &= A_{s-1} A_{s-2} \cdots A_0 Bp^0 \\ &= A_{s-1} A_{s-2} \cdots A_0 p^0 = p^s. \end{aligned}$$

Thus the question of periodicity reduces to that of the matrix equation $q' = Bq$ where q is a column of m probabilities. Since $N > m$, an inspection of (3.8) shows that all the zero elements become progressively filled up in the course of the multiplication by elements which (in view of our assumptions of positiveness of a_s) are themselves positive. Therefore, all the elements in B are positive. Finally, it is a stochastic matrix, the columns adding up to unity. Consequently, the standard theorems[5] apply to the matrix equation $q' = Bq$, from which our conclusion follows: a $q' = q$ exists for which $q = Bq$.

In the deterministic flow model the corresponding theorem is true when there is a periodic inflow; but it hardly merits detailed discussion.

The "relaxation method" is simply the procedure for approximating the periodic solution by starting with any initial values and calculating the corresponding solution as time increases, until it seems to reproduce itself. The numerical effort is greatly reduced by a reasonable choice of the initial values.

APPENDIX B

MATRIX PROPERTIES USED IN CHAPTER 3

In order to establish equations (3.6) and (3.10), we first examine the m determinants $(n = 1, 2, \dots, m)$

$$D_n(a) = \begin{vmatrix} a & 1 & 0 & 0 & \dots & 0 & 0 \\ \frac{a^2}{2!} & a & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{a^{n-1}}{(n-1)!} & \frac{a^{n-2}}{(n-2)!} & \dots & \dots & \dots & a & 1 \\ 1 & 1 & \dots & \dots & \dots & 1 & 1 \end{vmatrix} ; \; D_1(a) = 1.$$

The first step in its evaluation is to differentiate $D_n(a)$ when $n > 1$ with respect to a, using the rule of differentiation by rows. The derivative of the last row is a row of zeros. The derivative of the next to the last row gives a row equal to the one above it; and this will continue to be true for all rows except the first. Therefore, the only non-vanishing determinant is the one in which the first row is differentiated; and this is evidently $D_{n-1}(a)$. Therefore, for $n > 1$, $D_n'(a) = D_{n-1}(a)$, whereas $D_1'(a) = 0$. Consequently, for the k'th derivative

$$D_n^{(k)}(a) = D_{n-k}(a) \;, \quad (k \leqslant n-1) \;; \; D_n^{(n)}(a) = 0. \tag{B.1}$$

The second step in the evaluation is based on the obvious fact that $D_n(a)$ is a polynomial in a; let it be expanded about $a = 0$ by the Taylor formula:

$$D_n(a) = D_n(0) + D_n'(0)a + D_n''(0)\frac{a^2}{2!} + \dots + D^{(n-1)}(0)\frac{a^{n-1}}{(n-1)!} + \dots$$

Applying (B.1), together with the obvious fact that $D_n(0) = (-1)^{n-1}$, we obtain, for $n \geqslant 1$,

$$D_n(a) = \frac{a^{n-1}}{(n-1)!} - \frac{a^{n-2}}{(n-2)!} + \dots + (-1)^n a + (-1)^{n-1}. \tag{B.2}$$

It is now easy to establish formula (3.10). Upon the determinant of the matrix A(a) of (3.8), we operate as follows: First, add the first $m-1$ rows to the last row, which thereby becomes a row of units. Second, take out from each of the first $m-1$ rows the factor e^{-a}, calling the resulting determinant, which contains no exponential factor, $\Delta_m(a)$; so we have

$$\det A(a) = e^{-(m-1)a}\,\Delta_m(a).$$

Third, expand $\Delta_m(a)$ in terms of the elements of its first row, obtaining

$$\Delta_m(a) = D_m(a) + D_{m-1}(a).$$

On combining these two formulas with (B.2), equation (3.10) follows.

We turn now to the proof of (3.6). The augmented matrix $\| p^{s+1}, M \|$ as found from (3.5) is the following matrix of m + 1 rows and m + 2 columns

$$\left\| \begin{array}{cccccc} p_0^{s+1} & e^{-a} & e^{-a} & 0 & \cdots & 0 \\ p_1^{s+1} & ae^{-a} & ae^{-a} & e^{-a} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ p_{m-1}^{s+1} & \frac{a^{m-1}}{(m-1)!}e^{-a} & \frac{a^{m-1}}{(m-1)!}e^{-a} & \frac{a^{m-2}}{(m-2)!}e^{-a} & \cdots & e^{-a} \\ p_m^{s+1} & u_m & u_m & u_{m-1} & \cdots & u_1 \end{array} \right\|$$

For consistency of the equations (3.5) it is necessary and sufficient that the p^{s+1} quantities be such that the rank of this matrix be the same as the rank of the square matrix M obtained from the above by removing the first column.[6]

To apply this criterion, it is sufficient to apply it to the equivalent modified matrix obtained from the above one by the elementary transformation consisting of adding all the first m rows to the last, which thereby becomes a row of units. The m + 1 rowed determinant on the right, equivalent to M, vanishes, having its first two columns equal. On the other hand, the lowest m rowed determinant on the right is not zero and, in fact, can be evaluated by first taking out the (m – 1) factors e^{-a} from the rows, and then noting that the result is $D_m(a)$. The value is therefore $e^{-(m-1)} D_m(a)$. Thus the rank of M is m.

In order that the augmented matrix also have the rank m, it is necessary and sufficient to show that every m + 1 rowed determinant in the augmented matrix (or its equivalent) having the non-vanishing m rowed determinant considered above as a minor vanishes.[6] Clearly, this is equivalent to the equation obtained by erasing the second column of the (modified) augmented matrix, and setting the resulting m + 1 rowed determinant equal to zero. On expanding this determinant in terms of the elements of the first column, the desired linear relation (3.6) is obtained.

We have already obtained the coefficient of p_0^{s+1}. To find the coefficient of p_n^{s+1}, we must take $(-1)^n$ times the determinant having the last m columns of the (modified) augmented matrix, but with the single row numbered by n removed. Consider the first n rows of the latter determinant. All but one of the n rowed determinants in these rows is

…in solving (3.1) directly — at an acceptable level of reliability of the answers; i.e., a level suitable for decisions involving large expenditures for improving the air transportation system.

zero, the exception being the one composed of the first n columns; and we see at once that this is unity. Therefore, the Laplace expansion of the required determinant in terms of the determinants in these first n rows gives, simply, this unit determinant, multiplied by the determinant in the last $m - n$ rows and columns. This latter is evidently $D_{m-n}(a)$. Therefore, the coefficient of p_n^{s+1} is $(-1)^n e^{-(m-1)a} D_{m-n}(a)$. Finally, the co-factor of the (unit) last element in the first column is $(-1)^m e^{-ma}$. On writing the resulting linear relation and dividing out $e^{-(m-1)a}$, the required formula (3.6) follows.

It may be noted that in this reasoning it has been assumed that when certain determinants are not identically zero, they are not equal to zero; this assumption causes no restriction in the present applications.

APPENDIX C

CALCULATION OF EXPECTED REFUSALS

There are three cases to consider, corresponding with the three single queue models of Chapter 3: random arrivals and service; random arrivals and regular service; and regular arrivals and service – the deterministic flow.

The first case is reasonably simple. Consider a small time interval $(t, t + \Delta t)$; neglecting higher order terms in Δt, the expected number of refusals during this interval is the probability of a full queue at t times the probability of an arrival during the short interval. This elementary expected value is therefore $P_m(t)\,\lambda(t)\,\Delta(t)$. The total expected number of refusals during any time interval I is the time integral of $P_m(t)\,\lambda(t)$ over I; for the 24-hour day, over T.

In the second case the formulas are more complicated. Starting with the interval of length c between s and s + 1, we must first find the expected number of refusals at s + 1, due to arrivals between s and s + 1. Since the event of any such refusal implies a full queue at s + 1, the desired expected number is the product of p_m^{s+1} times the conditional expected number of disappointed candidates presenting themselves during (s, s + 1). This number is evidently less than the total expected number of arrivals during this interval, i.e., a_s since some of these can be accommodated; hence our answer is *less* than $p_m^{s+1}\,a_s$, which, since $a_s = c\lambda$, corresponds with the expression of the last paragraph in which $\Delta t = c$.

A more accurate appraisal of the above expected value is necessarily based on the exact formula derived as follows:

If the queue is empty at s, the probability of k rejections at s + 1 is the probability of m + k arrivals, i.e., $a^{m+k}\,e^{-a}/(m + k)!$. Hence the conditional expected number rejected is the sum over all positive k of k times this expression. Replacing k by (m + k) – m and separating terms, etc., we obtain, with the earlier formula for u_r, the expression $au_m - mu_{m+1}$. Similarly, the conditional expected number rejected if there are n in the queue at s is $a\,u_{m-n+1} - (m - n + 1)\,u_{m-n+2}$. Therefore, the unconditional expected number rejected is the sum, for n = 0, 1, ..., m, of the products of p_n^s times the corresponding conditional expected values just derived.

The computer program for calculating the quantities p_n^s can be made to calculate these expected numbers rejected.

In the case of the deterministic flow model, the expected number in the queue at time t coincides with the actual number. Aircraft will be rejected only during those periods between the time t′ that the graph of the queue length cuts the horizontal line of height m on its way up, and the time t″ when it next cuts the latter on the way down. The expected number rejected between t′ and t″ is the integral of $\lambda(t)$ between these limits, minus the number serviced, $\mu(t'' - t')$.

 PRECEDING PAGE BLANK

When the statistical data of arrival rates are given for each hour (as at the angular points of Figure 4.1), a simple manual computation quickly gives the number of rejects without going through the integration. This method, which was developed under the present contract, may be useful to apply as an approximate estimate in certain practical cases.

REFERENCES

1. See *A Study of Air Traffic Control System Capacity,* Arthur D. Little, Inc., Interim Report, Sept.-Aug. 1970, Report No. FAA-RD-70-70 [FAA Contract No. FA70WA-2141], Chapter 6.

2. Tables of this incomplete Gamma function are to be found in *Handbook of Mathematical Functions,* National Bureau of Standards, Applied Mathematics Series 55.

3. Reference 1, Chapter 3.

4. Contained in an unpublished *FAA Staff Study of Airport Congestion in Major Metropolitan Areas,* made available by the Federal Aviation Administration in connection with contract work by Arthur D. Little, Inc. The statistics were graphed as follows: for each interval between consecutive hours (e.g., 10 and 11) the average number of events was found and plotted as the ordinate of a point whose abscissa was the corresponding midpoint (e.g., 10:30). The resulting 24 points were then joined by straight line segments.

5. See, e.g., D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes,* 3.11-3.12, John Wiley, New York, 1965.

6. See, e.g., M. Bôcher, *Introduction to Higher Algebra,* Chapter IV Theorem 1, Chapter V Theorem 1, McMillan, New York, 1915.