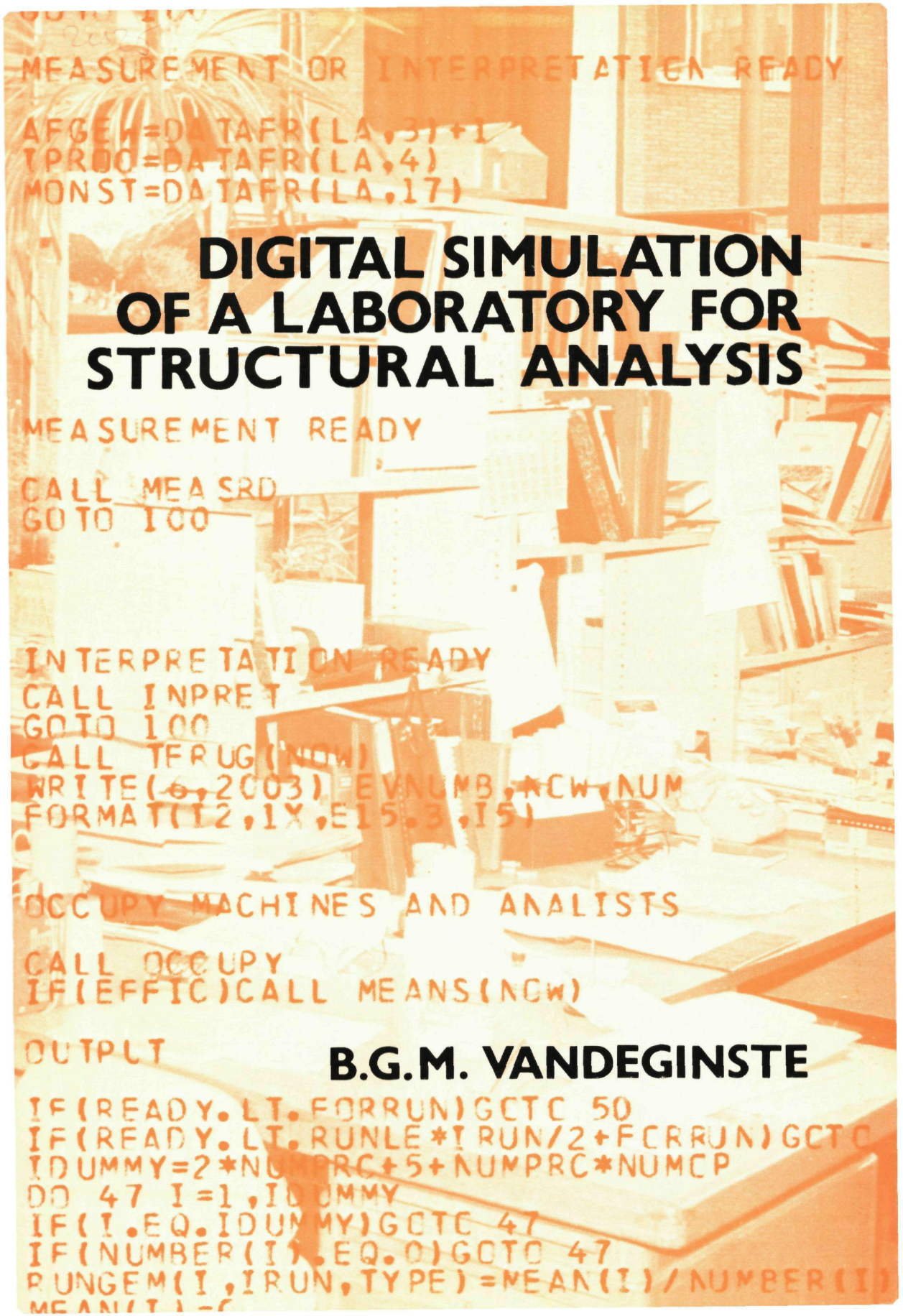


DIGITAL SIMULATION OF A LABORATORY FOR STRUCTURAL ANALYSIS

B.G.M. VANDEGINSTE

DIGITAL SIMULATION
OF A LABORATORY FOR STRUCTURAL ANALYSIS

PROMOTOR.

PROF. DRS. G. KATEMAN

DIGITAL SIMULATION
OF A LABORATORY FOR STRUCTURAL ANALYSIS

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
IN DE WISKUNDE EN NATUURWETENSCHAPPEN
AAN DE KATHOLIEKE UNIVERSITEIT TE NIJMEGEN, OP GEZAG VAN
DE RECTOR MAGNIFICUS, PROF.DR. P.G.A.B. WIJDEVELD,
VOLGENS BESLUIT VAN HET COLLEGE VAN DECANEN
IN HET OPENBAAR TE VERDEDIGEN
OP VRIJDAG 11 JANUARI 1980
DES NAMIDDAGS TE 4 UUR

door

BERNARD GABRIEL MARIE VANDEGINSTE
geboren te Kortrijk

Druk: Krips Repro Meppel

Allen die op enigerlei wijze hebben bijgedragen aan de totstandkoming van dit proefschrift wil ik hierbij van harte bedanken. Met name alle medewerkers van de afdeling Analytisch Onderzoek van Philips Duphar B.V. te Weesp, voor hun stimulerende bijdragen aan de discussies van de resultaten van het onderzoek, in het bijzonder drs. F.J. Spruit, zonder wiens medewerking en stimulans dit onderzoek niet uitgevoerd had kunnen worden.

ontwerp omslag: H.J.M. Spruyt, afdeling fotografie

lay-out: W.J.M. Philipse

Voor Veronique,
Tom en Pieter

CONTENTS

CHAPTER 1

INTRODUCTION

1. General introduction

Analytical chemists should produce qualified and relevant information about products and processes in an optimal way [GO72, KA74] or at least provide the strategies to do so [KA79].
Basically, the main tool for the extraction of the requested information, available to the analytical chemist, is the analytical procedure, which is the way the determination of the identity and/or amount of the compounds is effected.
The selection of the best analytical method for solving a given problem is a task, with which the analytical chemist is confronted daily. Although many comparisons between various analytical methods are reported frequently, still no systematical strategy exists to select the best method.
Vandeginste [VA77] made an attempt to select between atomic absorption and u.v.-v.i.s. absorption spectrometry by pattern recognition. This research clearly demonstrates that before trying to solve the method-selection problem, the analytical chemist has to collect a large amount of data that relates the analytical procedures to their analytical problems. To my knowledge, up to now such collections have never been assembled to this purpose.
Apart from the creation of strategies for procedure selection, the optimization

of existing and development of new analytical procedures remains extremely important. Moreover, nowadays,sopnisticated data nandling is more and more required for the extraction of useful information from single or combined analytical results [VI77, RE73, VS76]. Consequently the analytical chemist must study these data handling methods and pay attention to their limitations [VA75] and applicability.

Clearly, the amount of information of a single analysis or a series of analyses depends on the difference of the uncertainty about the product or process, before and after the analytical result is obtained.

Müskens [MU78] and others [GR73, VJ77] have shown that the sampling frequency, the accuracy of the analysis and the dead time, or delay time of the analysis determine the possibility of controlling a process. Optimal control is achieved when the sum of analysis and process costs is minimal. The analysis costs are mainly determined by the analytical accuracy, sampling frequency and delay time. Therefore, it is important to minimize costs of analytical response time and accuracy. The delay time consists of two parts: the real analysis time and the waiting time in the laboratory, which is the time lag between the moment the sample is received at the laboratory and the start of the analysis. The delay of samples in an analytical laboratory normally exceeds the analysis time greatly as most of the time is spent waiting. Therefore, the study of the organization of analytical laboratories is important, especially the process of queueing.

Delay times are directly related to the costs of the analysis, as there exists a relation between the amount of facilities (analysts- and apparatus) and delay time. This leads to the next level of optimization the analytical chemist is concerned with, namely the optimization of combinations of analytical procedures.

All these developments led to the inclusion of new mathematical techniques such as information theory, pattern recognition, operations research and control theory, in the scope of analytical chemistry [KO78]. This helps the analytical chemist in producing better analytical information in an optimal way.

2. Problem formulation

Clearly, the quantitative study of waiting line situations in analytical laboratories should permit a better use of the capacity of laboratories and

a reduction of delays. However, up to now only a few studies have been reported on laboratory activities [VA74, SC76, SC77].

Obviously, an analytical laboratory is a complex organization, which can be defined as follows: "It is the rational coordination of the activities of a number of people for the achievement of some common explicit analysis or analytical goals, through division of labour and function and through a hierarchy of authority and responsibility", Cook [CO76].

It is evident that the goal of a search for quantitative relations between several variables in the laboratory should never be a substitute for human creativity or human flexibility. However, it should be an aid for the analytical chemist in decision-making. A complex organization such as an analytical laboratory really is, can never be simplified to a model governed by a set of strictly mathematical rules. Clearly human response under different conditions is difficultly predictable. In contrast, it is impossible to study the mentioned relations from experiments with the real laboratory itself. Therefore, some alternative system should be used, a so-called model, which is similar to the real system in the characteristics of interest. This alternative system cannot be expected to exactly reproduce these characteristics of interest. This is the price to be paid for simplicity and accessibility of the alternative system.

In the work described here, a model is constructed and validated for an existing laboratory for molecular spectroscopic analysis, dealing with i.r., p.m.r., ^{13}C-n.m.r. and m.s. analyses. In this model the time lag between the arrival of a sample and the production of the analytical result is studied.

In an analytical laboratory and especially a spectroscopic laboratory various questions should be answered:

First of all the question about forecasting the delay time as a function of the mean sample traffic, number of analysts and instruments. Thereafter, strategies should be determined for the selection of the analytical method taking into account the estimated probability the various analytical methods might solve the requested structure, and the state (queue lengths, available capacity) of the laboratory. Also decision rules must indicate the best mode of action after an analytical method fails to elucidate the structure.

Finally, the assignment of priorities involves priority between samples from different research groups, priority between samples unsuccessfully analyzed, and priority of 'easy' (short analysis times) over 'difficult' problems (long analysis times).

There are various types of priority disciplines:
When an absolute priority rule is applied, all samples of a higher priority are always analyzed before samples of a lower priority irrespective of their waiting time. In contrast, the application of a time dependent priority rule or dynamic priority, has the effect of considering some samples to have a higher priority than others, but takes into account the undesirability of having low priority samples wait too long.
Another effect that should be clarified is the influence of other activities of the analyst, who interrupts the analytical process, while still samples are waiting.
Generally speaking, two types of models are suitable for this study.
In the first place, strictly mathematical models with theoretically deduced solutions, developed in queueing theory. Secondly, simulation models which describe the operation of the real system in terms of individual events of the individual elements or compounds of the system.
In complex systems of networks of queues, such as analytical laboratories, consisting of multiserver nodes and governed by state-dependent decision rules, queueing theory cannot provide exact results. However, queueing theory gives a good picture of the behaviour of queues in very simple single server systems. Because an investigation which is not based upon a theory or a formal hypothesis is just blind groping in the dark, the effect of various variables and strategies for those simple single server systems were calculated firstly by queueing theory, giving a hypothesis about the effects to be expected from the simulation experiments. Furthermore, simulations of those simple systems confirmed the validity of the simulation model.
Model building requires a knowledge of computer programming, statistics, probability theory and experimental optimization techniques. Because in a simulation model a great number of variables is involved, a good experimental optimization method is very important in order to obtain the desired information.
Computer simulation experiments and modelling in general, usually consists of the following stages [NA66]:
After the formulation of the problem (Chapter I), laboratory data should be collected and processed (Chapter II), such as the interarrival times of the samples, the mean down time of the instruments and the delay times of the samples. Some of these observations, such as those concerning the arrival of samples and the delivery of the analytical results can be obtained from

the book-keeping of the laboratory. Other data can be obtained from interviews with the analysts, e.g. to find out which priority policies are used to select an analytical method in the laboratory.

The most difficult and time-consuming stage of computer simulation is the formulation of the mathematical model. Because here all variables, parameters and relationships must be specified (Chapter IV). The variables are selected on the basis of an estimate of their relative importance (Chapter III). If one or more important variables are missed, the simulation results become inaccurate. In contrast the inclusion of too many variables renders the computer simulation needlessly complex.

The next stage is to estimate the parameters of the distributions of several variables (Chapter II), including tests for autocorrelation. To do this, various statistical tests can be used.

The most important stage of the simulation is the validation of the computer model (Chapter V). Some assurance of the validity of forecasts of future behaviour of the modelled laboratory, can be provided by a demonstration that for at least one alternative version of the simulated system and one set of conditions, the model produces results that are consistent with the known performance of the investigated laboratory.

The ultimate test of a computer simulation model is the degree of accuracy with which the model predicts the behaviour of the actual system in the future (Chapter V).

Once the validity of the computer model is satisfactory, the model can be used to conduct actual simulation experiments, which may be designed by experimental design techniques (Chapter VI).

CO76 C.F. Cook, Anal. Chem., 48 (1976), 724A.

GO72 G. Gottschalk, Z. Anal. Chem., 258 (1972),1.

GR73 P.M.E.M. van der Grinten and J.M.H. Lenoir, Statistische Procesbeheersing, Spectrum, Utrecht, 1973.

KA74 R.E. Kaiser, Z. Anal. Chem., 272 (1974), 186.

KA79 G. Kateman, A. Dijkstra, Z. Anal. Chem., 297 (1979), 249.

KO78 B.R. Kowalski, Anal. Chem., 50 (1978), 1309A.

MJ78 P.J.W.M. Müskens, Dutch Thesis, Nijmegen, 1978.

NA66 T.H. Naylor, J.L. Balintfy, D.S. Burdick, K. Chu, Computer Simulation Techniques, J. Wiley & Sons, New York, 1966.

RE73 J.D. Remijnse, B.G.M. Vandeginste, B.M. Wepster, Rec. Trav. Chim., 92 (1973), 804.

SC76 B. Schmidt, G.I.T. Fachz., 20 (1976), 1167.

SC77 B. Schmidt, Z. Anal. Chem., 287 (1977), 157.

VA74 I.K. Väänänen, S. Kivirikko, J. Koskenniemi, J. Koskimies and A. Relander, Meth. Inform. med., 13 (1974), 158.

VA75 B.G.M. Vandeginste, L. de Galan, Anal. Chem., 47 (1975) 2124.

VA77 B.G.M. Vandeginste, Anal. Letters, 10 (1977), 661.

VI77 B.G.M. Vandeginste, P.B.W. van Iersel, Proc. Analyt. Div. Chem. Soc., (1978), 10.

VJ77 B.G.M. Vandeginste, T.A.H.M. Janse, Z. Anal. Chem., 286 (1977), 321.

VS76 B.G.M. Vandeginste, P.J.M. Salemink, J.C. Duinker, Neth. J. Sea Res., 10 (1976), 59

CHAPTER 2

THE DESCRIPTION OF DELAY AND QUEUE BEHAVIOUR IN A LABORATORY FOR STRUCTURAL ANALYSIS

A first step towards modelling reality is the collection and processing of data from the system considered. During the period 1/6/1976 - 1/6/1978, the arrival dates of all samples at the various analytical sections and the dates of completion of the analysis were registered in the spectroscopic laboratory for structural analysis at Philips-Duphar B.V. at Weesp, the Netherlands.

The data from 1/6/1976 - 1/6/1977 were used to determine various parameters of the system and to validate the simulation model, described in Chapter IV. The data from 1/6/1977 - 1/6/1978 were used to test the predicting power of the simulation model.

The laboratory receives about 3000 samples per year. These samples are analyzed in four sections: infrared (i.r.), proton magnetic resonance (p.m.r.), mass (m.s.), and carbon nuclear magnetic resonance (^{13}C-n.m.r.) spectrometry. The analysis consists of two steps: the measurement and the interpretation of the spectrum.

From the collected observations the statistical properties of the flow through the network and the queue levels in each section could be calculated.

Published in part in:

B.G.M. Vandeginste, Analyt. Chim. Acta, CTO, accepted for publication.

B.G.M. Vandeginste, Communication presented at the IUPAC Congress, Helsinki (1979)

The laboratory under investigation can be considered as an open network, consisting of 4 nodes (the 4 sections) which receive samples from two different origins. In an open network, samples arrive from external sources (the environment) and each sample eventually leaves the system. In contrast, in a closed network, the samples circulate through the network without external arrivals or departures [LE77]. An open network where the samples visit a node only once is called an open acyclic network.
In Fig. II-1, a sketch is given of the network of the analytical sections. The arrows connecting each section indicate the direction of the sample flow. The mean flow (samples per day) towards and from each section is also given.

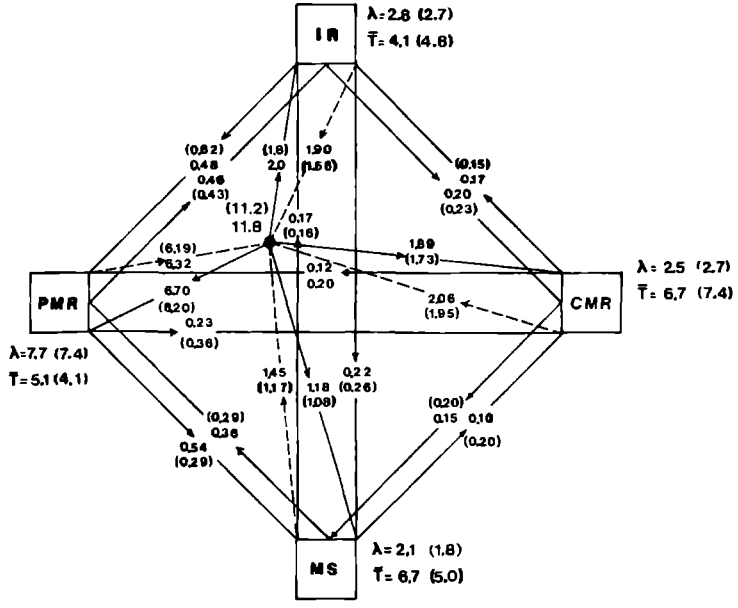


Fig. II-1: The sample flow (samples/day) through the laboratory network. (numbers in parentheses are simulated data) T̄: mean delay time, λ: sample flow

Although similar networks of servers are quite common, a relatively small theoretical basis exists for analyzing networks of queues. An excellent review and critique of the results available for modelling networks of queues with random flows is given by Lemoine [LE77].
The total sample flow to and from each node (section) is tabulated in

table II-1. These observations indicate that the total flow in the network exceeds the number of received samples, as on the average more than 1 analysis (1.28) is done on a sample.

Table II-1

Mean, and variance of input and output flow (samples per day) at the spectroscopic laboratory of Philips-Duphar, period 1/6/1976 - 1/6/1977

Section	input (samples/day)		output (samples/day)	
	mean	variance	mean	variance
I.r.	2.8	13.2	2.8	7.5
P.m.r.	7.7	20.3	7.7	50.0
M.s.	2.1	4.2	2.1	5.9
^{13}C-n.m.r.	2.5	4.2	2.5	11.1
Total	15.1		15.1	
Lab	11.8	44.1	11.8	77.6

From the flow through the network, the conditional probability (p_{ji}) of transfer of samples from one to another section could be established (Table II-2). Moreover, the probability could be calculated that a method will be selected in the first instance or after one or more other methods failed (Table II-3).

Table II-2

Conditional probabilities (p_{ji}) of transfer of samples from one to another section, and the probability (q_i) that a sample in node (i) leaves the system

from \ to	OUT (qi)	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.
OUT	-	0.17	0.57	0.10	0.16
I.r.	0.68	-	0.17	0.08	0.07
P.m.r.	0.83	0.06	-	0.07	0.03
M.s.	0.70	0.08	0.17	-	0.05
^{13}C-n.m.r.	0.83	0.07	0.05	0.06	-

Table II-3

Probabilities of the methods to be selected

	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.	% of samples completed
first selection	0.17	0.57	0.10	0.16	79.6
% good	0.65	0.84	0.72	0.84	
second	0.27	0.29	0.29	0.15	94.5
% good	0.76	0.76	0.64	0.78	
third	0.21	0.29	0.28	0.22	98.3
% good	0.58	0.72	0.78	0.69	
fourth	0.25	0.27	0.25	0.23	99.8
% good	0.92	0.85	0.83	0.91	

Jackson [JA57] derived a balance or conservation equation [Eqn. II-1] for open networks, describing the equilibrium rate of flow through node i, α_i, as the sum of the external input rate λ_i, and the total rate of internal transfers to node i, $\sum_{j=1}^{N} p_{ji}\alpha_j$

$$\alpha_i = \lambda_i + \sum_{j=1}^{N} p_{ji}\alpha_j \qquad i = 1, \ldots N \qquad \text{II-1}$$

with p_{ji} the probability that a sample leaves node j towards node i. For each section (i), the external arrival rate (λ_i) and total arrival rate (α_i) were determined (Table II-4).

The departure flow rate from section (i) equals $\alpha_i q_i$, where q_i represents the probability that a sample in section (i) leaves the system.

These values are tabulated in Tables II-2 and II-4. Substitution of the values α_i, λ_i and p_{ji} in Eqn. II-1 demonstrates that the conservation equation is valid for the laboratory. Moreover, as q(i) and λ(i) are different from zero for each section the network is open.

Table II-4

Verification of the balance equation

Section	λ-observed	α-ooserved	α-calculated	$\alpha_i q_i = \gamma_i$
I.r.	2.01	2.8	2.81	1.90
P.m.r.	6.73	7.7	7.68	6.39
M.s.	1.18	2.1	2.09	1.47
^{13}C-n.m.r.	1.89	2.5	2.42	2.08
Sum	11.81			11.84

At equilibrium the total external input flow rate to the network equals the total external departure flow rate.

Thus

$$\sum_{i=1}^{N} \lambda_i = \sum_{i=1}^{N} \alpha_i q_i \qquad \text{II-2}$$

The validity of Eqn. II-2 in our system is shown in Table II-4.

2. *Input and output of the laboratory*

2.1 Probability density functions

When the aggregate effect of a large number of individuals or particles is under observation in nature, often Poisson processes appear.

A Poisson process appears when the following conditions are fulfilled:

(i) the number of events (the number of arrivals) is a random variable which is independent for non overlapping time intervals.

(ii) the probability of a definite number of events during a certain time interval is only dependent on the length of that interval. For all intervals of constant length this probability is equal, and is independent of the absolute time t.

(iii) the probability of a single event during a small interval is proportional to the length of that interval. The probability of more than one event in such an interval is negligible.

The Poisson process is widely used in queueing theory. Numerous examples have been shown that in many queueing problems (e.g. telephone calls [KL75], airplane arrivals [AC68], patient arrivals in a hospital), the arrival process (e.g. the number of arrivals per day) can be modelled by a Poisson distribution.

When the external input to an open network has a Poisson distribution, and the external input streams are assumed to be independent, and the analysis rate is also independent of the sample arrivals, then [JA57] each node in the network behaves as an independent queue with Poisson input. This facilitates the study of networks of queues considerably.

Therefore, the arrival processes of the samples at the laboratory and at each individual section were compared to the theoretical Poisson distribution.

Fig. II-2 shows the calculated histograms and best fitting Poisson distribution.

A goodness of fit test by means of a Chi-square test (a suitable test for discrete distributions [KR70]) indicated that the arrival distribution functions were significantly different from the Poisson distribution.

In most cases, days with a smal number of arrivals occurred too frequently.

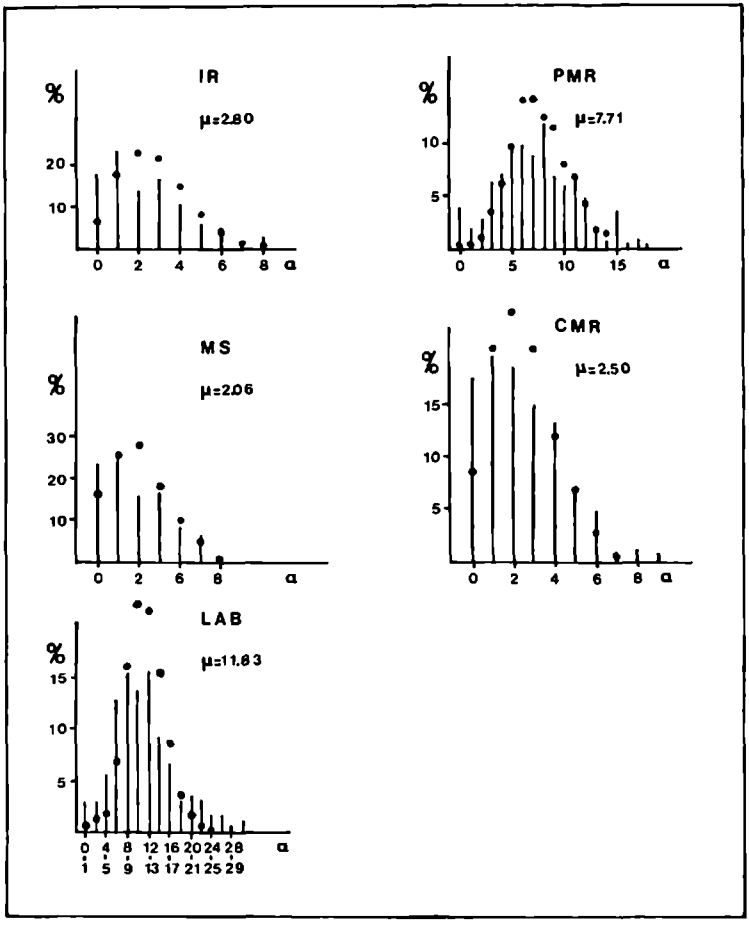


Fig. II-2: Histograms of the probability (%) of an input density α (samples/ day) to the sections of the laboratory. — experimental data, • Fitted Poisson distribution (same mean)

Consequently, the possibility to get exact results from the application of queueing theory becomes hampered. Fig. II-3 indicates clearly that the probability density function (p.d.f.) of the output rate is completely different from the p.d.f. of the input rate, and is far from a Poisson process.

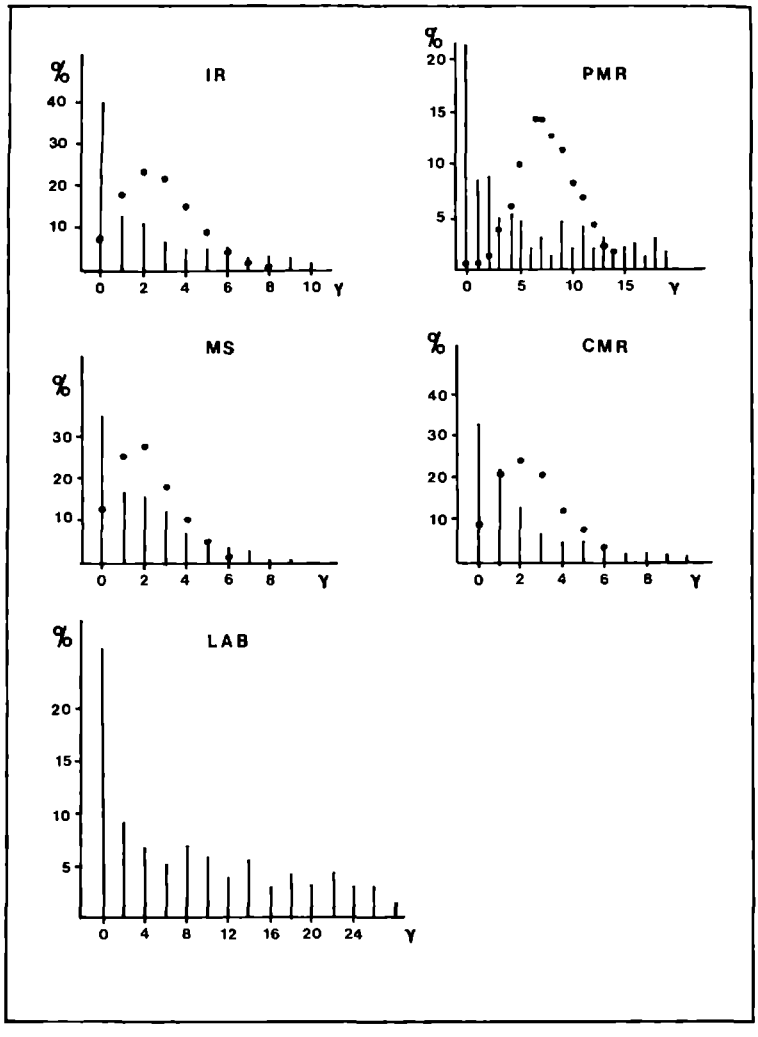


Fig. II-3: Histograms of the probability (%) of an output density γ (samples/day) from the sections of the laboratory. • Fitted Poisson distribution (same mean), — experimental data

2.2 Autocorrelation functions of input and output flow

An autocorrelation analysis is applied to investigate the fluctuations of a process variable during a certain observation period. Since the autocorrelation function of a stochastic stationary variable tends to zero, deterministic parts of the signal are easily detected [MU78].

Autocorrelograms of the input and output rate were calculated for two reasons: (i) when the input and output streams meet certain conditions (such as the

independency of the number of arrivals each day), analytical results may be easier obtained from queueing theory [LE77]. (ii) The statistical description of a system variable is not complete, when only the probability density function is known. Here, the correlation between the value of a system variable at different times should be determined. This completes the description of the input to the laboratory, necessary to create a similar input process in the simulation model (Chapter IV).

2.2.1 Theory

When a stochastic process generates random variables (e.g. the number of samples arriving per day) and there is a random variable x_t for each time t, a time series is observed. A common property of time series is the covariance of $[x_{t+\tau} - \mu]$ and $[x_t - \mu]$, where τ is the time lag, i.e. the number of time intervals between the respective values of the time series. For stationary processes, the covariance function $\psi_{xx}(\tau)$ and autocorrelation function $\phi_{xx}(\tau)$ for a time lag τ are respectively defined as [GR73]:

$$\psi_{xx}(\tau) = E[(x_t-\mu) \cdot (x_{t+\tau}-\mu)] \qquad \text{II-3}$$

$$\phi_{xx}(\tau) = \psi_{xx}(\tau)/\psi_{xx}(0) \qquad \text{II-4}$$

The autocorrelation function of a first order stochastic stationary time series has an exponential shape [MU78]

$$\phi_{xx}(\tau) = e^{-|\tau|/T_x} \qquad \text{II-5}$$

T_x is called the time constant of the time series. T_x^{-1} is a measure for the frequency of the fluctuations of the time series. A large value of T_x indicates that the variations are slow. In contrast, a low value of T_x is found when the variations are fast (Fig. II-4). The autocorrelation function of a discrete time series can be estimated from Eqn. II-6

$$\phi_{xx}(\tau) = \frac{1}{N-\tau-1} \sum_{N=1}^{N-\tau} (x_i-\bar{x})(x_{i+\tau}-\bar{x})/s_x^2 \qquad \text{II-6}$$

where s_x^2 is the estimated variance of x and N the number of observations. When the time series is composed from a deterministic part and a stochastic part, the stochastic part of the time series does not contribute to the autocorrelation for sufficiently large values of τ and the deterministic part

can be detected. For example, Müskens [MU78] calculated the autocorrelation function of a time series composed from a stochastic part (e_t) and a deterministic part

$$f_t = A_p \sin(2\pi t/L + B_p) \qquad \text{II-7}$$

with $E[e_t] = 0$, and $E[x_t] = f(t)$
The autocorrelation function of x_t approximately equals

$$\phi_{xx}(\tau) = [\tfrac{1}{2}A_p^2 \cos(2\pi\tau/L) + \sigma_e^2 \cdot \phi_{ee}(\tau)]/(\tfrac{1}{2}A_p^2 + \sigma_e^2) \qquad \text{II-8}$$

Because the autocorrelation function has to be calculated from a finite time series, the accuracy of the estimation of this function should be established. Bartlett [BA46, BO70] derived the variance of the estimated autocorrelation $\phi_{xx}(\tau)$

$$\sigma^2[\phi_{xx}(\tau)] = \frac{1}{N-\tau} \sum_{j=-\infty}^{j=+\infty} [\phi_{xx}^2(j) + \phi_{xx}(j-\tau) \cdot \phi_{xx}(j+\tau) + 2\phi_{xx}^2(\tau) \cdot \phi_{xx}^2(j) - 4\phi_{xx}(\tau) \cdot \phi_{xx}(j) \cdot \phi_{xx}(j-\tau)] \qquad \text{II-9}$$

Since the theoretical autocorrelation function has the exponential shape (Eqn. II-5), the variance equals

$$\sigma^2[\phi_{xx}(\tau)] = \frac{1}{N-\tau} \left[\frac{(1+\phi_{xx}(1))(1-\phi_{xx}^2(\tau))}{(1-\phi_{xx}(1))} - 2\tau\phi_{xx}^2(\tau)\right] \qquad \text{II-10}$$

Consequently, the estimation of the autocorrelation function equals

$$\phi_{xx}(\tau) \pm u(P) \cdot \sigma[\phi_{xx}(\tau)] \qquad \text{II-11}$$

with u(P) the excentricity of the normal distribution with a confidence of P%.

2.2.2 The observed autocorrelograms of sample input and output of the laboratory.
The autocorrelograms of the number of daily arrivals at and daily departures from the individual sections and the total laboratory were calculated according to Eqn. II-6 and are shown in Fig. II-4 and II-5. The autocorrelograms of the input indicate that the input processes are not autocorrelated.
These observations indicate that the number of arrivals on one day cannot be

forecasted from the number of arrivals at the preceeding day. In contrast to these observations the autocorrelograms of the output process (Fig. II-5) show a distinct deviation from the exponential shape.

The autocorrelogram of the output of the laboratory reveals a periodicity of 5 days, suspecting the laboratory to release the analytical results with a periodicity of 5 days.

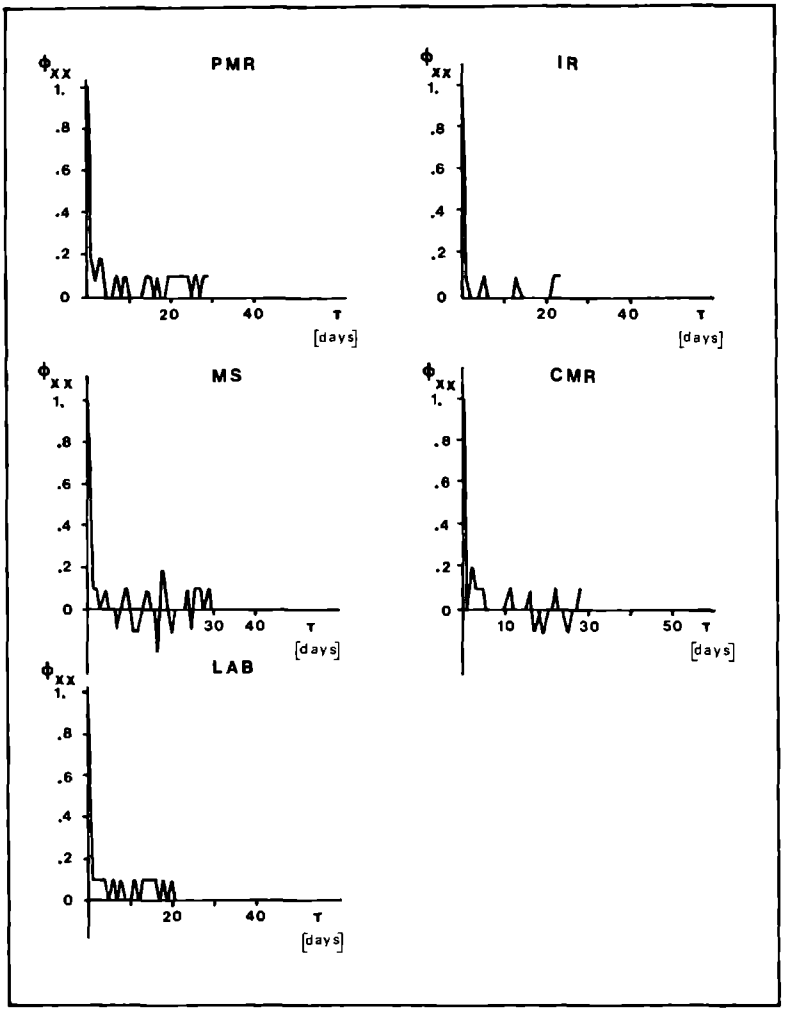


Fig. II-4: Autocorrelograms of the input density to the sections and the entire laboratory

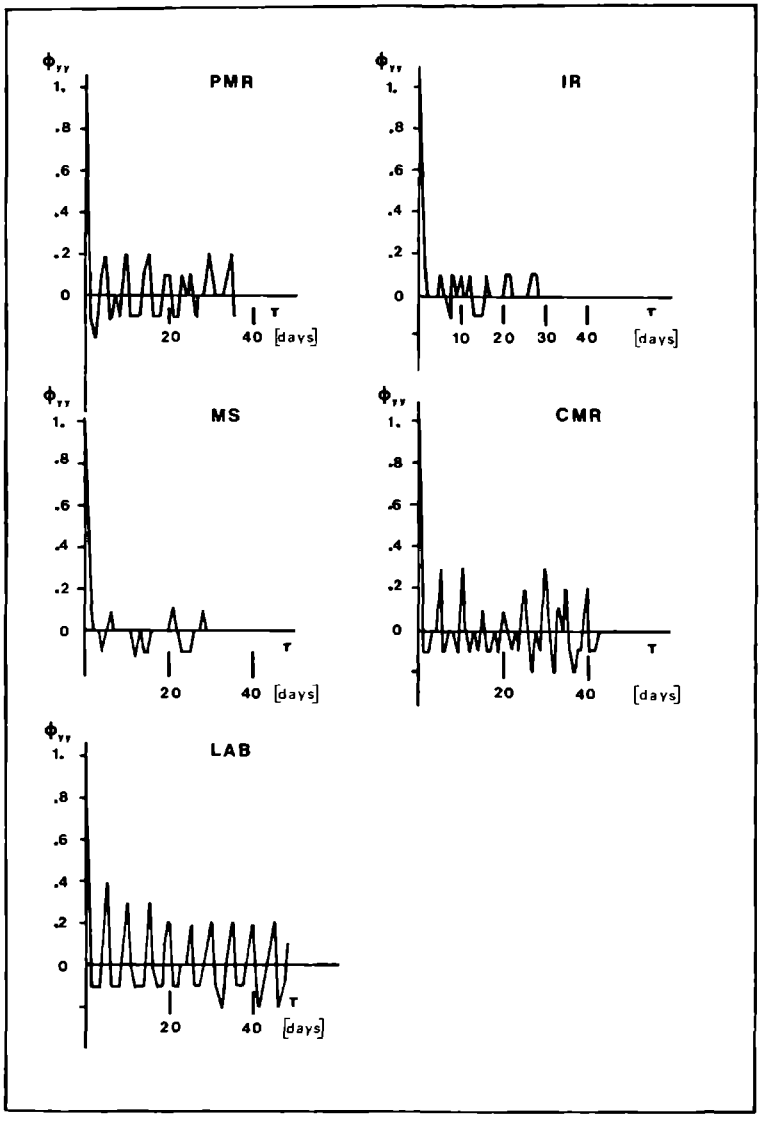


Fig. II-5: Autocorrelograms of the output density from the sections and the entire laboratory.

3. The queue levels in the network

3.1 Autocorrelograms and histograms of the queue levels

Since data on queue sizes were not explicitely available of this network of analytical sections, the queue sizes were calculated from the differences between input and output each day.

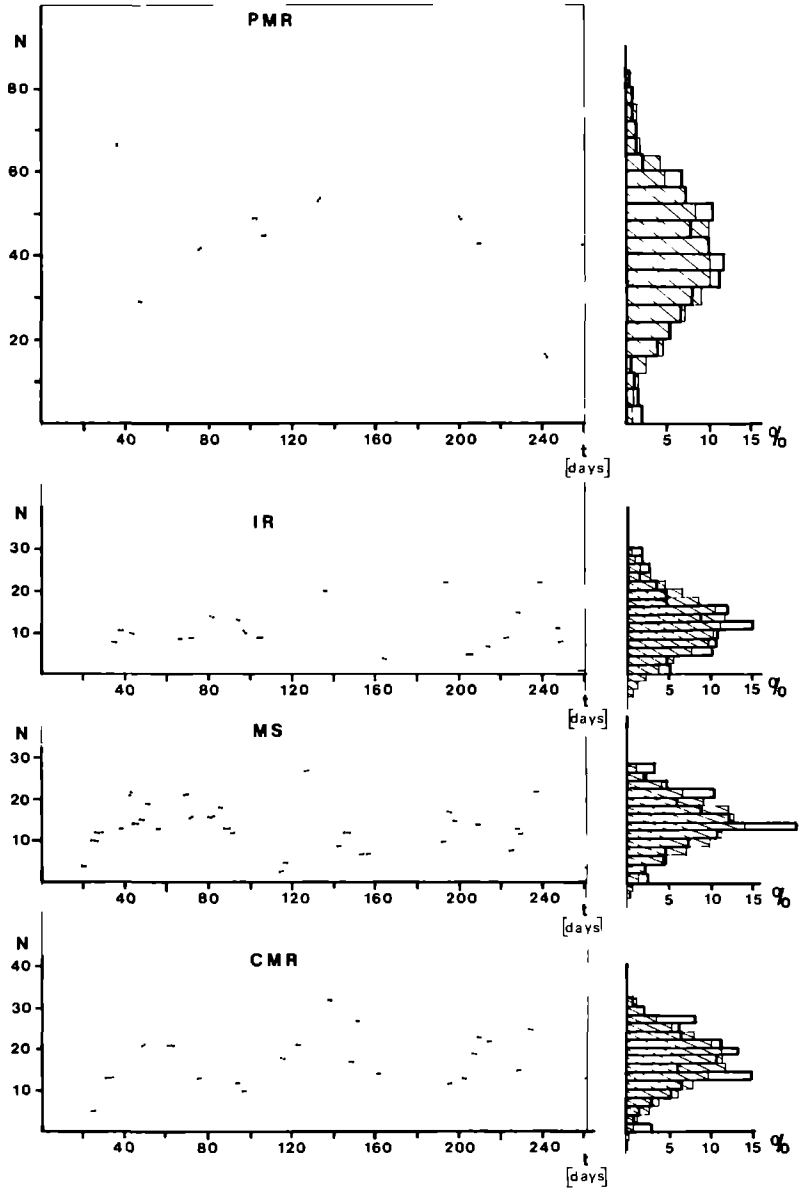


Fig. II-6 The number (N) of samples in the sections of the laboratory as a function of time (t) (days), with histograms. Shaded histograms are fitted Gaussian distributions with same mean and standard deviation as the observations.

The queue size here is defined as the total number of samples, including spectra and samples under investigation, which are present in the laboratory, or in a node of the network. As an analytical section (e.g. mass spectrometry) is represented as a node of the network, no discrimination was made between queues of samples waiting for an instrument or analyst, and the various piles of spectra belonging to each analyst.

Fig. II-6 shows the fluctuations of the number of samples in queue, with a sample interval of 1 day, with their histograms.

Kolmogorov-Smirnov tests [K-S] [KR70] executed on the relative cumulative density functions indicated that no significant difference could be detected with the Gaussian relative cumulative density function having the same mean and standard deviation.

Strictly the K-S test may only be applied on uncorrelated observations. For correlated observations a greater probability that a given value will be exceeded, should be taken into account

From the calculated autocorrelation functions of the queue sizes, shown in Fig. II-7 it is easily seen that the queue sizes are highly autocorrelated. Because testing procedures for comparing historical data and data obtained by digital simulation can be characterized by autoregressive models, and because serial correlation in time is itself often an important characteristic of the simulated system, the parameters of autoregressive models of queue sizes were calculated.

The results of fitting an autoregressive model of order 1 (AR(1)) to these time series are presented in Table II-5.

The algorithm of an AR model of order p equals

$$N_t = \phi_1 N_{t-1} + \phi_2 N_{t-2} + \ldots \quad \phi_p N_{t-p} + a_t \qquad \text{II-12}$$

where $N_t = n_t - \bar{n}$ is the difference in the queue level from the mean level at time t and a_t is a normal random variable with mean zero and variance σ_a^2, the so called residual variance, $\phi_1 \ldots \phi_p$ are the parameters of a AR(p) model.

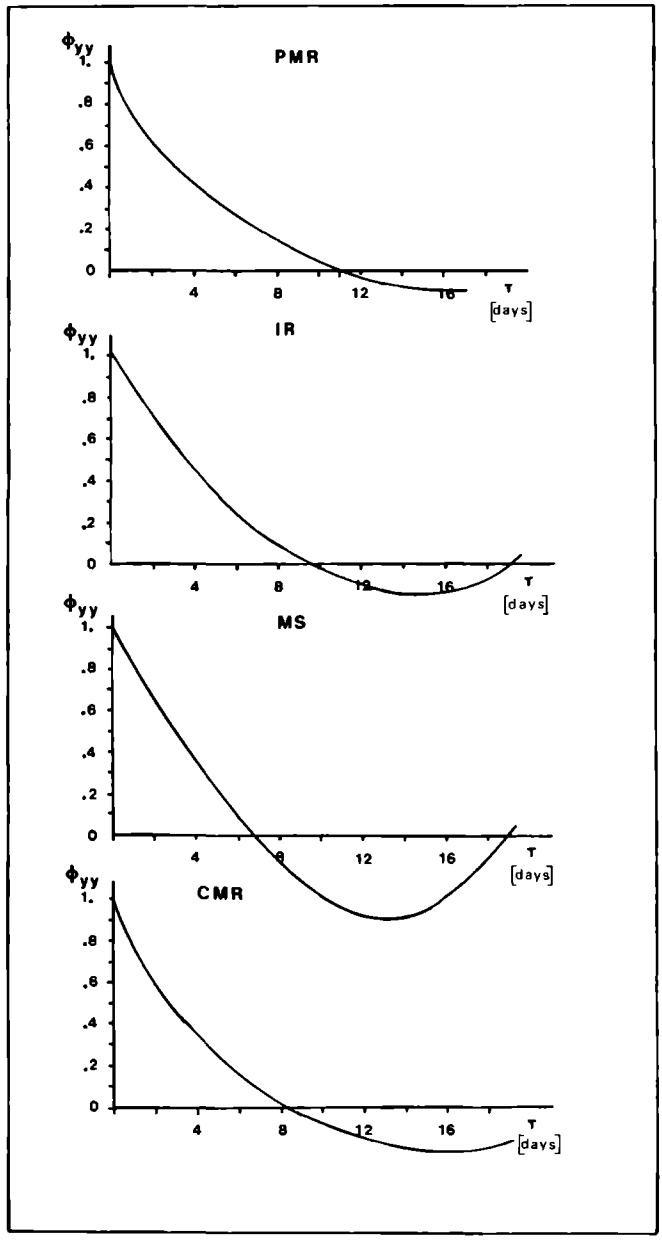


Fig. II-7: Autocorrelograms of the queue sizes (N) in the sections.

Table II-5

Summary of the AR(1) model parameters of the queue lengths (number of waiting samples)

parameters	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.	LAB
mean of series	11.5	39.1	13.9	16.8	81.4
variance of series	46.2	228.0	34.8	43.6	582.7
ϕ_1	0.83	0.81	0.85	0.81	0.83
95%	$\pm$ 0.17	$\pm$ 0.17	$\pm$ 0.17	$\pm$ 0.17	$\pm$ 0.17
var [a_t]	14.37	78.4	9.7	15.0	181.3

When the AR(1) model is adequate, the dependence of N_t on the past history is completely accounted for by the term $\phi_1 N_{t-1}$ in the model.
For a first order model, ϕ_1 is equal to the autocorrelation function at $\tau=1$ ($\phi_{xx}(1)$) and all higher order terms ($\phi_2 \ldots \phi_n$) are not significantly different from zero.
The parameters ϕ_1 and eventually ϕ_2 follow from the autocorrelation values $\phi_{xx}(1)$ and $\phi_{xx}(2)$ according to the Yule-Walker equations [BO70]

$$\phi_1 = [\phi_{xx}(1)(1 - \phi_{xx}(2))]/(1 - \phi_{xx}(1)^2) \qquad \text{II-13}$$

$$\phi_2 = [\phi_{xx}(2) - \phi_{xx}(1)^2]/(1 - \phi_{xx}(1)^2) \qquad \text{II-14}$$

For a first order AR model $\phi_{xx}(2) \approx \phi_{xx}(1)^2$ (Eqn. II-5) and therefore, $\phi_2 = 0$ and $\phi_1 = \phi_{xx}(1)$.
The residuals a_t are independent variables, thus a_t does not depend upon its own past history a_{t-1}, a_{t-2} ..., or $E[a_t, a_{t-\tau}] = 0$ for $\tau \neq 0$
From the equation $a_t = N_t - \phi_1 N_{t-1}$ it follows that the residual variance equals

$$\sigma_a^2 = \sum_{t=1}^{N} (a_t^2) = \sum_{t=1}^{N} (N_t - \phi_1 N_{t-1})^2 \qquad \text{II-15}$$

Straight forward elaboration of the latter equation yields for AR(1) models

$$\sigma_a^2 = \sigma_N^2(1 - \phi_1^2) \qquad \text{II-16}$$

where σ_N^2 is the variance of the time series.
Verification of the relation $\phi_{xx}(2) = \phi_{xx}(1)^2$ proved the validity of the AR(1) model. A remarkable result from the model parameters listed in Table II-5 is that although the number of facilities in the various analytical sections is quite different, and also the queue levels are very differing, the time constants of the time series of the queues are quite similar.

4. The analysis times at the departments

The measuring times and interpretation times of every fourth sample arriving at the laboratory were recorded during approximately two months by the analyst who does the analysis.
The sample preparation was included in the measuring time, but the transfer times and administration times, directly coupled with the analysis of the sample were excluded. Because the means of these times were calculated during a relatively short period (in comparison to the observation period of the laboratory), and the small size of the sample, only rough estimations ($s_{\bar{x}} \approx$ 10 - 15%) of these means could be obtained (Table II-6). Because the low accuracy of the measured standard error (s_x^2) (F-test) only a rough estimate of the variation coefficient was obtained (Table II-6).

Table II-6
Statistical parameters of the measuring and interpretation times

Section	measurement time (hrs)					interpretation time (hrs)			
	number of obs.	mean	$s_{\bar{x}}$	s_x^2	c_x^2	mean	$s_{\bar{x}}$	s_x^2	c_x^2
I.r.	18	0.42	0.04	0.029	0.2 $\pm$0.5	0.72	0.11	0.20	0.4 $\pm$ 1.2
P.m.r.	132	0.48	0.01	0.020	0.1 $\pm$0.05	0.93	0.07	0.67	0.8 $\pm$ 0.7
M.s.	30	0.62	0.07	0.152	0.4 $\pm$0.8	0.88	0.12	0.45	0.6 $\pm$ 1.2
^{13}C-n.m.r.	15	1.83	0.18	0.476	0.14 $\pm$0.5	1.57	0.37	2.04	0.8 $\pm$ 3.2

5. The delays in the network

5.1 Statistical parameters of the delay times
Various delay times can be distinguished in the laboratory, depending on which group or class of samples is considered.

Firstly, the delay time of all samples with the same final analytical method, or subjected to the same number of analyses can be distinguished. Secondly, the delays in the sections, and the mean overall delay of all samples.
Considering the delays in one section, samples which were or were not analyzed before in another section can be distinguished. The means and variances of mentioned delay times are presented in Table II-7.

Table II-7

Summary of the parameters of the delays in the investigated laboratory

	Section				
parameter	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.	LAB
mean (days)	4.1	5.1	6.7	6.7	6.88
variance	24.7	21.2	47.4	37.0	55.25
best fitting F(x)	2-Er	2-Er	2-Er	2-Er	
$d_{max(\%)}$	8.43	4.27	4.09	3.52	
$d_{0.95(\%)}$	5.98	4.34	6.73	6.24	
	samples with the same final method				
	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.	
mean (days)	6.2	5.9	9.8	8.3	
variance	37.0	68.7	33.2	96.6	
	samples with the same number of analysis				
	1	2	3	4	
mean (days)	5.1	11.1	19.7	26.8	
variance	18.6	62.8	165.6	356.0	

The histograms of these delay times are presented in Fig. II-8 en II-9.
Kolmogorov-Smirnov tests executed on the delay times in the sections indicated that the two-stage Erlangian distribution was the best fitting distribution.
The probability density function of the r-stage Erlang function equals [KL75]

$$f(x) = \frac{r\mu(r\mu x)^{r-1}e^{-r\mu x}}{(r-1)!} \qquad \text{II-17}$$

for r=2 $f(x) = 4\mu^2 x e^{-2\mu x}$, with a mean $\frac{1}{\mu}$, and with a cumulative density

function equa. to

$$F(x) = 1-(2\mu x + 1)e^{-2\mu x}$$

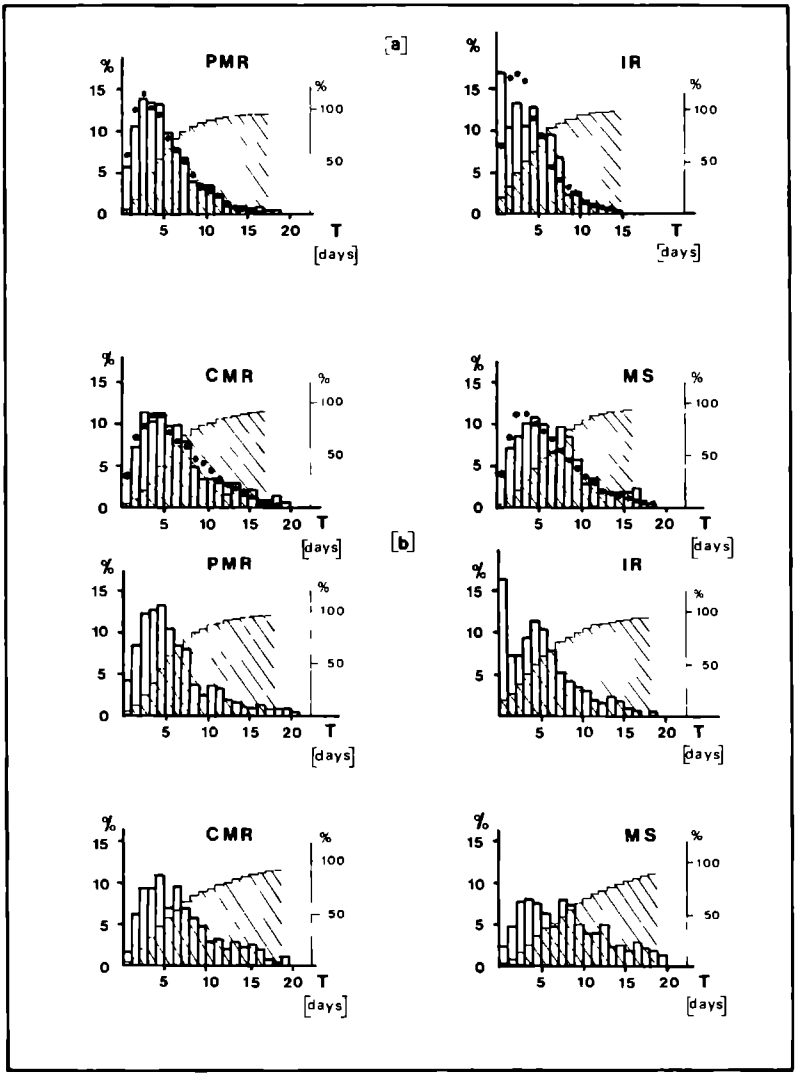


Fig. II-8: (a) Histograms of the delays (T) (days) in the sections of the laboratory. • Fitted two stage Erlangian distribution. (b) Histograms of the delays (T) (days) of the samples with the same final analytical method. Shaded figures are the cumulative density functions of the histograms.

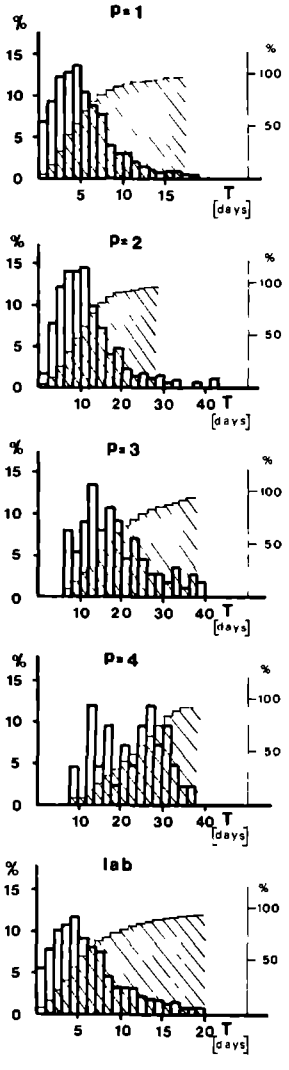


Fig. II-9: Histograms of the delays (T) (days) of the samples, which visited the same number (p) of sections Shaded figures are cumulative density functions of the histograms

In Chapter III, it will be demonstrated that this exponential shape of the cumulative density function of the waiting time is very characteristic for many waiting time systems.

The histogram of the overall delay in the laboratory indicates that the 'a priori' probability to obtain the result within 21 days equals 95%. The delay time of an individual analytical result may depend on the state of the laboratory, especially the number of waiting samples in front of the arriving sample. Fig. II-10 describes the conditional probability function of the delay, as a function of the total number of samples in the system at the moment of the arrival of the sample. It is seen that no dependence of the delay on the state of the laboratory can be detected.

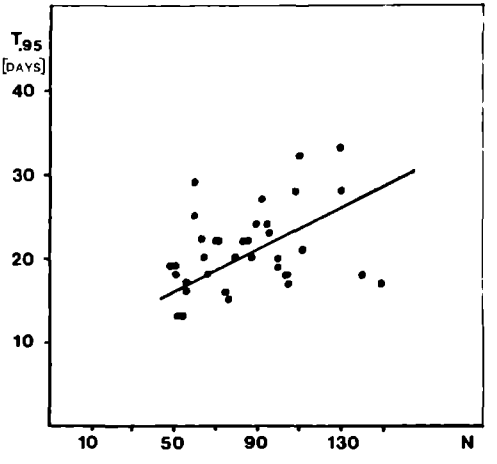


Fig. II-10: Conditional 0.95 probability limit for a sample to be analyzed within a delay ($T_{0.95}$) as a function of the number (N) of samples in the laboratory. correlation coefficient: ρ = 0.23

Table II-8 also demonstrates that the strategy on sample priorities, as applied in the laboratory, results in longer delay times at the m.s. and ^{13}C-n.m.r. sections for samples which were submitted unsuccessfully to other analytical methods.

Table II-8

Comparison of the delay times at the analytical sections for samples which were directly received from the environment, and samples which were unsuccessfully submitted to other analytical methods

Section	delay (days)				Student's t
	not submitted before		submitted before		
	mean	var($\bar{T}$)	mean	var($\bar{T}$)	value
I.r.	3.9	0.07	4.7	0.25	0.45
P.m.r.	5.1	0.02	5.3	0.20	1.46
M.s.	5.8	0.13	7.9	0.32	3.5 *
^{13}C-n.m.r.	6.1	0.10	8.9	0.43	3.6 *
AB	6.1	0.10	6.5	0.08	4.5 *

* $t_{0.005\ (200)} = 2.6$

The variance of the estimation of $\bar{T}$ (var($\bar{T}$)) in Table II-7 is corrected for autocorrelation in the data. Assuming a first order AR process, the variance of $\bar{T}$ (var($\bar{T}$)) can be calculated according to [WA75, M067],

$$\mathrm{var}(\bar{T}) = \frac{\mathrm{var}(T)}{N}\left[1 + \frac{2\phi_1}{1-\phi_1}\left(1 - \frac{1-\phi_1^N}{N(1-\phi_1)}\right)\right] \qquad \text{II-10}$$

There are two principal customers of the facilities of the analytical laboratory, denoted as F1 and F2. When the samples of both users have the same priority, one should expect their delays to be equal.
The data compiled in Table II-9, demonstrate that the F1 samples have a smaller overall delay than the F2 samples.
However, the data base of both groups was to small to allow a conclusion on which section(s) give different priority.

Table II-9

Comparison of the delays of F1 and F2 samples

Section	F1		F2		Student's t
	$\bar{T}$	var($\bar{T}$)	$\bar{T}$	var($\bar{T}$)	value
I.r.	3.81	0.04	4.92	0.35	1.76
P.m.r.	4.98	0.03	5.29	0.06	1.0
M.s.	6.68	0.15	6.98	0.52	0.37
^{13}C-n.m.r.	6.30	0.12	8.1	0.52	2.25
samples with the same exit node					
I.r.	5.33	0.24	8.88	1.14	3.02
P.m.r.	5.79	0.05	6.25	0.11	1.0
M.s.	9.49	0.43	11.42	1.65	1.3
^{13}C-n.m.r.	7.96	0.25	9.47	1.16	1.3
samples with the same number of visited sections					
1	4.82	0.02	5.63	0.074	2.67
2	11.16	0.36	11.12	0.59	0.04
3	19.41	1.5	20.67	7.08	0.43
4	28.0	14.8	24.67	15.0	0.61
overall	6.63	0.03	7.50	0.10	2.41

$t_{0.01}$(1 side) = 2.33

6. Cross correlations in the network

A property of an open network of queues is that the number of samples at the various nodes at each time point is an independent random variable [LE77] i.e. the fluctuations of the number of samples in each analytical department should not be correlated. Moreover, it is proven [LE77] that the traffic flows on the various exit arcs of the network are independent processes under equilibrium conditions. Both properties were verified by calculating the cross correlograms of the sample flow from and to each analytical section according to the algorithm:

$$\phi_{yx}(\tau) = \frac{E[(y_t-\mu_y)(x_{t+\tau}-\mu_x)]}{\sqrt{E[y-\mu_y]^2 E[x-\mu_x]^2}} \qquad \text{II-20}$$

The cross correlogram of a finite discrete time series can be estimated according to

$$\phi_{yx}(\tau) = \sum_{i=1}^{N-\tau} \frac{(y_i-\bar{y})(x_{i+\tau}-\bar{x})}{(N-\tau-1)s_x.s_y} \qquad \text{II-21}$$

The cross correlation is significantly differing from zero when

$$\phi_{yx}(\tau) \geq u(P)\sigma[\phi_{yx}(\tau)]$$

with

$$\sigma^2[\phi_{yx}(\tau)] \quad \frac{1}{N-\tau} \cdot \frac{1+\phi_{xx}(1).\phi_{yy}(1)}{1-\phi_{xx}(1).\phi_{yy}(1)} \qquad [\text{MU78}] \qquad \text{II-22}$$

The results of this test are listed in Table II-10. It shows that the number of samples in each section are not mutually correlated. This implies that a large number of waiting samples in one section does not necessarily mean that the other sections are also saturated.

Table II-10

Correlation at τ=0 between the number of samples in each section (99% confidence interval)

	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.
I.r.	1	0.3 ± 0.39	0.3 ± 0.36	0.3 ± 0.38
P.m.r.		1	0.3 ± 0.39	0.3 ± 0.38
M.s.			1	0.2 ± 0.36
^{13}C-n.m.r.				1

Cross correlations between the input flow and number of samples in the system (table II-11) show that the arrival processes do hardly depend on the state of the system. The number of samples sent to a section does not depend on the saturation of that section. Thus samples are not preferably moved to that section with the lowest saturation.
The correlations between the number of samples (x) in a section and the

delay (y) of the samples arriving at a section are significantly different from zero (Table II-12), but the correlations are too small to allow the conclusion that the fluctuations of the delay time are completely explained by the fluctuations in the number of waiting samples, whereas the residual variance is more than 90% of the total variance.

Table II-11

Correlation between the input flow (x) and number of samples in the system

	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.
ϕ_{xy}	0.301	0.405	0.263	0.303
	± 0.163	± 0.163	± 0.163	± 0.163
residual variance	0.91	0.84	0.93	0.91

Table II-12

Correlation between the number of samples (x) in a section and the delay (y) of the samples, arriving at the laboratory (99% confidence interval)

	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.	LAB
ϕ_{yx}	0.247(-5)*	0.19(-3)	0.19(-3)	0.28(-6)	0.22(11)
	± 0.18	± 0.18	± 0.17	± 0.20	± 0.21
residual variance	0.94	0.96	0.96	0.92	0.94

*time lag (τ) for which maximal correlation is observed is given in parentheses.

6. Conclusions

In this section several statistical properties on the sample flow through the spectroscopic laboratory for structural analysis of Philips Duphar have been determined. It has been demonstrated that this laboratory can be represented by a network of queues, having properties which are generally valid for networks in an equilibrium state: i.e. the conservation equation, and the independency of the number of samples in the various sections.
The queue levels in the laboratory, which are Gaussian distributed, can adequately been described by a first order Autoregressive model.

AC68 R.L. Ackhoff, M.w. Sasieni, Fundamentals of Operation Research, J. Wiley & Sons, New York, 1968

BA46 M.S. Bartlett, J. Royal Stat. Soc., B8 (1946), 27.

BO70 G.E.P. Box and G.M. Jenkins, Time series analysis. Forecasting and Control, Holden-Day, San Francisco, 1970.

GR73 P.M.E.M. van der Grinten and L.M.H. Lenoir, Statistische Procesbeheersing, Spectrum, Utrecht, 1973.

JA57 J.R. Jackson, Opns. Res., 5 (1957), 518.

KL75 L. Kleinrock, Queueing Systems, Vol I: Theory, J. Wiley & Sons, New York, 1975.

KR70 E. Kreyszig, Introductory Mathematical Statistics, J. Wiley & Sons, New York, 1970.

LE77 A.J. Lemoine, Management Sci., 24 (1977) 464.

MO67 A.L. van der Mooren, Dutch Thesis, Delft, 1967.

MU78 P.J.W.M. Muskens, Dutch Thesis, Nijmegen, 1978.

ST77 H.J. Steudel, S.M. Pandit, S.M. Wu, Management Sci., 24 (1977), 456.

WA75 H.A. Wagner, Principles of Operation Research, Prentice Hall, Inc., London, 1975.

CHAPTER 3

DETERMINATION OF THE EFFECT OF SOME DESCRIPTORS OF DELAY IN AN ANALYTICAL LABORATORY BY QUEUEING THEORY

Applications of queueing theory in analytical chemistry are restricted to the rather general ones mentioned in the introduction by Adelberg and Doerffel [AD75]. Most of all analytical results obtained with queueing theory, are derived for systems in a 'steady state'. That means that the arriving stream and service time are stochastic variables, which are described in terms of time-independent probability distribution functions. Consequently, many laboratories cannot be studied by queueing theory. For example, the sample input of some clinical and industrial control laboratories is described by a time dependent probability distribution. In the early morning the laboratory is almost empty, and by the evening all samples have been processed. Jackson [JA57] and Baskett [BA75] demonstrated that for open networks of queues, where the arrival processes do not depend on the state of the system, each node can be considered individually. Ch.II demonstrated that the arrival processes to the sections of the laboratory are indeed independent of the number of samples in the laboratory. However, here, no exact analytical results can be obtained as no theoretical results are known for complex systems: i.e. systems with batch input and output, where analyses are interrupted for other activities, and where eventually the expertise of the analysts is different. However queueing theory reveals the important variables in queueing systems. From these theoretical considerations, the relative importance of various variables can be estimated.

Published in part in:

B.G.M. Vandeginste, Communication presented at the IUPAC Congress, Helsinki (1979)

B.G.M. Vandeginste, Analyt. Chim. Acta, CTO, accepted for publication.

D.L. Massart, A. Dijkstra and L. Kaufman, Evaluation and Optimization of Laboratory Methods and Analytical Procedures, Elsevier Sc. Pu. Co.,A'dam (1978)

1. General results

1.1 Relation between the mean number of samples in the system, the mean arrival rate and the mean delay time.

A general relation exists between the mean number of samples in a queueing system, the mean input flow and mean delay time. This relation is independent of the distributions of the input flow and analysis time, and therefore is valid for all kinds of queueing systems. Furthermore, the relation depends neither on the number of analysts in the system, nor on the particular queueing discipline in the system.

Little [LI61] derived that:

$$\bar{N} = \alpha\bar{T} \qquad \text{III-1}$$

where $\bar{N}$ is the mean number of samples in the laboratory (or section), α is the mean input flow and $\bar{T}$, the mean delay time.

The validity of Eqn. III-1 for the laboratory under investigation is demonstrated in Table III-1. This relationship applies also for each individual section and each priority class of samples.

Table III-1

Verification of Little's result

All samples section	mean flow (α)	$\bar{T}$ (days)	$\bar{N}$(observed)	$\bar{N}$(calculated)	$\bar{T}$(calc)
I.r	2.8	4.1	11.5	11.5	
P.m.r.	7.7	5.1	39.1	39.3	
M.s.	2.1	6.7	13.9	14.1	
^{13}C-n.m.r.	2.5	6.7	16.8	16.7	
Lab $\Sigma\alpha_i\bar{T}_i = 81.6$	11.8	6.88	81.3	81.2	6.92
F-1 samples					
I.r.	2.07	3.81	7.9	7.9	
P.m.r.	5.2	4.97	25.6	25.8	
M.s.	1.68	6.7	11.2	11.3	
^{13}C-n.m.r.	1.93	6.3	12.1	12.2	
Lab	8.60	6.63	56.8	57.0	6.64

$\Sigma\alpha_i\bar{T}_i = 57.14$					
F-2 samples					
I.r.	0.76	4.9	3.7	3.7	
P.m.r.	2.6	5.3	13.8	13.8	
M.s.	0.41	7.0	2.9	2.9	
^{13}C-n.m.r.	0.60	8.1	4.9	4.9	
Lab	3.38	7.5	25.3	25.3	7.46
$\Sigma\alpha_i\bar{T}_i = 25.23$					

The mean delay time of all samples in an open network of q nodes can be calculated from the mean number of samples in each node and the external input flow to each node, according to

$$\bar{T} = \sum_{i=1}^{q} \bar{N}_i / \sum_{i=1}^{q} \lambda_i \qquad \text{III-2}$$

where λ_i is the external input to node i. Substituting $\bar{N}_i = \alpha_i\bar{T}_i$ in Eqn. III-2 we find that

$$\bar{T} = \sum_{i=1}^{q} \alpha_i\bar{T}_i \Big/ \sum_{i=1}^{q} \lambda_i = \sum_{i=1}^{q} \alpha_i\bar{T}_i / \lambda \qquad \text{III-3}$$

where λ is the total external flow to the laboratory.
The calculated value $\bar{T}$=6.92, using Eqn. III-3, agrees reasonably well with the observed mean delay in the laboratory. Furthermore from Table III-1, it is clear that Eqn. III-3 applies also for each class of samples: i.e. the samples originating from user F-1 and F-2. In this way the average delay in the laboratory is decomposed into its single channel components. The analysis problem therefore reduces simply to the calculation of the delay time ($\bar{T}_i$) in each section.

1.2 The utilization factor

A basic parameter in queueing systems is the utilization factor (ρ). It is the ratio of the rate at which samples enter the system to the maximum rate at which the system can perform the work, that the samples bring into the system. For a single server system, the definition of ρ becomes:

$\rho \triangleq$ average arrival rate of samples * average analysis time

$$\rho \triangleq \lambda E[AT] \qquad \text{III-4}$$

Eqn. III-4 applies only when the average analysis time is independent of the system state. Obviously a single server system can only reach a steady state when $0<\rho<1$, because for $\rho>1$, more samples arrive in the laboratory than can be analyzed, causing the number of waiting samples to grow in an unlimited fashion. This factor can be interpreted as the fraction of the time the server (analyst) is busy: $(1-\rho)$ is the fraction that the section is idle, waiting for the next sample.

Therefore ρ = E[busy time]/{E[busy time]+E[idle time]} III-5

For a system with several analysts (m) in the section, the utilization factor is defined as:

$$\rho \triangleq \lambda E[AT]/m \qquad \text{III-6}$$

2. *The basic model.*

2.1 The behaviour of systems with Markovian input.

In a system with equally spaced interarrival times and constant analysis times, no queues are formed when the utilization factor is less than or equal to one. Here, the analysis is always finished before the next sample arrives. Otherwise, for $\rho>1$, no steady state is reached and the waiting time is infinite. In all other systems queueing occurs as a consequence of the probability that a sample arrives before the analysis of the preceeding sample is finished.

2.1.1. Mean values and distribution function of waiting and system time for a M/M/1 system.

Queues are described by a shorthand notation A/B/m, where A, B and m represent the distributions for interarrival time (IAT) and service time, and the number of channels. For example, in the M/M/1 system both the interarrival time and service time are exponentially distributed, and there is only one service channel. That is the system for which most mathematical results are available. It can be easily demonstrated that a system with exponentially distributed interarrival times is a Poisson process [KL75], which has no memory. This means that the probability for a particular interarrival time does not depend on the last interarrival time. From this memoryless property of a Poisson process, with E[IAT], it follows that an idle analyst has on the average to wait a time equal to E[IAT], until a new sample arrives in the section.

Thus E[idle] = E[IAT] III-7

From Eqn. III-5 it follows that

$$E[\text{busy}] = E[IAT]\rho/(1-\rho) \qquad \text{III-8}$$

For a M/M/1 system where the samples are analyzed in the sequence of arrival at the system (First-in First-out (FIFO) rule), the mean waiting time equals [KL75]:

$$\bar{W} = \overline{AT}\rho/(1-\rho) \qquad \text{III-9}$$

and the total delay or system time equals:

$$\bar{T} = \bar{W} + \overline{AT} = \overline{AT}/(1-\rho)$$

The asymptotical shape of the representation of Eqn. III-9 in Fig. III-1 is very characteristic for all kinds of queueing systems. For the M/M/1 system it is clear that the mean waiting time depends strongly on the value of ρ.

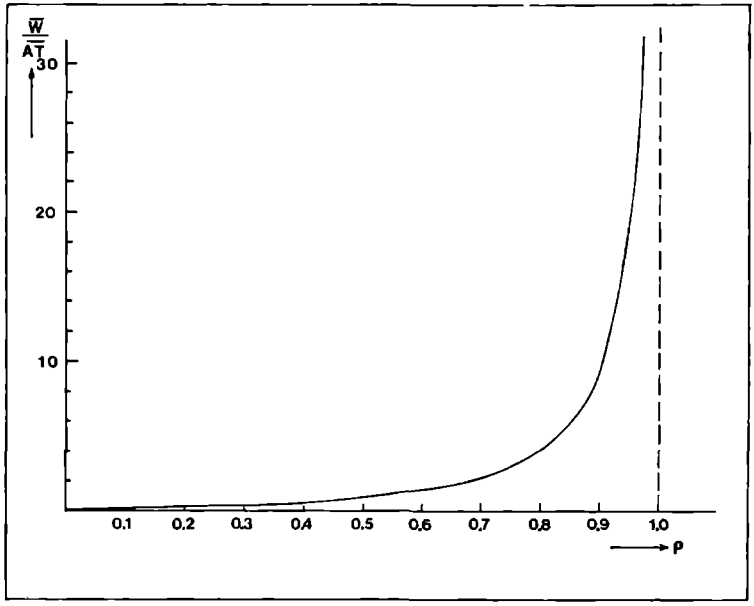


Fig. III-1: The ratio between the average waiting time ($\bar{W}$) and the average analysis time ($\overline{AT}$) as a function of the utilization factor (ρ) for a system with exponentially distributed interarrival times and analysis times (M/M/1 system).

For $\rho>0.85$ small variations in the laboratory organization may provide a serious change in waiting time. Therefore it is worthwile to investigate which sections in the laboratory are highly saturated. In contrast, channels with overcapacity ($\rho<0.5$) will be relatively insensitive to alterations of the organization. When the FIFO rule is applied in the laboratory, samples with a small analysis time have the same mean waiting time as samples with a large analysis time. The cumulative density functions of the waiting time and system time are exponential for a FIFO M/M/1 system (Fig. III-2) [KL76].

$$P(W\leq y) = 1 - \rho\exp[-(1-\rho)y/\overline{AT}] \qquad \text{III-10}$$

$$P(T\leq y) = 1 - \exp[-(1-\rho)y/\overline{AT}]$$

The graphical representation of these equations in Fig. III-2 is very similar to the shape of the cumulative density functions of the delay times observed in the laboratory (Fig. II-8). This agrees with the general constatation of

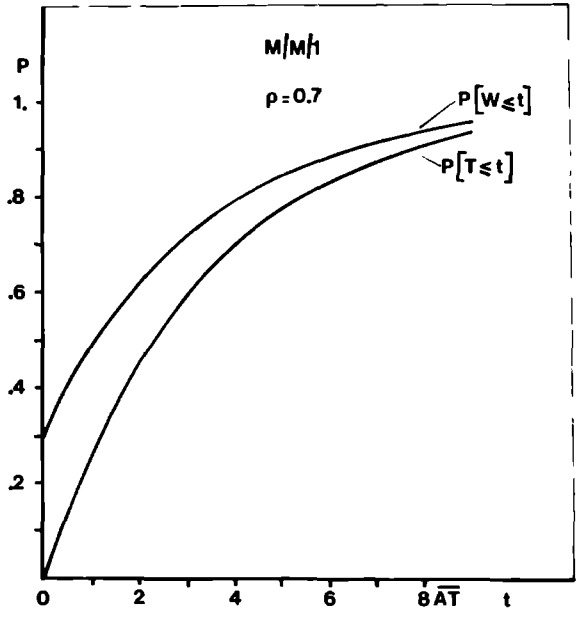


Fig. III-2: Probability (P) that the waiting (W) and system (T) time of a sample are less or equal to t. (t is expressed in units of mean analysis time), for a M/M/1 system with $\rho=0.7$.

Kleinrock [KL75] that the cumulative density function of the waiting time for

many kinds of queueing systems, has a characteristically exponential shape. The probability of finding k samples in a M/M/1 system equals [KL75]

$$p(k) = \rho^k(1-\rho) \qquad \text{III-11}$$

It is interesting to note that the probability of finding zero samples in the system equals 1-ρ. The general functional relationship $p(k)=z^k(1-z)$ is characteristic for all kinds of queueing systems, and is even derived for the general G/G/m system in a heavy traffic situation (ρ>0.9), where z is a function of ρ. However, the observed Gaussian distribution of the number of samples in the laboratory and in each section disagrees with mentioned theoretical expectation. Apparently, the backlog in the laboratory is large and the sections become never idle. Mentioned discrepancy will be explained by simulation experiments presented in Ch. VI. The conditional probability $P(t\leq y|k)$ of a sample to have a system time less than or equal to y, when it finds k samples before it at its arrival in the M/M/1 system, is given as (Appendix A):

$$P(T\leq y|k)=1 - \exp(-y/\overline{AT})\sum_{r=0}^{k} (y/\overline{AT})^{k-r}/(k-r)! \qquad \text{III-12}$$

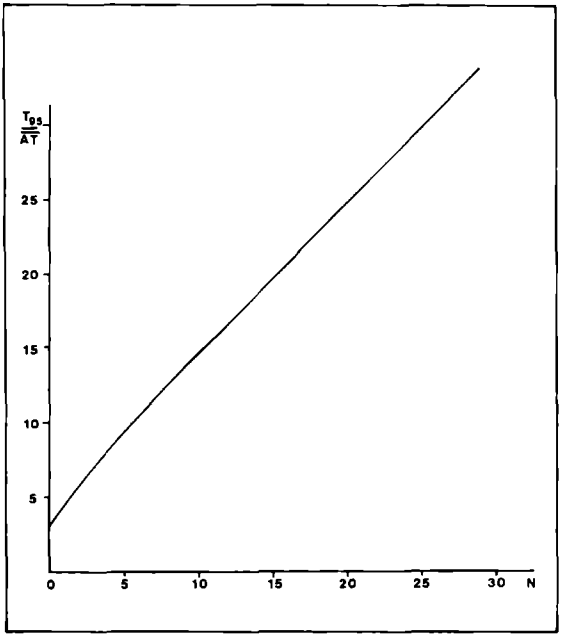


Fig. III-3: Conditional 0.95 probability limit that a sample is analyzed within a delay ($T_{.95}/\overline{AT}$) as a function of the number (N) of samples in a M/M/1 system.

In fig III-3 the maximal waiting time (95% probability) of a sample is plotted against the number of samples at its arrival in the M/M/1 system. It is clear that for large values of k, the value $dP(T \leqslant y|k)/dk$ equals the mean analysis time. Again, a major difference is found between the investigated laboratory system and a simple M/M/1 (FIFO) system, as no correlation was found between the number of samples in the laboratory at the arrival of a sample and its delay (Table II-12). A possible explanation of this difference will be given in Ch. VI.

2.1.2 Interruptions of the analysis for other activities

Two important features of queueing systems are the mean lengths of the busy and of the idle periods of the channel (or analyst). According to Eqn. III-7 the mean idle time, defined as the mean time during which no samples are present in the channel, is independent of the utilization factor. The substitution of the value $\rho=0.9$ in Eqn. III-7 indicates that the mean length of the busy time of a M/M/1 system equals 9xE[IAT]. Therefore an analyst which receives one sample per hour on the average remains 9 hrs busy. Obviously, in laboratory practice, it may happen that no analyses are done, although samples are waiting, i.e. the service is interrupted by other activities. The question arises how the mean interruption time (β) and the mean time between the interruptions (α) affect the mean waiting time. Furthermore, one may desire to compare the case where the 'on' and 'off' times are exponentially distributed with that where they are constant or scheduled. The effect depends on how the interruptions are scheduled. Two situations can be distinguished:

(i) As usual the analyst is busy as long as there are samples in the laboratory (section). However, as soon as the analyst becomes idle he starts the other activities. The duration of the other activities is a random variable with a known distribution function.

Two models may be considered:

a) The arrival of a sample during the period of other activities does not end these activities prematurely. After finishing the other activity, the analyst returns to the main queue and begins to analyze the samples, if any, that have arrived during his absence. If no samples are waiting, the analyst waits for the first arrival. When the durations of the other activities are exponentially distributed, the mean waiting time equals according to Levy [LE75]:

$$\bar{W} = \bar{W}_{fifo} + \lambda\beta^2/[\,1/(1+\lambda\beta) + \lambda\beta] \qquad \text{III-13}$$

where $\lambda\beta$ represents the ratio between the mean interruption time and mean idle time.

b) Contrary to this model, the analyst immediately starts a new period of other activities, when he finds the system empty at the end of a vacation period . According to Levy [LE75] the mean waiting time is increased by β for exponentially distributed interruptions and by $\beta/2$ for constant interruptions:

(ii) In the second situation, the analyst may also start the other activities during a busy period. The factor (R) with which the mean waiting time of the system is multiplied, when interruptions of the analyses are permitted, depends on the probability density functions of α and β. According to Fisher [FI77], when α and β are exponentially distributed R equals:

$$R = [\,\rho+(\alpha\beta)^2/(\overline{AT}(\alpha+\beta)^3)][\,1-\rho]/[\,\rho(\alpha/(\alpha+\beta)-\rho)] \qquad \text{III-14}$$

and for α and β being constant (or scheduled)

$$R = (1-\rho)/[\,\alpha/(\alpha+\beta)-\rho] \qquad \text{III-15}$$

These equations indicate that the mean waiting time in a steady state would be larger for random breakdowns than for scheduled breakdowns, with a factor (F)

$$F = 1+(\alpha\beta)^2/[\,\overline{AT}(\alpha+\beta)^3\rho] \qquad \text{III-16}$$

Here condition for a steady state is that the sum of the utilization factor of the analyst and the relative time spent for other activities is less or equal unity.

Thus $\beta/(\alpha+\beta)+\rho \leq 1$

Other activities, permitted during a busy period have a very strong influence on the mean waiting time as is demonstrated by the plots of Eqns. III-14 and III-15 in Fig. III-4. In fact, they cause a system with a low utilization factor to behave as a system with a high utilization factor. This means that the waiting time becomes asymptotically dependent on ρ, even for low utilization factors, when $\rho \to \alpha/(\alpha+\beta)$. The influence of these breakdowns in the investigated laboratory is clearly demonstrated for the i.r. section, which resembles a M/M/1 system. The observed waiting time is considerably higher

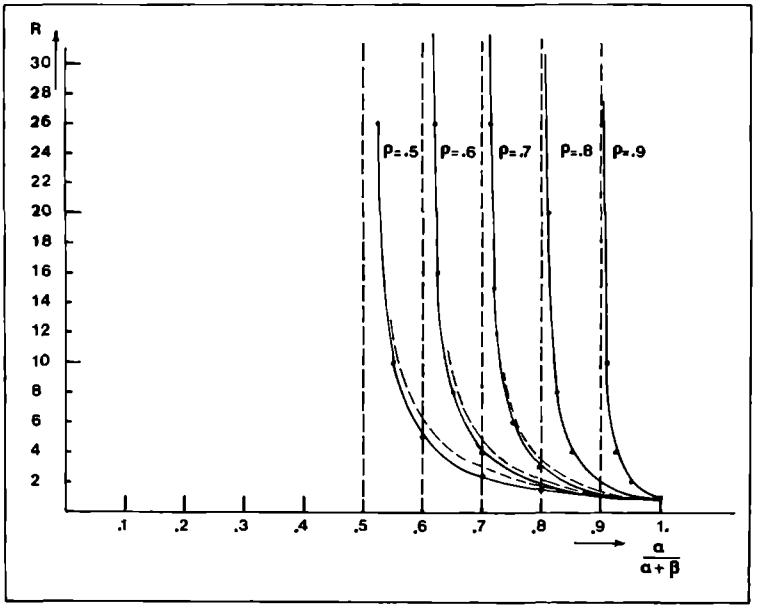


Fig. III-4: The ratio (R) between the average system time for interrupted and not interrupted analyses as a function of the available time [α/(α+β)] for analysis in a M/M/1 system with various utilization factors (ρ=0.9... 0.5) (——) time between the interruptions (α) and interruption time (β) are exponentially distributed. (---) α and β are constant.

than forecasted by Eqn. III-9. However, inclusion of 23% of exponentially distributed other activities results in a calculated waiting time that agrees reasonably well with the observed value.

Table III-2:

The effect of other activities on the mean delay.

section	flow samples per day	mean analysis time (including the transfer times)	ρ	$\bar{T}_{obs}$ (days)	$\bar{T}_{calc}$ M/M/1	$\bar{T}_{calc}$ α=15.4 β=4.6 exponential
I.r.	2.8	∿0.25hr	0.7	3.9	0.6	3.9

On the other hand, when these other activities in the i.r. section are scheduled (e.g. at the end of the day), the mean waiting time decreases to 2.7 days. Permitting the other activities exclusively during the periods

that no samples are present in that section, will cause a further decrease of the waiting time to 1.1 day. Here the relative time spent to other activities is not altered, but the mean time between the other activities is defined by the queueing process itself.

2.1.3. Influence of the distribution function of the analysis time (M/G/m). For non exponentially distributed analysis times, the mean system time depends linearly on the coefficient of variation of the distribution, defined as $C_{AT}^2 \triangleq \sigma_{AT}^2/(\overline{AT})^2$.

$$\overline{T}/\overline{AT} = 1 + \rho(1+C_{AT}^2)/[2(1-\rho)] \qquad \text{III-17}$$

Eqn. III-17 is the well known Pollaczeck-Khinchin mean value formula. Using this equation the effect of reducing the analysis time can be compared to the effect of decreasing the variance of the analysis time. Supposing that the alteration of the mean analysis time does not affect the shape of the distribution function, which means that C_{AT}^2 remains constant, then the reduction of the system time by decreasing the analysis time to x times the original value, equals:

$$\overline{T}_x/\overline{T} = [x(1-\rho)(2+x\rho(C_{AT}^2-1))]/[(1-x\rho)(2+\rho(C_{AT}^2-1))] \qquad 0<x<1 \qquad \text{III-18}$$

When C_{AT}^2 is decreased to y times the original value, then

$$\overline{T}_y/\overline{T} = [2+\rho(yC_{AT}^2-1)]/[2+\rho(C_{AT}^2-1)] \qquad 0<y<1 \qquad \text{III-19}$$

The comparison of the diagrams of both equations in Fig. III-5 demonstrates that in general, a reduction of the analysis time will improve the system more than a decrease of the variance. Under certain conditions, however, ($x>0.8$ and $y<0.2$) it will be beneficial to reduce the variance.

The analysis time can be decreased in various ways: e.g. only a single result may be presented rather than duplicates. The coefficient of variation can be decreased by standardizing or automating parts of the analytical procedure. However, an alteration of the analytical procedure may influence the accuracy of the analytical result. Therefore cost-profit analyses should indicate whether the profit of obtaining the analytical result within a shorter time balances against the costs of the eventually introduced inaccuracy. In the particular case of structural analysis, measurements are not normally duplicated, and standardization is difficult. Here the only way to influence the

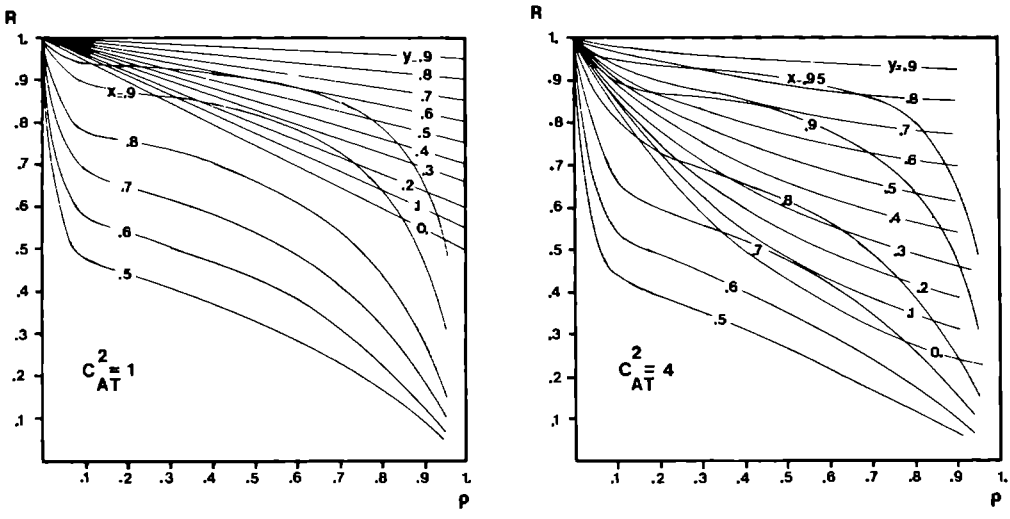


Fig. III-5: Reduction (R) of the system time for two M/M/1 systems with different variation coefficients of the analysis time (C^2_{AT}), as a function of the utilization factor (ρ), when (i) the mean analysis time ($\overline{AT}$) is reduced to $x\overline{AT}$ ($0 < x < 1$) and (ii) the variation coefficient (C^2_{AT}) of the analysis time is reduced to yC_{AT} ($0 < y < 1$).

parameters of the probability density function of the analysis time is to disrupt the interpretation of the spectra after a certain time (X) and to urge the analyst to transfer the problem to another spectroscopic method. Then the original exponential distribution of the analysis time takes the shape shown in Fig. II-6, with a mean equal to

$$\overline{AT}(1 - \exp(-X/\overline{AT})) \qquad \text{III-20}$$

and a coefficient of variation equal to:

$$2[\exp(-X/\overline{AT})(X-\overline{AT}) + \overline{AT}] / [\overline{AT}(1 - \exp(-X/\overline{AT}))^2] - 1 \text{ (Appendix B)} \qquad \text{III-21}$$

When these terms are substituted in Eqn. III-17, the reduction of the waiting time is found as a function of the ratio between the maximal and original mean analysis time, for various values of the utilization factor. As Fig. III-7 demonstrates, the truncation of the analysis time has the greatest effect at high ρ levels. However, in the laboratory for structural analysis, two types of analyses are truncated: analyses that should be successfully finished, if the analyst was allowed to study the spectrum for a longer time,

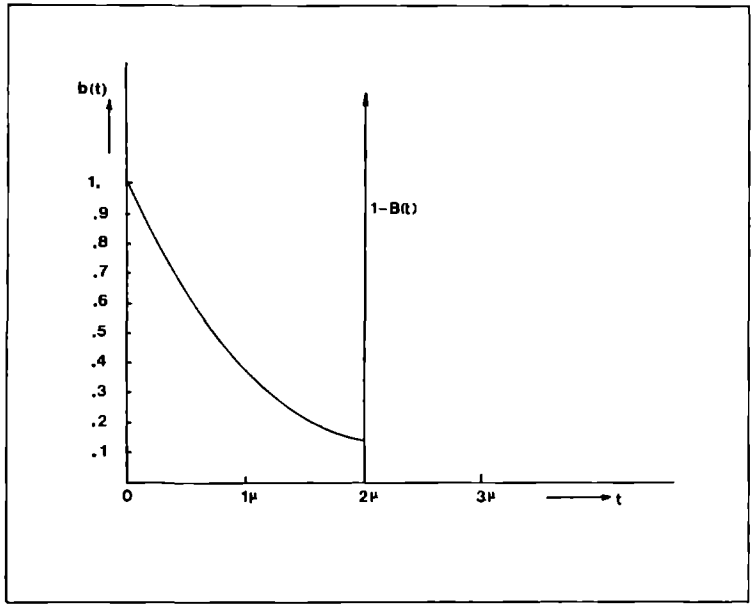


Fig. III-6: The truncated analysis time probability density function: $b(t) = \mu \exp(-\mu t)$ *for* $t < 2\mu$. $b(t=2\mu) = 1 - B(2\mu)$ *for* $t = 2\mu$, *with* $B(2\mu) = \int_0^{2\mu} \mu e^{-\mu t} dt$

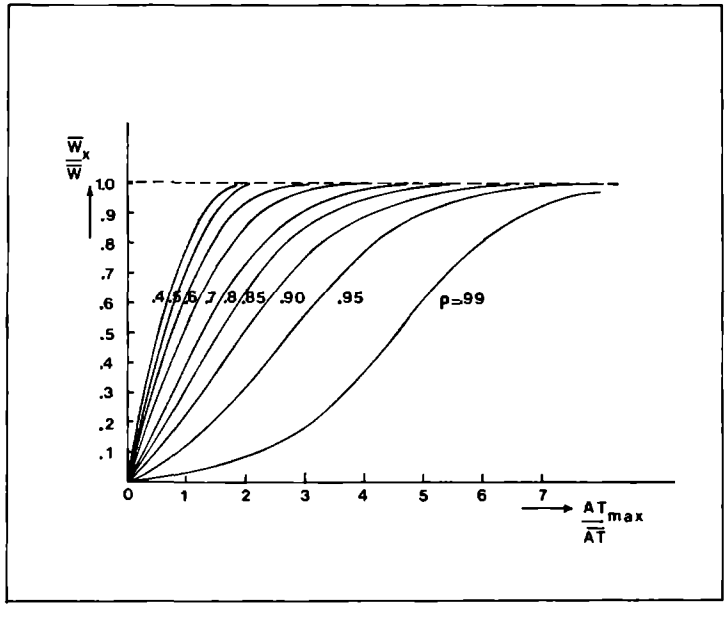


Fig. III-7: The reduction of the mean waiting time $(\bar{w}_x/\bar{w})$ *as a function of the maximal allowed analysis time* $(AT_{max}/\overline{AT})$ *for various utilization factors* ρ.

and analyses which should not. Consequently, the truncation of analyses in one section may increase the flow towards the other sections, and, therefore, increase the overall waiting time. It depends on the utilization factors of the other sections whether the performance of the total system is improved or not, when analyses are truncated, as the simulations will demonstrate in Chapter VI.

2.1.4 Batch input systems

An alternative for separate arrivals of the samples to the laboratory, is to collect the samples during a certain period, and sent them simultaneously to the laboratory. This transforms the original M/M/1 system with a mean interarrival time E[IAT], and mean analysis time E[AT] to a M/G/1 or D/G/1 system. In the latter system the batches of samples arrive at equispaced times in the laboratory. As mentioned before 'G' means that the analysis time has a general probability density function. The mean analysis time of a batch ($\overline{AT}_b$) with a mean size $\bar{r}$, equals $\bar{r} \cdot \overline{AT}$. The variance of the analysis time equals var[AT]($\bar{r}$ + var(r)), and the mean interarrival time of the batches is $\bar{r} \cdot \overline{IAT}$.

In this model the delay seen by a sample now consists of two independent components: the delay of the first member of his batch to be analyzed, and that due to the analysis times of the preceeding members of his batch [BU75]. When a comparison is made between two models with the same total sample flow, one without batch and the other with batch input, then the ratios compiled in Table III-3, between the mean system times are found (Appendix C). The equations in Table III-3 are derived with the assumption that the 'overhead' is not changed because of 'batch' analysis. This means that the fact that during the analysis of one sample, the sample preparation of the next sample may be started, was not accounted for.

Table III-3

Ratio between the delay ($\bar{T}_{batch}/\bar{T}$) without batch input ($\bar{T}$) and with batch input ($\bar{T}_{batch}$).

p.d.f. of the batchsize	p.d.f. interarrival time of the batches	
	Constant	Exponential
Poisson	$0.5\bar{r}(1-\rho)+1$	$1+0.5\bar{r}$
constant	$0.5\bar{r}(1-\rho)+0.5$	$0.5+0.5\bar{r}$
exponential	$0.5\bar{r}(2-\rho)+0.5$	$0.5+\bar{r}$
Gaussian	$\frac{1}{2}\frac{\sigma_r^2}{\bar{r}} + \frac{1}{2}\bar{r} + \frac{1}{2} - 0.5\bar{r}\rho$	$\frac{1}{2}\frac{\sigma_r^2}{\bar{r}} + \frac{1}{2}\bar{r} + \frac{1}{2}$

At a first sight, one might expect that a batch input would result in a better performance of the system. However, Table III-3, indicates that the mean delay time will only be improved when all batches have an equal size or are Gaussian distributed, with the condition that the batches enter the laboratory at equidistant times. Furthermore, for Gaussian distributed batch sizes, improvement is only achieved under certain conditions of $\sigma_r^2/\bar{r}$ (Fig. III-8).

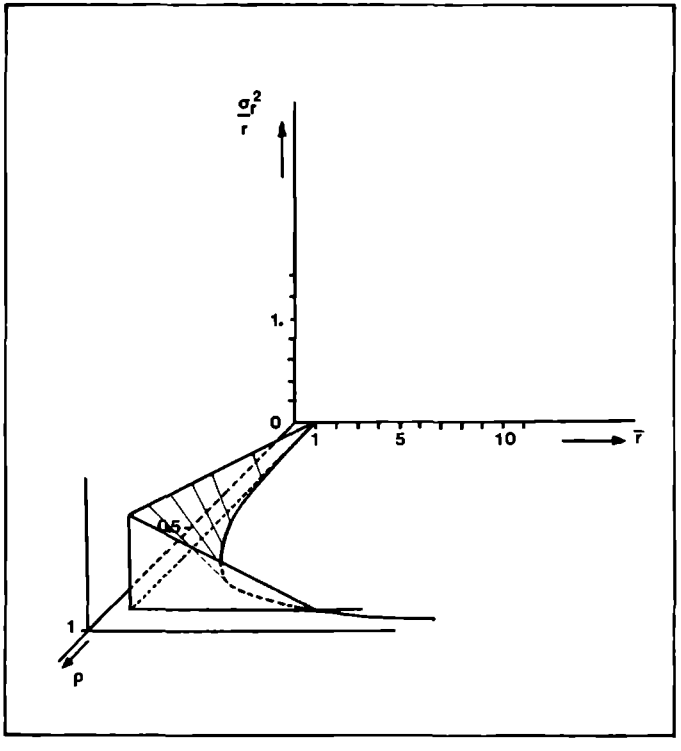


Fig. III-8: Upper bound (shaded plane) of the variation coefficient (σ_r^2/r) and the mean batchsize ($\bar{r}$) of the sample input, as a function of the utilization factor (ρ) to obtain a reduction of the mean delay, by transforming a M/M/1 system in a batch input system with Gaussian distributed batchsizes and equidistant arrivals

As mentioned before (section II), the arrivals in a M/M/1 system have a Poisson distribution. When all the samples enter the laboratory simultaneously, once a day, then, the batch sizes are also Poisson distributed. Here no improvement of the system is obtained, and the results become even worse when

the mean sample flow (λ) in the system is high ($\bar{r} = \lambda$).
The equations of Table III-3 were verified for a M/M/1 system, with $\overline{IAT} = 1$, $\overline{AT} = 0.7$ and $\bar{r} = 1$, which was transformed to a batch input system with $\bar{r} = 8$ and $\overline{IAT} = 8$. The results of the simulations agree reasonably well with the theoretical forecast (Table III-4)

Table III-4

$\bar{T}_{batch}/\bar{T}_{single}$

p.d.f. of the batchsize	p.d.f. of the interarrival times of the batches				
		exponential		constant	
	σ_r	theoretical	simulated	theoretical	simulated
Poisson	2.8	5	8.5	2.1	2.8
Gaussian	1	4.5	4.9	1.7	1.7
	2.8	5.0	4.7	2.2	2.0
	4	5.5	5.2	2.7	2.1
exponential	8	8.5	5.6	4.6	3.8
Gaussian ($\bar{r}$=2; $\overline{IAT}$=2)	0			0.8	0.8

In practice, a spectroscopic analysis consists of two parts: the measurement and interpretation of the spectrum. However, that M/M/1 model does not account for the fact that at first, the total batch of spectra is measured, whereafter, the interpretation is started. Simulation experiments, presented in Chapter VI account for this fact and for the mentioned reduction of overhead.

2.2 Priority queueing

The equations in above paragraphs are derived for first-in-first-out (FiFo) disciplines: i.e. all samples are analyzed in the sequence of arrival at the service channel. Of course, there may be many reasons to deviate from this FiFo rule, some samples being given priority. In an analytical laboratory, priority can be given to samples depending on their origins (different research groups), their analysis time ('easy' and 'difficult' samples), or their history in the laboratory (first, second analytical method which is tried). A particular priority difference can be obtained by attributing urgency numbers to the samples, which may be a function of the waiting time.

For example [KL76]

$$q_p^{(r)} = (t - \tau)^r b_p \qquad \text{III-22}$$

where $t - \tau$ is the waiting time of the sample at time t and b_p is the urgency parameter for priority class p. The samples are analyzed in the sequence of decreasing urgency numbers (q). The advantage of this priority system is that analytical results are available in queueing theory. However, in the practice of an analytical laboratory, this priority rule is hard to operate. since each time the analysis of a new sample is to be started, the priority value q_p should be calculated of each sample in queue, in order to find the sample with the highest priority. For large values of r, the sequence becomes FIFO, while for small values of r an absolute priority discipline is obtained, because all urgency numbers become equal to b_p. In a system of head-of-the line (H.O.L.) or absolute priority, samples queue according to priority groups and are strictly separated on the basis of the group to which they belong. Kleinrock [KL76,KF76] calculated the mean waiting time for a M/M/1 system with time dependent priorities (Eqn. III-23)

$$\bar{W}_p = [W_0/(1-\rho) - \sum_{i=1}^{p-1} \rho_i \bar{W}_i (1-(b_i/b_p)^{1/r})] / [1 - \sum_{i=p+1}^{P} \rho_i (1-(b_p/b_i)^{1/r})] \qquad \text{III-23}$$

with p=1, 2, ...P and $b_1 < b_2 < \ldots\ldots < b_p$

where $W_0 = \sum_{i=1}^{P} \lambda_i \overline{x_i^2}/2$ with $\overline{x_i^2}$ equal to the second moment of the analysis times of samples from group i.

For an exponentially distributed analysis time $\overline{x_i^2}$ equals $2(\bar{x}_i)^2$, consequently

$$W_0 = \sum_{i=1}^{P} \rho_i \bar{x}_i$$

Substituting r=0 in Eqn. III-23, the expression for the absolute priority discipline is obtained

$$\bar{W}_p = (W_0/(1-\rho) - \sum_{i=1}^{p-1} \rho_i \bar{W}_i)/(1-\sum_{i=p+1}^{P} \rho_i) \qquad \text{III-24}$$

For a system with only two priority groups Eqn. III-24 becomes:

$$\bar{W}_1 \text{ (low priority)} = \sum_{i=1}^{2} \rho_i \bar{x}_i / [(1-\rho)(1-\rho_2)] \qquad \text{III-25}$$

$$\bar{W}_p \text{ (high priority)} = \sum_{i=1}^{2} \rho_i \bar{x}_i / (1-\rho_p)$$

From the graphical representation of Eqn. III-23 in Fig. III-9, it is easily seen that by choosing appropriate values of r and b, each ratio between the mean delay times of the various classes can be obtained, ranging from equal priority to an absolute priority discipline. Moreover, from Eqn. III-23 it can be calculated that for equal mean analysis times of the various priority groups, the mean waiting time of the entire population of samples is not influenced by any priority discipline, and equals the delay of the M/M/1 system. Accordingly optimization of b and r is not achieved by minimizing the mean waiting time exclusively, but should include cost functions for waiting [KL76] for the different priority groups of samples.

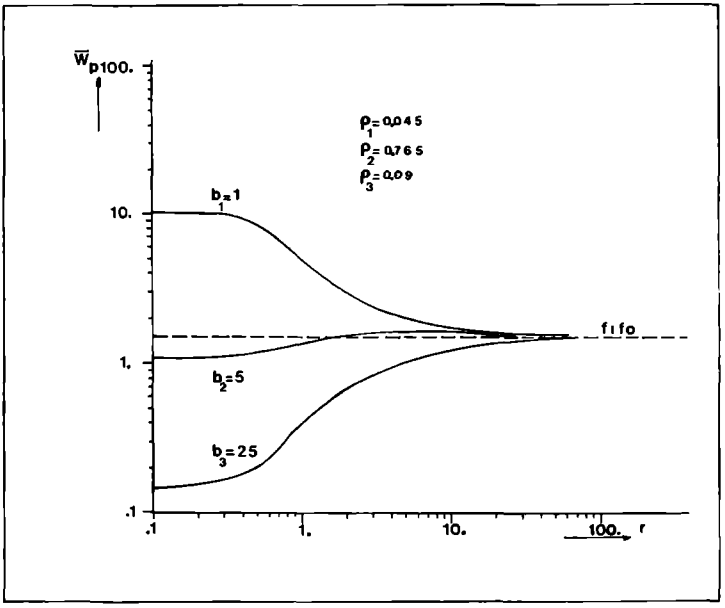


Fig. III-9: Time dependent priorities, varying from absolute priority (r=0.1) to a First-in First-out (FIFO) discipline (r=100) for three priority groups with $b_1:b_2:b_3$ = 1:5:25

A very interesting property of priority queueing is that analytical results are obtainable for systems with utilization factors greater than 1. For a system where group (p+1) should give absolute priority to group (p), Eqn. III-24 can be rewritten as:

$$\bar{W}_p = \sum_{i=1}^{P} \rho_i \bar{x}_i / [(1-\sum_{i=1}^{p} \rho_i)(1-\sum_{i=1}^{p} \rho_i)]$$ III-26

This equation demonstrates that all priority groups (1...p), for which $\sum_{i=1}^{p} \rho_i < 1$ reach a steady state. All other groups (p+1,...P) are oversaturated and have an infinite delay. This effect is demonstrated in Fig. III-10, where the groups get successively saturated with increasing input flow to the system, under the condition of unchanged flow ratio to the various groups.

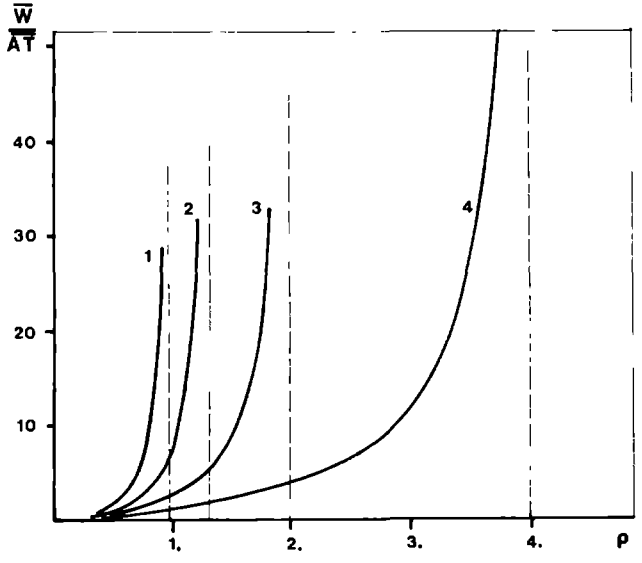


Fig. III-10: The mean delay ($\bar{W}/\bar{AT}$) of 4 priority groups in a M/M/1 system as a function of the total utilization factor (ρ) of the system where (i) each priority group has the same utilization factor (ρ_i) and (ii) the (i+1)-th group has absolute priority on the i-th group.

For a system with two priority groups with an equal mean analysis time, but with a different input flow to the system, it is interesting to investigate which of the two groups is the most sensitive to the applied priority rule (i.e. an absolute priority is attributed to group1 or to group 2). In Fig. III-11, the relative delays ($\bar{T}_1/\bar{T}$ and $\bar{T}_2/\bar{T}$) of two sample groups (1 and 2) are plotted versus the ratio of the input flow of both groups of samples. Fig. III-12 shows the plots of the increase of the delay for both groups when their priority is inverted. These graphs demonstrate that the delay of

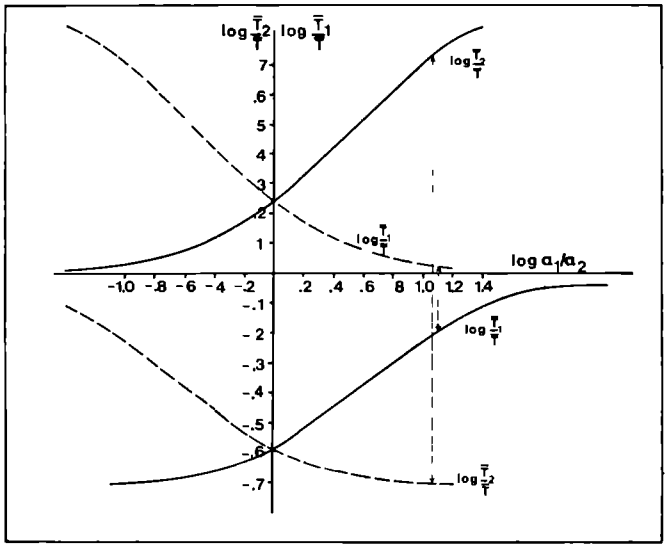


Fig. III-11: The system time of two groups of samples with equal analysis times as a function of the ratio of their input density (a_1/a_2) in a system with a total utilization factor $\rho=0.9$. (——) group 1 has absolute priority, (---) group 2 has absolute priority.

the group with the largest (group 1) input flow is the less sensitive to the priority rule. For example: suppose that the input flow of the first group is ten times that of the second group, and attribute the absolute priority, first assigned to the first group, now to the second group, then, the delay of the latter group (smallest input flow) will be reduced by a factor 10, while the delay of the former group is only doubled.

Holtzman [HO70] analyzed a dynamic priority discipline. Arriving samples at a queue are assigned urgency numbers, just as in a time dependent priority rule. However, the sample with the smallest sum of urgency number (b_p) and arrival time (t) is analyzed next ($q_p=t+b_p$). This service discipline has also the effect of considering some samples to have higher priority than others, but takes into account the undesirability of having low priority samples to wait too long. Identically to the time dependent priority discipline, the dynamic priority discipline may be altered from a FIFO ($b_p=0$), to an absolute priority discipline (b_p is large). Unfortunately, no exact analytical results are known for this priority discipline. Only upper and lower bounds of the

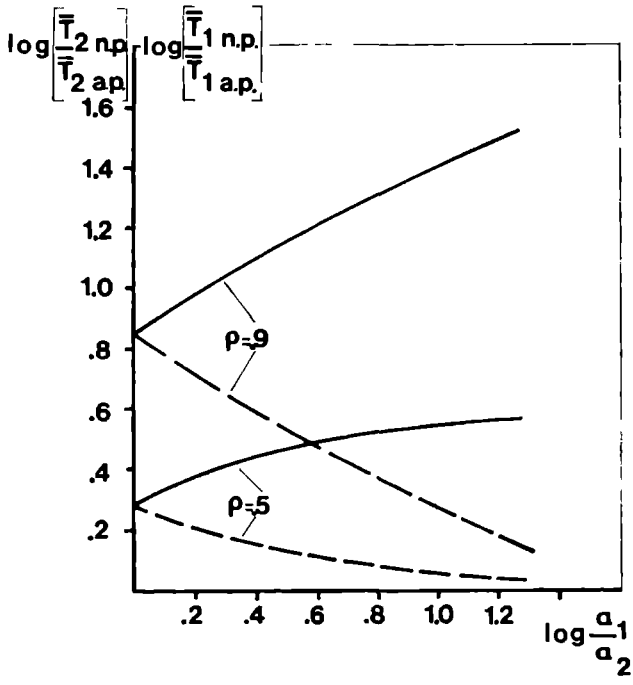


Fig. III-12: A system with two groups of samples: The influence of the attributed priority on the delay of both sample groups as a function of the ratio of their input density. (——) T_1 n.p./T_1 a.p. : ratio between the delays of the first group having no priority (n.p.) and absolute priority (a.p.) (---) T_2n.p./T_2 a.p. : idem for the second group of samples.

waiting time under equilibrium conditions can be given [HO70]. However, the advantage of this priority rule is that the sample is immediately scheduled in the queue at a fixed position.

When the mean analysis times of the various priority groups are different, the lowest overall mean waiting time is found when the samples with the shortest analysis time get absolute priority (Eqn. III-24). In an analytical laboratory, this situation occurs when an analyst does two different analyses, or when samples can be subdivided into two groups: e.g. so called 'easy' and 'difficult' samples, with 'small' and 'large' analysis times respectively: i.e. the interpretation of a spectrum may be easy or difficult. Particularly, Conway [CO67] indicated in many examples, that this separation into two groups provides a considerable reduction of the mean waiting time, as opposed to the FIFO system. The following example demonstrates the effect of subdividing the samples into two categories. Starting from an exponentially distributed analysis time, with

a mean $\overline{AT}$, the samples are subdivided into two categories: the first one with analysis times in between 0 and x and a second category with $x<AT<\infty$ (Fig. III -13).

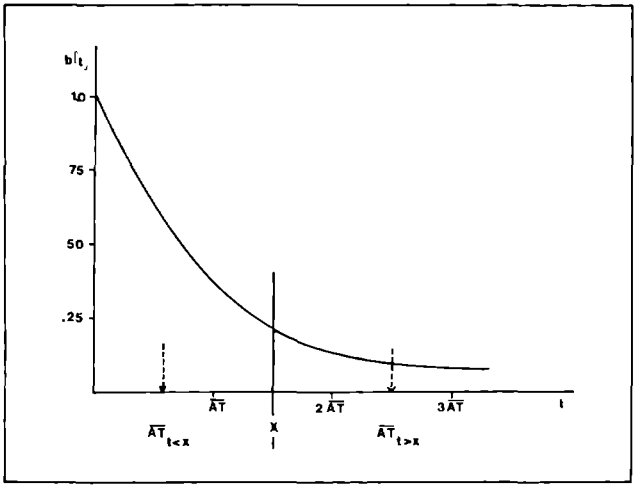
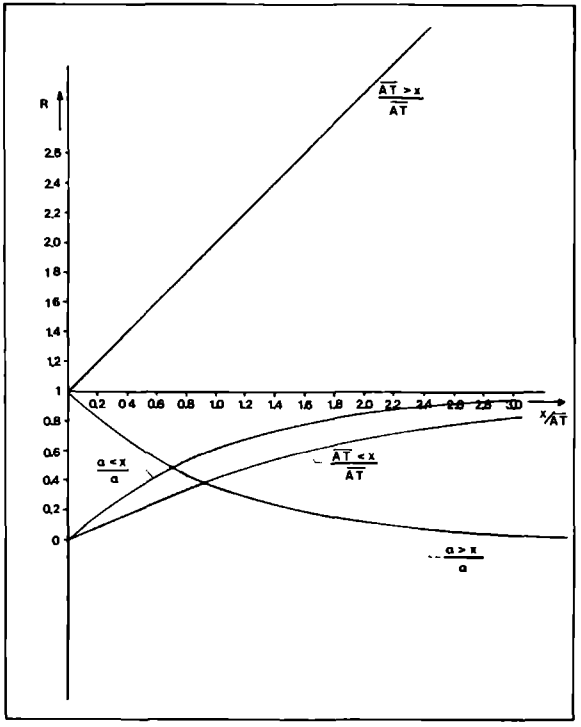


Fig. III-13: Distribution of the samples over two groups: (1) 'easy' samples for which $AT<x$, with a mean $\overline{AT}<x$, (2) 'difficult' samples for which $AT>x$, with a mean $\overline{AT}>x$.

Fig. III-14 and 15 show the mean analysis times and the sample flows of both categories of samples as a function of x for exponentially and k=4 Erlangian distributed analysis times. The plots of the calculated overall waiting time when the first category of samples has absolute priority (Fig. III-16 and 17) clearly show that the total mean system time is approximately halved for high utilization factors (ρ=0.9). The effect is maximal when a small number of samples (10%) with high analysis times must give absolute priority to all other samples. A minor reduction is obtained for the k=4 Erlangian system (30%) (see Appendix D for the derivation of the equations). Furthermore, the mean system times of both priority groups differ considerably: i.e. in a M/M/1 system with ρ=0.9, the mean waiting time of the samples with a high priority (60% of all samples) is only 7% of that of the samples with low priority. It should be stressed here that for these calculations, the correct class was assumed to be determined for each sample. The influence of inaccurate estimations of the analysis time of the samples will be demonstrated by the simulations presented in Chapter VI. However, clearly, when the analysis time cannot be estimated at all, it has no sense to divide the samples into



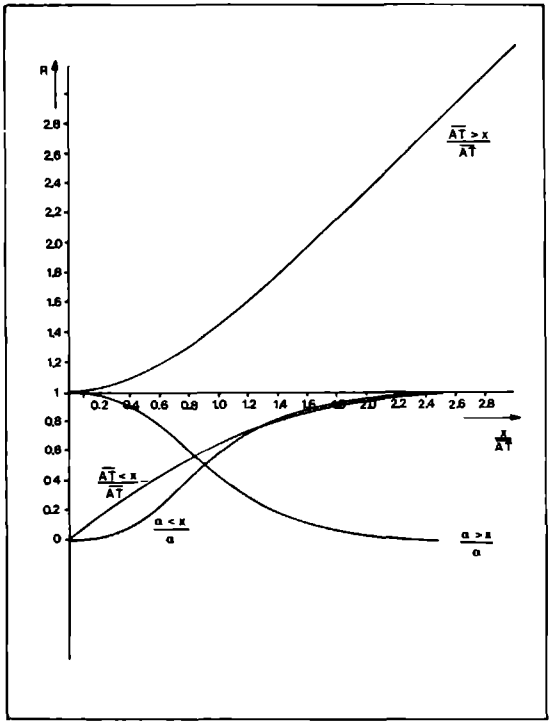


Fig. III-14: Distribution of samples with exponentially distributed analysis times over two groups.
Fig. III-15: Distribution of samples with k=4 Erlangian distributed analysis times over two groups. The mean analysis time ($R=\overline{AT}<x/\overline{AT}$) and fraction ($R=\alpha<x/\alpha$) of 'easy' samples as a function of $x/\overline{AT}$.

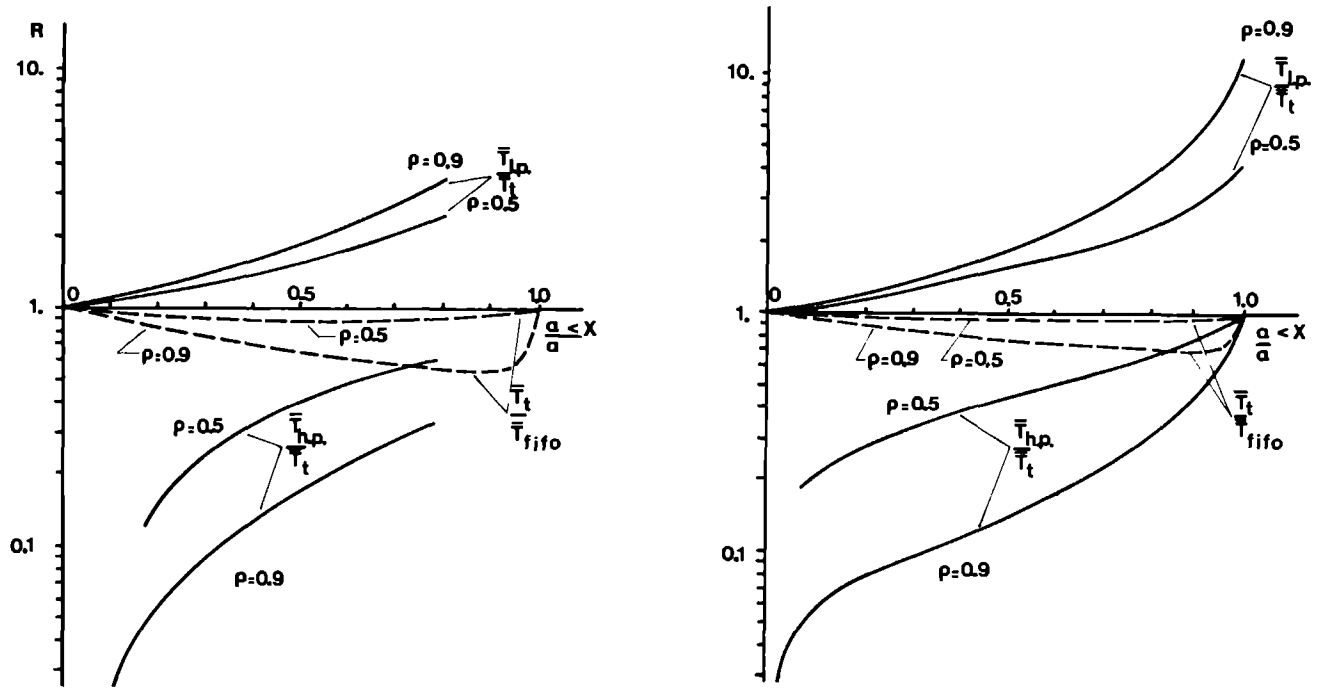


Fig. III-16: Ratio (R) of the mean delay of the 'difficult' ($\bar{T}_{l.p.}/\bar{T}_t$) samples, the 'easy' ($\bar{T}_{h.p.}/\bar{T}_t$) samples and of the overall delay, as a function of the fraction of 'easy' samples' ($\alpha<x/\alpha$) in a M/M/1 system, where the easy samples have absolute priority. (——) Reduction of the overall mean system time ($\bar{T}_t/\bar{T}_{fifo}$) by discriminating easy and difficult samples.

Fig. III-17: Idem for a $M/E_4/1$ system.

two categories on the basis of an estimated analysis time,as then, the samples are positioned at random in the queue. The parameters of the waiting time (mean and variance) are equal for a random and FIFO queue discipline (Durr [DU71]). Another way to obtain a discrimination in favour of short analyses, is to transfer the spectrum after a fixed interpretation time q to a pile of unfinished spectra and to start the measurement of the next sample or the interpretation of the next spectrum. This is a common method of job handling in time shared computer systems, which is known Round Robin (R.R.) scheduling. Kleinrock [KL76] demonstrated that R.R. scheduling has no effect on the mean waiting time of the total population of samples, but diminishes the mean waiting time of the easy samples at the expense of the difficult ones. For q approaching zero, the waiting time of a sample (W_x) in a M/M/1 system becomes linearly proportional to its analysis time(x).

$$W_x = x\rho/(1-\rho) \qquad \text{III-27}$$

However, subdividing the interpretation time of a spectrum in infinitessimally small steps is unrealistic. Therefore the effect of applying R.R. scheduling was simulated here only during the interpretation step, and for large values of q. Fig. III-18a,b shows the results of simulations of a laboratory system where the mean interpretation time for difficult samples is twice that for easy samples. For a negligibly small measurement time 'easy' samples wait for shorter times only when the interpretation is done in steps smaller than 0.1th of the mean interpretation time (q=0.1). Moreover, the effect is considerably decreased when the measurement and interpretation times are approximately equal, and only the interpretation step is partitioned (Fig. III-18b). These simulations led to the conclusion that in spectroscopic analysis waiting time can be improved considerably only by giving priority to the 'easy' samples which have previously been recognized as such.

One may imagine that in some cases one is not interested in a minimal delay time, but prefers a uniform response time to users. In such situation, the variance performance measure is important and the question of minimizing the variance of delay time should be tackled. Merten and Muller [ME72] have shown that the sequence that minimizes the variance of waiting time is antithetical to the sequence that minimizes the mean waiting time. They proposed some heuristic method to schedule the samples in a V-shaped sequence: i.e. the samples must be arranged in descending order of analysis time if they are placed before the shortest job, but in ascending order of analysis times if

placed after it.

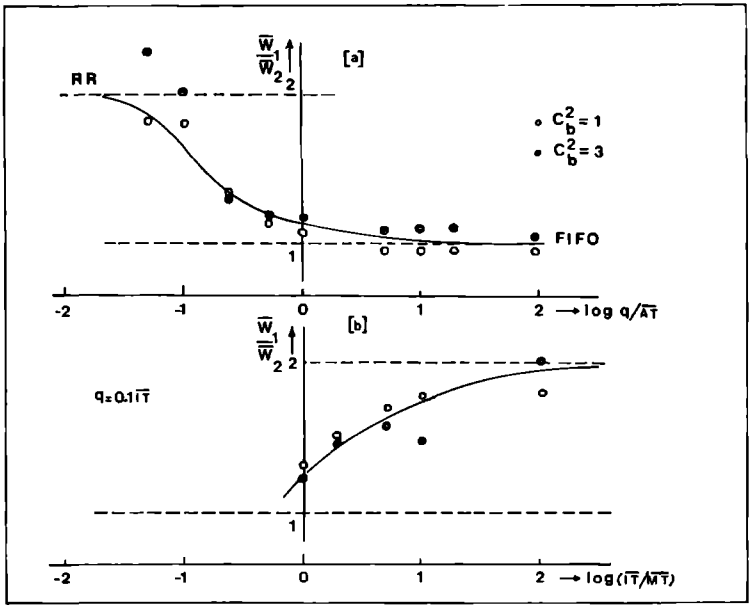


Fig. III-18a: The effect of applying Round Robin (R.R.) scheduling in a system with two groups of samples, where $(\overline{AT})_1=2(\overline{AT})_2$, on the waiting time of both groups, as a function of the length $(q/\overline{AT})$ of the steps with which the samples are interpreted. O the variation coefficient of the analysis time $(C_b^2)=1$; ● the variation coefficient of the analysis time $(C_b^2)=3$.

b: The effect of applying RR scheduling with $q=0.1\overline{IT}$ as a function of the ratio between the mean interpretation time $(\overline{IT})$ and the mean measuring time $(\overline{MT})$ of the samples.

2.3. Dynamic aspects of M/M/1 systems.

In Chapter II a time series approach was used as a practical means to obtain a model of the queues in the laboratory. To date, very little has been done in applying time series techniques to analyze queueing systems, and the theoretical expressions (Eqn. III-28) for the autocorrelation function of the number of samples are only derived for M/M/1 systems [M055].

$$\phi_{xx}(\tau) = \exp(-\tau(1-\rho)^2/\rho\overline{AT}) \qquad \text{III-28}$$

with a time constant $T_x = \overline{AT}\rho/(1-\rho)^2$ III-29

The time constant T_x represents the mean time for the queue to return from any deviation from the mean level ($\bar{N}$) back to 0.368 of this deviation. Therefore, the time constant is a measure of the rate of the queue size fluctuations, Fig. III-19 demonstrates that the utilization factor of the system has a considerable influence on the time constant of the system.

Recalling the fact that the i.r. section resembles the most to a M/M/1 system with the parameters presented in Table III-2, a time constant of 1,9 days should be observed. The larger value of 5 days found is probably due to the other activities of the analyst, causing the system to behave as a system with a larger utilization factor.

From the study of Bhat [BH72], who describes the transient behaviour of queueing systems, time constants for other systems could be calculated (Fig. III-19). Although these step responses were given in terms of the number of departures needed for the value $\rho(1 - \rho)^{-1}$ to reach 0.623 of its 'steady' state value, a similar relationship between T_x and ρ is found as given by Eqn. III-29. Comparing the lines calculated for $E_5/M/1$ and $M/E_{10}/1$ systems it seems that the strong dependence of T_x from ρ seems a general characteristic of all queueing systems.

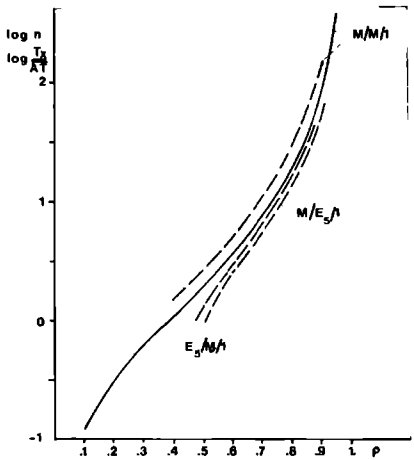


Fig. III-19: (——) the fractional time constant $T_x/\overline{\Delta T}$ of the number of samples in a M/M/1 system as a function of the utilization factor (ρ)
(---) the number of departures (n) needed for the mean number of waiting samples to reach 0.62 of its 'steady state' value as a function of ρ for a M/M/1, $M/E_5/1$, and $E_5/M/1$ system.

In a congested system the observations in a sample record are correlated. This complicates seriously the statistical analysis of simulated queueing data (Chapter V).

Studies involving time series analysis in order to model queue data, have been executed by Steudel [ST76]. He described the number of samples in a M/M/1 system with a discrete autoregressive model of order one, AR (1), in the form of $n_t = \phi_1 n_{t-1} + a_t$. Our calculations in Chapter II proved that this AR (1) model is also valid for more complex systems that the M/M/1 system.

2.4 Many server systems

Considering m-server systems, it should be indicated that analytical results are often not available, and the derivation of the relationship between the mean waiting time and the utilization factor is very difficult. For example very few substantive results can be given for the M/G/m and G/G/m system [KL76]. Therefore, several upper and lower bounds were derived for these systems. Perhaps the most important of these is the lower bound for the G/G/m system, being [KL76]

$$\bar{W} > \frac{\rho^2(C_b^2)-\rho(2-\rho)}{2\lambda(1-\rho)} - \frac{[(m-1)/m]\overline{AT^2}}{2\overline{AT}} \qquad \text{III-30}$$

For m=1, approximations for heavy-traffic situations can be derived. These results are extremely robust and give the general behaviour of queues with long waiting times.

The average waiting time is given by

$$\bar{W} \simeq \frac{(\sigma_a^2+\sigma_b^2)}{2(1-\rho)\overline{AT}} \qquad \text{III-31}$$

with σ_a^2: variance of the interarrival times
σ_b^2: variance of the analysis times

And the probability that $W \leq y$ equals

$$P(W \leq y) = 1-\exp\left(\frac{-2\overline{AT}(1-\rho)}{\sigma_a^2+\sigma_b^2} \cdot y\right) = 1-\exp(-y/\bar{W}) \qquad \text{III-32}$$

However, analytical results are available for M/M/m systems.

The average waiting time equals

$$\bar{W} = p_o \frac{(m\rho)^m}{m!(1-\rho)^2} \cdot \overline{AT} \qquad \text{III-33}$$

where

$$P_o = \left[\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \frac{(m\rho)^m}{m!} \frac{1}{1-\rho} \right]^{-1} \qquad \text{III-34}$$

and

$$\rho = \frac{\lambda . \overline{AT}}{m} \qquad \text{III-35}$$

When the service time is k-Erlang distributed the approximation of Maalφe [MA70] is very useful

$$W(\rho, m>1, k>1) = p_o \frac{(m\rho)^m}{m!(1-\rho)^2} \cdot \overline{AT} \cdot \frac{(1+1/k)}{2} \qquad \text{III-36}$$

The graphs of Eqn. III-33 in Fig. III-20 where the utilization factor of each analyst is independent of m, indicate that the asymptotical rise of the waiting time for large systems, with one queue served by several analysts starts at higher ρ values ($\rho > 0.8$ for $m = 3$).

3. Conclusions

The available queueing models, giving analytical results are generally too simple to fit problems, encountered in practical situations. For example, frequently analytical results or approximations can be obtained for models with only a minor deviation from the basic A/B/m systems: e.g. a M/M/1 with interruptions or with batch input; a M/M/1 system with absolute or time dependent priority for samples with short analysis times.
Serious problems arise when these systems are imbedded in a network of queues, where a part of the output of a node is the input of another one. Systems with batch input and interrupted analyses, where different priorities are attributed to various classes of samples, are considerably complex. Often, the solution of such models requires a high level of mathematics, having little sense to the practician. Mentioned solutions will often be given in terms of transforms, excluding a practical application.

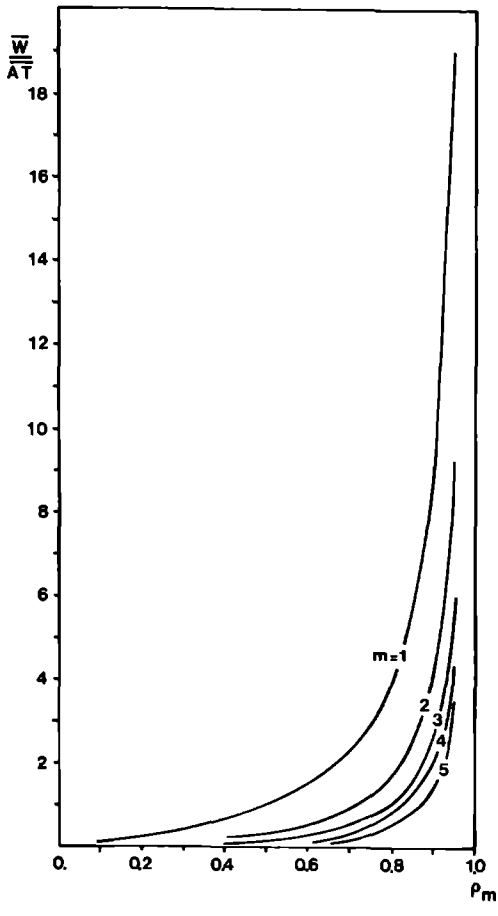


Fig. III-20: The mean waiting time ($\bar{W}/\bar{AT}$) of a multi server system (M/M/m) as a function of the utilization factor (ρ_m) of the analysts for various m-values.

As H.J. Steudel [ST76] states: "In many ways the subject of queueing appears to have gotted bagged down in a quagmire of intractable mathematics". However, in this section we did not aim to furnish analytical results for the laboratory under investigation, but to show the relevant parameters, with their influence, and eventually to formulate some generally valid statements. From the study of simple queueing systems completed with the observations of the real laboratory, one concludes:

- the utilization factor (ρ) is a dominating factor, determining the waiting time

- a reduction of the analysis time is more important for the reduction of the waiting time than a reduction of the variation coefficient of the analysis time;
- interruptions of the analysis, while samples are waiting cause the system with a low utilization factor to behave as a saturated system (large T_x, and asymptotical dependency of T_x on ρ).
  Allowing the start of other activities only during the idle period has practically no influence on system performance;
- modification of a M/M/1 system to a system with batch input improves the system performance only in these situations where the batches enter the system equidistantly and the distribution of the batch size is Gaussian or constant (supposing no change of the overhead);
- system performance improves, when attributing absolute priority to the samples with the shortest analysis time. The mean waiting time of all samples is not influenced by attributing different priority to groups of samples with the same mean analysis time. Optimization of such systems is only achieved by including cost functions;
- the application of Round Robin scheduling in analytical laboratories is not feasible;
- the cumulative density function of the waiting time has an exponential shape for many kinds of queueing systems;
- a sampled queue has a stationary first order autoregressive behaviour.

Appendix A.

Calculation of P(T<y|k) for a FIFO, M/M/1 system.

According to Kleinrock [KL75], the Laplace transform of the probability that the sample's total delay in the system is equal to y, when it finds k samples in system ahead of it, equals:

$$S^*(s|k) = [\,1/\overline{AT}/(1/\overline{AT}+s)]^{k+1}$$

The inversion of this equation gives:

$$P(T=y|k) = (1/\overline{AT})^{k+1}\frac{y^k}{k!}\exp(-y/\overline{AT})$$

Integrating that equation, we have:

$$P(T\leq y|k) = \int_0^y 1/\overline{AT}^{k+1}\frac{y^k}{k!}\exp(-y/\overline{AT})dy = [-(1/\overline{AT})^{k+1}\exp(-y/\overline{AT})\sum_{r=0}^{k} y^{k-r}\overline{AT}/(k-r)!\,]_0^y$$

$$= -(1/\overline{AT})^{k+1}\exp(-y/\overline{AT})\sum_{r=0}^{k} y^{k-r}\overline{AT}^{r+1}/(k-r)! + 1$$

$$= \exp(-y/\overline{AT})\sum_{r=0}^{k} y^{k-r}\overline{AT}^{r-k}/(k-r)! + 1$$

$$P(T\leq y|k) = 1 - \exp(-y/\overline{AT})\sum_{r=0}^{k}(y/\overline{AT})^{k-r}/(k-r)!$$

Appendix B

Calculation of the first and second moments of a truncated exponential distribution.

1. first moment:

Let $1/\overline{AT}$ be equal to μ. the p.d.f. of the truncated analysis time equals:

$b(\tau) = \mu\exp(-\mu t)$ for $t<x$

$b(x) = 1-B(x)$ with $B(x)=\int_0^x \mu\exp(-\mu t)dt$ $t=x$

From the definition of the first moment we have:

$$\overline{AT}_x = \mu\int_0^x t\exp(-\mu t)dt + x[\,1-B(x)]$$

$$= \mu\int_0^x t\exp(-\mu t)dt + x[\,1-\int_0^x \mu\exp(-\mu t)dt]$$

$$= \mu[\exp(-\mu t).(-\mu t-1)/\mu^2]_0^x + x + x[\exp(-\mu t)]_0^x$$

$$= (1-\exp(-\mu x))/\mu = \overline{AT}(1-\exp(-x/\overline{AT}))$$

2. second moment

$$\overline{AT^2_x} \triangleq \int_0^x t^2 b(t)dt + x^2(1-B(x))$$

$$= \int_0^x t^2 \mu \exp(-\mu t)dt + x^2[1-\int_0^x \mu \exp(-\mu t)dt]$$

$$= \mu[\exp(-\mu t)(-\frac{t^2}{\mu} + \frac{2t}{\mu^2} - \frac{2}{\mu^3})]_0^x + x^2 + x^2[\exp(-\mu t)]_0^x$$

$$= \frac{2}{\mu}[\exp(-\mu x)(x-1/\mu) + 1/\mu] = 2\overline{AT}[\exp(-x/\overline{AT})(x-\overline{AT}) + \overline{AT}]$$

The coefficient of variation equals: $\sigma^2_{AT_x}/(\overline{AT_x})^2 = \frac{\overline{AT^2}}{(\overline{AT_x})^2} - 1 =$

$$\frac{2[\exp(-x/\overline{AT})(x-\overline{AT})+\overline{AT}]}{\overline{AT}(1-\exp(-x/\overline{AT}))^2} - 1$$

Appendix C.

Calculation of the delay of batch input systems.

According to Burke[BU75] the average delay of a sample equals the sum of the delay of the first member of the batch and the delay due to the analysis times of the members of his batch analyzed before him.

$$\bar{T} = \bar{T}_1 + \frac{\overline{AT}}{2}\left[\frac{E(r^2)}{E(r)} - 1\right] = \bar{W}_1 + \frac{\overline{AT}}{2}\left[\frac{E(r^2)}{E(r)} + 1\right] \qquad \text{III-c1}$$

where $\bar{T}_1$ is the delay of the first member of the batch, $\overline{AT}$ is the mean analysis time pro sample and $E(r^2)$ is the second moment of the p.d.f. of the batch size. The second term in Eqn. III-c1 gives the average delay due to the analysis times of the members of the batch analyzed before the sample. Table III-c1 gives the expressions for the two moments of the various considered p.d.f. of the batch size (r).

Table III-c1

The two first moments for several p.d.f. of the batch size (r) and mean delay.

p.d.f.	E(r)	$E(r^2)$	$\bar{T}$
constant	$\bar{r}$	$\bar{r}$	$\bar{W}_1+\overline{AT}(\bar{r}+1)/2$
exponential	$\bar{r}$	$2(\bar{r})^2$	$\bar{W}_1+\overline{AT}(2\bar{r}+1)/2$
Poisson	$\bar{r}$	$\bar{r}+\bar{r}^2$	$\bar{W}_1+\overline{AT}(\bar{r}+2)/2$
Gaussian	$\bar{r}$	$\sigma_r^2+(\bar{r})^2$	$\bar{W}_1+\overline{AT}[\sigma_r^2/\bar{r}+\bar{r}+1]$

When the interarrival times of the batches are exponentially distributed, $\bar{W}_1$ can be calculated using the Pollaczek-Khincnir formula [KL75]

$\bar{W}_1 = \frac{\overline{AT_b}}{2}(1+C^2_{AT_b})\rho/(1-\rho)$ with $\overline{AT_b}$ the mean analysis time of a batch and

$C^2_{AT_b} = var(AT_b)/(\overline{AT_b})^2 = var(AT)[\bar{r}+var(r)]/[\bar{r}^2(\overline{AT})^2]$

For exponentially distributed analysis times $var(AT)=(\overline{AT})^2$

Thus $C^2_{AT_b} = 1/\bar{r} + var(r)/(\bar{r})^2$

Because $\overline{AT_b} = \bar{r}\overline{AT}$ we find that $\bar{W}_1 = \rho\bar{r}.\overline{AT}(1+1/\bar{r}+var(r)/(\bar{r})^2)/[2(1-\rho)]$

When the interarrival times of the batches are equidistant $\bar{W}_1$ can be calculated using the heavy-traffic approximation of Kingman[KI62]

$\bar{W}_1 = \frac{\sigma^2_{IAT_b} + \sigma^2_{AT_b}}{2\overline{IAT_b}(1-\rho)}$ where $\sigma^2_{IAT_b}$ is the variance of the interarrival times of the batches and $\sigma^2_{AT_b}$ is the variance of the analysis times of the batches.

For the considered batch system $\sigma^2_{IAT_b}=0$, $\overline{IAT_b}=\bar{r}.\overline{IAT}$ and $\sigma^2_{AT_b}=var(AT)(r+var(r))$ $=(\overline{AT})^2(r+var(r))$

Thus $\bar{W}_1 = \frac{(\overline{AT})^2(\bar{r}+\sigma^2_r)}{2\bar{r}.\overline{IAT}(1-\rho)} = \frac{\overline{AT}(\bar{r}+\sigma^2_r)}{2\bar{r}(1-\rho)}$

The ratio between the delay of a M/M/1 system modified to a batch input system ($\bar{T}_b$) and the original M/M/1 system can now easily been calculated.

1. the interarrival time of the batches is exponentially distributed.

$$\bar{T}_b/\bar{T} = \frac{\frac{\rho\bar{r}.\overline{AT}(1+1/\bar{r}+var(r)/(\bar{r})^2}{2(1-\rho)} + \frac{\overline{AT}}{2}\left[\frac{E(r^2)}{E(r)}+1\right]}{\overline{AT} + \overline{AT}\rho/(1-\rho)}$$

$$= \rho\frac{\bar{r}}{2}(1+1/\bar{r}+var(r)/(\bar{r})^2) + \frac{(1-\rho)}{2}\left[\frac{E(r^2)}{E(r)}+1\right]$$

2. the batches arrive at equidistant times.

$$\bar{T}_b/\bar{T} = \left[\frac{\overline{AT}\rho(\bar{r}+\sigma^2_r)}{2\bar{r}(1-\rho)} + \frac{\overline{AT}}{2}\left(\frac{E(r^2)}{E(r)}+1\right)\right](1-\rho)/\overline{AT}$$

$$\bar{T}_b/\bar{T} = \frac{\rho}{2}(1+\sigma^2_r/\bar{r}) + \frac{(1-\rho)}{2}\left(\frac{E(r^2)}{E(r)}+1\right)$$

Calculation of the effect of subdividing the samples into two categories: small and large analysis times.

1. M/M/1 system

Let be:

	large analysis times	small analysis times	total
input density	λ_1	λ_2	λ
mean analysis time	$\overline{AT}_1$	$\overline{AT}_2$	$\overline{AT}$
mean delay	$\bar{T}_1$	$\bar{T}_2$	T

Then:

$$\lambda_2 = \frac{\lambda \int_0^x \mu \exp(-\mu t)dt}{\int_0^\infty \mu \exp(-\mu t)dt} = \lambda(1-\exp(-x/\overline{AT})$$

$$\lambda_1 = \lambda - \lambda(1-\exp(-x/\overline{AT})) = \lambda\exp(-x/\overline{AT})$$

$$\overline{AT}_2 = \frac{\mu \int_0^x t\exp(-\mu t)dt}{\mu \int_0^x \exp(-\mu t)dt} = \frac{1/\mu - (\frac{1+\mu x}{\mu})\exp(-\mu x)}{1 - \exp(-x\mu)} = \frac{\overline{AT} - (x+\overline{AT})\exp(-x/\overline{AT})}{(1 - \exp(-x/\overline{AT}))}$$

$$\overline{AT}_1 = \frac{\mu \int_x^\infty t\exp(-\mu t)dt}{\mu \int_x^\infty \exp(-\mu t)dt} = \frac{\exp(-x/\overline{AT})(\overline{AT}+x)}{\exp(-x/\overline{AT})} = \overline{AT} + x$$

$$\overline{AT^2}_2 = \frac{\mu \int_0^x t^2\exp(-\mu t)dt}{\mu \int_0^x t\exp(-\mu t)dt} = \frac{(-x^2+2x.\overline{AT}-2(\overline{AT})^2)\exp(-\mu x) + 2(\overline{AT})^2}{1 - \exp(-x/\overline{AT})}$$

$$\overline{AT^2}_1 = \frac{\mu \int_x^\infty t^2\exp(-\mu t)dt}{\mu \int_x^\infty t\exp(-\mu t)dt} = (x-\overline{AT})^2 + (\overline{AT})^2$$

$$\bar{T}_1 = \frac{\overline{AT}_1(1-\lambda_1\overline{AT}_1-\lambda_2\overline{AT}_2) + \lambda_1\overline{AT_1^2}/2 + \lambda_2\overline{AT_2^2}/2}{(1-\lambda_2\overline{AT}_2)(1-\lambda_1\overline{AT}_1-\lambda_2\overline{AT}_2)}$$

$$\bar{T}_2 = \frac{\overline{AT}_2(1-\lambda_2\overline{AT}_2) + \lambda_2\overline{AT_2^2}/2}{(1-\lambda_2\overline{AT}_2}$$

$$\bar{T} = \frac{\lambda_2(\bar{T}_2-\bar{T}_1) + \lambda\bar{T}_1}{\lambda}$$

2. $M/E_4/1$ system

In all integrals of foregoing paragraph, the term $\mu\exp(-\mu t)$ should be replaced by $\frac{(4\mu)^4}{3}t^3\exp(-4\mu t)$

Thereafter all terms ($\overline{AT}_2,\ldots\bar{T}$) can be calculated straight forward.

AD75 V. Adeberg and K. Doerffel, Communication presented at the Euroanalysis Congress, Budapest, 1975.

BA75 F. Baskett, K.M. Chandy, R.R. Muntz and F.G. Palacios, J. Assoc. Comput. Mach., 22 (1975), 248.

BH72 U.N. Bhat, Opns. Res., 20 (1972), 357.

BU75 P.J. Burke, Opns. Res.,23 (1975),830.

CO67 R.W. Conway, W.L. Maxwell, L.W. Miller, Theory of Scheduling, Addison Wesley, Mass., 1967

DU71 L. Durr, Opns. Res., 19 (1971) 453.

FI77 M.J. Fisher, Management Sci., 24 (1977), 338.

HO70 J.M. Holtzman, Opns. Res., 19 (1971), 461.

JA57 J.R. Jackson, Opns. Res., 5 (1957), 518.

KI62 J.F.C. Kingman, J. Royal Stat. Soc., B24 (1962), 383.

KL67 L. Kleinrock, Opns. Res., 15 (1967), 304.

KL75 L. Kleinrock, Queueing Systems, Vol I, Theory, Wiley & Sons, New York, 1975.

KL76 L. Kleinrock, Queueing Systems, Vol II, Computer Applications, Wiley & Sons, New York, 1976.

KF76 L. Kleinrock and R.P. Finkelstein, Opns. Res., 15 (1967), 104.

LE75 J. Levy and K. Yechiali, Management Sci., 22 (1975), 202.

LI61 J.D.C. Little, Opns Res., 9 (1961), 383.

MA72 E. Maaløe, Opns. Res., 20 (1972), 703.

MF72 A.G. Merten, M.L. Muller, Management Sci., 18 (1972), 518.

MO55 P.H. Morse, Opns Res., 3 (1955), 255.

RE68 J.F. Reynolds, J. Appl. Prob., 5 (1968), 467.

ST76 H.J. Steudel, S.M. Pandit, S.M. Wu, Management Sci., 24 (1977), 456

CHAPTER 4

THE SIMULATION MODEL

Simulation is defined [HU70] as a numerical technique for conducting experiments with certain types of mathematical models, describing the behaviour of a complex system in a digital computer over extended periods of time.

The starting point of any computer experiment is a model of the system to be simulated, which is characterized by (1) a structure, (2) many parameters and variables (deterministic and stochastic) (3) a response (or responses). This section is devoted to the description of the simulation model.

1. Structure of the model

1.1 Fixed characteristics

The flow chart in Fig. IV-1 may serve to clarify the structure of our model. It should be noted that this flow chart gives a crude picture of the decision processes inside the laboratory. The laboratory consists of 4 sections, each one having an input flow of samples, originating from outside the system and from the departments within the system. The samples originate from two sources (F(1) and F(2)). For each source a stochastic variate $IAT_{i,j}$ is defined, which is the time interval between the arrival in the laboratory

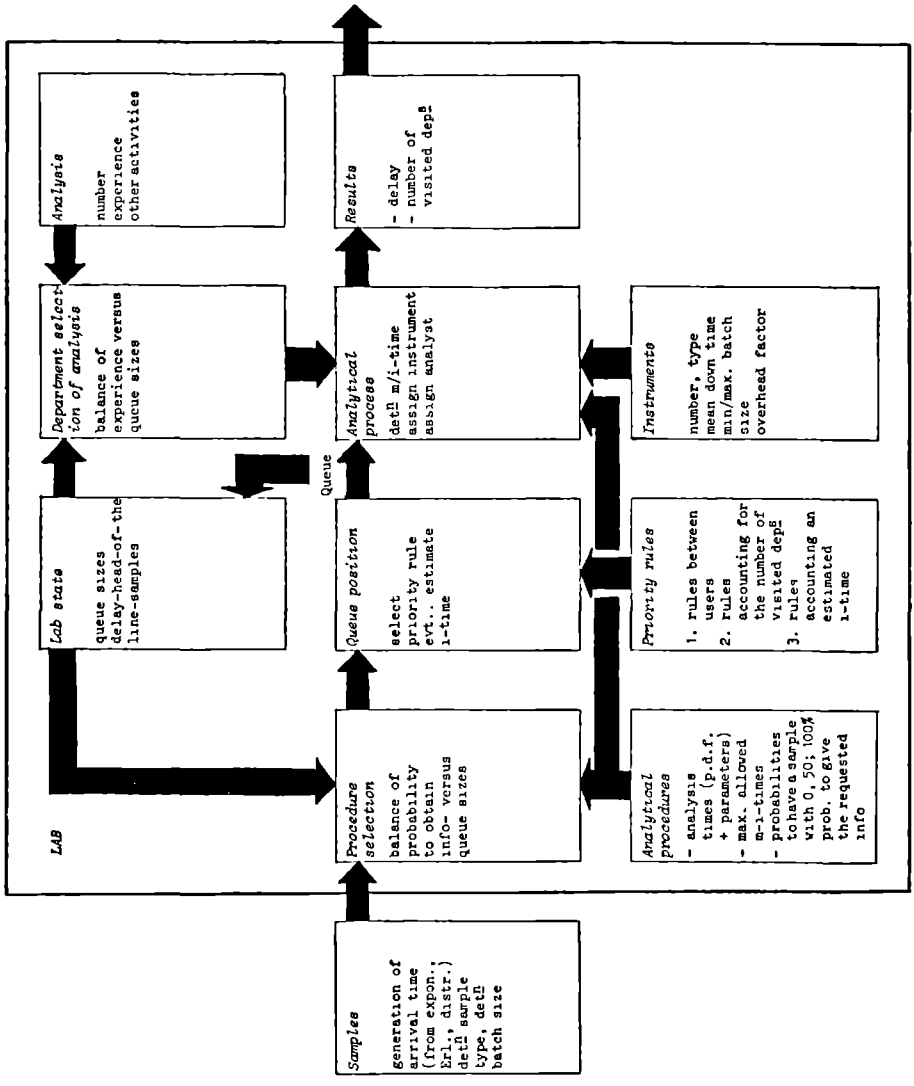


Fig. IV-1: Structure of the model

of the i-th sample and the (i-1)th sample from source j, with a known probability density function $f(IAT_j)$, expected value $E[IAT_j]$ and variance $var[IAT_j]$.

Exponential, Erlangian (of order k), and hyperexponential probability functions can be selected. The distribution of the newly arrived samples at the laboratory over the four sections is realized using various decision rules, described in paragraph (1.2). The arrival processes in the model do not depend on the state of the system. The analysis process has three stages: the measurement of the spectrum, the interpretation of the spectrum and the communication of the analytical result. In the model the interpretation and the measurement times of the samples are generated by taking random numbers from exponential or Erlangian density functions.

As many operating characteristics in the described model are given by probability functions, the model is called stochastic. In practice, a sequence of random numbers is required to generate a sequence of e.g. analysis times or interarrival times, having a given density function describing the actual statistic property of these variables.

The analysis rate is state independent: a sample's analysis time at a section is not permitted to depend upon its analysis time at previously visited sections

A batch input and/or batch analysis of the samples can be generated. The batch sizes and interarrival times of these batches can be taken from various probability density functions (Gauss, Poisson, constant). For each section, a minimal and maximal batch size for the analysis can be selected. When the minimal and maximal batch sizes are different, then the analyst waits until the minimal batch size is present before starting the measurement of samples. For a minimal batch size equal to one, the analyst starts the measurements when there are any samples waiting. He starts the interpretation of a spectrum after the measurement of a complete batch. In the model, results are only communicated to the user, when the pile of results has reached a given value, or when results wait longer than a preset time before communication. However, when the state of the laboratory is such that the analyst remains idle, while results are still waiting, then, results are communicated as well. In the model an instantaneous transfer of samples, spectra or results between the departements in the laboratory is assumed.

The analysis can be interrupted for other activities, coffee breaks, holidays, and machine breakdowns. For each of these four types of interruptions a mean interruption time and mean interval between the interruptions can be selected

from an exponential probability density function. The simulated laboratory has a maximum capacity of 20 analysts and 10 instruments. Each analyst has a given experience with the four procedures, denoted by a parameter between one (fully qualified) and zero (no experience at all). This factor in combination with a minimal required experience for a particular analytical procedure, ranges the organization of the laboratory from an open organization where all analysts can do all analysis, to a closed organization, where the analysts specialize in one analytical method only. Furthermore, there is a functional relationship in the model between the analysis time and the experience parameter (Exp) of the analyst (j) who executes the analysis (i).

$$(AT_i)_j = (AT_i)/_{Exp(j)}$$

In the model a dynamic priority rule is applied as described in Chapter III. The F(1) and F(2) samples are positioned in the queue according to the value of the sum of arrival date and product of priority factor (p) and urgency number (A). The priority difference between samples of both sources is not necessarily the same in each section. The laboratory is empty at the start of each simulation run. The simulation period for each run is 4000 completed samples. This corresponds to about 1 year operation of the laboratory.

In the model, measurement times and interpretation times were introduced which are higher than measured in the laboratory.

This augmentation accounts for tranfer times and administration times of each sample (or spectrum) which were not included in the data presented in Table II-6

Table IV-1 shows the statistical parameters which were used in the model.

Table IV-1

Statistical parameters of the measurement time (MT) and interpretation time (IT) in the model

section	measurement time (hrs)			interpretation time (hrs)		
	mean	$S_{\overline{MT}}$	$C^2_{MT} = \frac{S^2_{MT}}{\overline{MT}^2}$	mean	$S_{\overline{IT}}$	$C^2_{IT} = \frac{S^2_{IT}}{\overline{IT}^2}$
I.r.	0.4	0.01	1.0	1.2	0.05	1.5
P.m.r.	0.7	0.006	0.2	1.2	0.02	1.0
M.s.	0.9	0.03	1.0	1.5	0.07	1.5
^{13}C-n.m.r.	1.0	0.02	0.5	1.6	0.08	2.0

Probability density functions with a variation coefficient (C_b^2) smaller than one were generated by taking a r-stage Erlang distribution. Because for a r-stage Erlangian distribution C_b^2 equals $\frac{1}{r}$, only the values C_b^2 = 1, 0.5; 0.3; 0.25; 0.2 etc. can be selected. Values of $C_b^2 > 1$ are obtained by generating a hypergeometric distribution, which is a combination of two exponential functions: i.e. exponentially distributed variates are taken with the probabilities p and 1-p from distributions with the parameters $2p\alpha$ and $(1-p)2\alpha$ respectively. This process generates hypergeometric variates with mean $^1/\alpha$ and a density function f(x):

$$f(x) = 2p^2\alpha \exp(-2p\alpha x) + 2(1-p)^2\alpha \exp[2(1-p)\alpha x] \quad [NA66]$$

with a variance of x equal to $\frac{1}{\alpha}[\frac{1}{2p(1-p)}] - 1$
If the desired value of C_b^2 is known for a given value of $^1/\alpha$, p can be calculated from: $p = 0.5 - 0.5\ (1-2/(C_b^2+1))^{\frac{1}{2}}$ [NA66]

1.2 Variable operating characteristics

In the model strategies can be selected concerning the sample priorities, the route of the sample through the laboratory, the assignment of the analysts and the termination of the analysis.
The strategies considered in the model are as follows:

a. Strategies concerning sample priorities:

(1). the sample in queue with the earliest-laboratory-arrival date is selected first for analysis (ELAD)

(2). that sample in queue with the earliest arrival date at the analytical section is analyzed first (EAAD)

(3). the samples in each section are subdivided into groups according to the number of analyses unseccessfully done before. Priority (varying from FiFo to absolute priority) is assigned either to samples which have visited the largest number of sections, or to samples which have visited the smallest number of sections.

(4). samples receive priority according to the analysis time expected. There are two situations:
- the shortest-expected-analyzing-time-first (SEAT) discipline.
- the samples are grouped in two categories. All samples with an expected analysis time smaller than some defined value, have priority over the others. The discipline within a group is FiFo. The accuracy

of the 'a priori' estimation of the interpretation time of a sample can be varied in the model. If the accuracy is low, and the SEAT discipline is applied, then the analysis is accomplished in random sequence.

b. Strategies concerning the routing process:

(1). the analytical method with the highest estimated probability of a successful elucidation of the requested structure is selected. The choice between methods with the same probability is made randomly. This is a fixed routing procedure in which a sample path is uniquely determined from the properties of the sample itself. This is assumed to be the policy of the real laboratory, which is taken as a base for the comparison of alternative strategies.

(2). Not only the probability of success but the expected waiting times for the various sections are considered at the time that the sample arrives at the laboratory. The relative importance of both criteria is given by weighting factors. Because this routing algorithm bases its decisions on the observed traffic flow, it is called a dynamic or adaptive routing procedure.

c. Strategies for assignment of the analyst:

The same decision rule for assigning analysts to a section is used for all analysts. There are two main possibilities:

(1). complete centralization: when an analyst completes an analysis, he is available for reassignment to another analytical method. This assignment can be governed by the following work selection rules:
- assign analysts to the method with which they have most experience and which has an unmanned instrument
- assign the available analyst to that section with the oldest sample in its queue and for which the analyst has sufficient experience
- the experience of the analyst and the waiting time of the oldest sample in each queue are weighted.

(2). complete decentralization: the analyst is always assigned to the same analytical method, irrespective of the state of the laboratory, (that is the policy in the real laboratory).

d. Strategies concerning the termination of the analysis:

(1). there are no restrictions on the analysis time (existing policy)

(2). a maximal measuring and interpretation time is assigned to each analytical section, regardless of the originally estimated probability of success.

(3). the maximal measuring and interpretation time is a function of the

probability of success and the number of unsuccessfully applied methods.

(4). the maximal measuring and interpretation time also depends on the state of the laboratory, i.e. the number of waiting samples in the analytical section.

In each section the same strategy concerning the termination of the analysis is employed.

A typical property of the considered spectroscopic laboratory is that some samples (20%) are sequentially analyzed in several sections. All arriving samples are accompanied with an application form with data about their origin and requested information. Sometimes the applicant of the analysis indicates the desired analytical method by himself. For 80% of all samples sufficient information is obtained by one method only. When the analysis failed, the problem along with the intermediate results is passed to a next method. The various sections operate relatively independent from each other, as only few samples (10%) fail after a combination of two methods. As a result the sections in the model are designed as independent nodes in the network. The mean measurement and interpretation times in the model account for the transfer times of samples between the sections. Because the data base of the registered measurement- and interpretation times was too small, no functional relationships could be determined between those times and the number of unsuccessful methods tried before.

2. Generation of the flow through the laboratory

As a dynamic or adaptive routing procedure will be used, based on as well the properties of the sample as the state of the laboratory (the number of samples in each section), a routing algorithm had to be developed, that based its decisions on the observed traffic flow and probabilities that the underlying analytical problem can be solved by the various analytical methods. It was assumed that these probabilities were independent from the source (F(1) and F(2)) of the samples. In the actual situation a minor difference exists, but for reasons of simplicity, the number of parameters in the model was maintained as small as possible.

The generation of sample flows, by taking random numbers from exponential or Erlangian probability distributions, was only executed, for the sample traffic from the outside to the inside of the laboratory, and not for the traffic between the departments. As a result, the probability density functions

of the arrivals at the individual departments are defined by three factors:

a) the probability functions of the arrivals of F(1) and F(2) samples at the laboratory

b) the distribution process of these samples over the four departments

c) the departure processes of samples which are not successfully analyzed.

For each sample arriving at the laboratory, the probability of a successful analysis of the molecular structure of the sample is estimated on three levels:

j = 0: the analytical procedure is estimated to be incapable to furnish the requested structure

j = 0.5: the analytical procedure is estimated to give the structure with a probability of 0.5

j = 1: the estimated probability that an analytical method will furnish the structure = 1.

The fractions of samples, having j = 0; 0.5; and 1, denoted by p(i,0); p(i, 0.5) and p(i,1) are determined for each section (i), from the observed sample flow in the laboratory (Appendix IV-A), and are shown in Table IV-2.

The flow to the sections can be simulated, assuming that all samples, for which the estimated probability that the analytical method will furnish the requested information, are indeed successfully analyzed by that method. Another possibility is that only a given fraction of these samples are successfully analyzed: i.e. the probability of elucidation of the structure by a given method can be estimated less accurate. As a result, the effect of balancing this probability against the queue lengths in the model can be determined as a function of the accuracy of the estimated probability of success. The uncertainty that could be introduced in the 'a priori' forecast of the probability of success of a given method could be enhanced to a maximum of 16%: i.e. 16% of all samples, are unsuccessfully analyzed in a method, estimated before to give the requested information.

Table IV-2

The probability p(i,j) to find for a sample that the analytical procedure (i) will give the requested analytical result with probability (j)

section \ probability (j)	A*)			B*)		
	0	0.5	1	0	0.5	1
I.r.	0.62	0.26	0.12	0.07	0.78	0.15
P.m.r.	0.24	0.23	0.53	0.24	0	0.76
M.s.	0.87	0.02	0.11	0.46	0.41	0.13
^{13}C-n.m.r.	0.80	0	0.20	0.72	0	0.28

A*) assumption that all samples with j=1, directed to department (i) are successfully analyzed.
assumption that 50% of all samples with j=0.5, directed to department (i) are successfully analyzed.
assumption that no samples with j=0, directed to department (i) are successfully analyzed .

B*) assumption that only 84%** of all samples with j=1, directed to department (i) are successfully analyzed.
assumption that only 16%** of all samples with j=0, directed to department (i) are successfully analyzed.

**Upper and lower bound for this assumption, to reproduce the flow in the laboratory.

3. Simulation of the model

3.1 Time flow mechanisms

Two general types of methods have emerged for moving a model of a system through time on a computer: a fixed time, and variable time increment method [NA66].

With fixed time increment methods a clock is simulated by the computer, which is updated in uniform discrete intervals of time. Every unit of clock time, the system is scanned to determine whether any event occurred during that time. Underlying simulation model used the variable time increment method. This means that, when a particular event occurred in the laboratory, the clock time is advanced to the time at which the next event is to occur. The intervening time periods where no changes occur in the system are skipped. At the occurrence of each event, a number of activities must be executed by the model, listed in the event description. In the model five different kinds of events can take place: 1. a sample enters the laboratory; 2. the measurement or interpretation of a sample is completed; 3. an analyst finishes other activities; 4. an analyst returns to the laboratory after the communication of the results; 5. the down time of an instrument is over. It was not necessary to include two additional events marking the moment that analysts start other activities and instruments go down. These moments are calculated during the occurrence of event 3 and 5. The event description, associated with the five events accounts for the availability of the facilities. The event description generates the next events which should take place. In this way the model progresses in time automatically.

3.2 The simulation language

The task of writing simulation programs is simplified by the development of 'simulation languages'. Among the simulation languages that have been developed are: GPSSII [GP], SIMSCRIPT [MA62], GASP [KI63], SIMPAC [SI62], DYNAMO [PU63] and SIMULATE [HO64]. GPSSII and GASP are best suited to certain types of scheduling and waiting time problems. Because GASP is the only language which is written in Fortran IV, and can be recompiled with a Fortran IV compiler, we have written the simulation program in that language, consisting of several subroutine programs and function subprograms.
The used GASP version was described by Kerbosch [KE73].

3.3 Generation of random variates

Random variates drawn from a given probability distribution are generated by means of uniformly distributed random numbers (between 0 and 1) which were obtained from an IBM pseudo random number generator.
Among others, random variates x_i from some particular statistical population with a probability density function (f(x)) are generated by calculating the cumulative distribution F(x). Since F(x) is defined over the range 0-1, the value of x (say x_o) can be calculated from uniformly distributed random numbers (r), for which $F(x_o) = r_o$ (Fig. IV-2). From Fig. IV-2 it is easily seen that the probability of finding a value $x < x_o$ is equal to the probability of finding a value $r < r_o = F(x_o)$ Thus:

$$P(x < x_o) = P(r < F(x_o)) = P(F^{-1}(r) < F^{-1}F(x_o)) = P(F^{-}(r) < x_o)$$

where F^{-1} is the inverse function of F.
By this method, exponentially distributed variates were generated in the model. K-order Erlangian distributed variates were generated by adding k exponentially distributed numbers.
For the generation of Gaussian distributed numbers, a method based on the central limit theorema was used. Adding 12 independent, uniformly distributed random numbers and substracting 6, gives Gaussian distributed variates (x_i) with a mean zero and standard deviation equal to one. A Gaussian distribution with a mean $\bar{x}$ and σ_x is simply obtained by applying the algorithm $x_i * \sigma_x + \bar{x}$ on each variate (x_i).

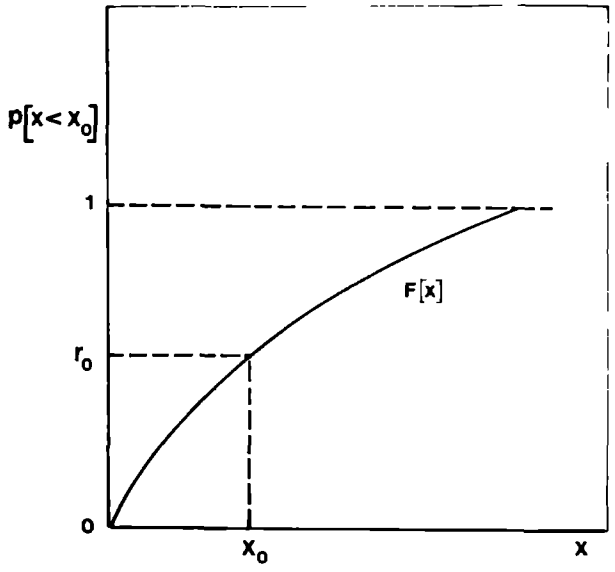


Fig. IV-2: The generation of random variates (x) with a cumulative density function F(x) from uniformly distributed random numbers (r)

3.4 Validation

The validation of a simulation model requires a comparison of actual versus simulated data. It is common practice to compare histograms of both data series, employing the standard χ^2, t, and F statistics.

However, in most cases the actual data as well as data from many simulations are serially correlated. This greatly complicates the application of above statistics. Moreover, serial correlation in time itself is often an important characteristic of the simulated system. Hsu and Hunter [HS77] suggested the comparison of historical and simulated data by identifying a time series model and estimating the parameters of this model by the techniques outlined by Box and Jenking [BO70]. Hereafter, the models are tested for differences in their means, autoregressive parameters and residual variances. This approach based on time series can very well be used for modelling the channel utilization [HS77] and number of samples in the laboratory [HS77, ST77].

However, application of a time series approach is difficult when delay times of the samples are involved because the sequences of departures and arrivals of samples are unequal. In general the delay times of samples in a sample record are correlated. An alternative method for the estimation of some

parameters of correlated observations is the batch means method which is described by several workers [FI78,C063] .

3.4.1. The batch means method.

The basic idea behind the batch means method is to combine the sample sequence of n observations into k batches of m observations, and to compute a sample mean ($y_{i,m}$) of each batch (i). With these means an estimate of the variance of the grand sample mean over all batches is calculated (Fig. IV-3).

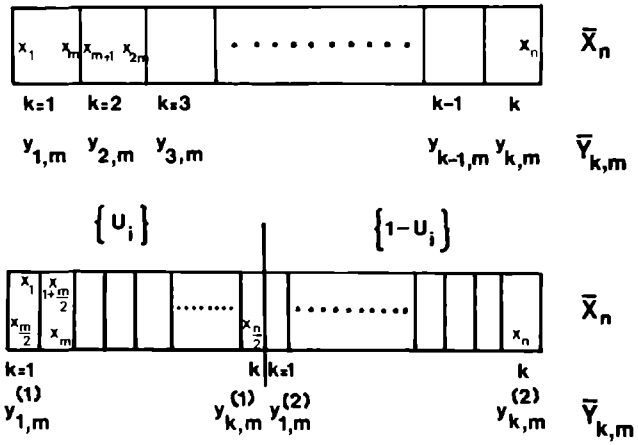


Fig. IV-3: Variance reduction by using antithetic variates.

The sample mean of a sample record of n observations gives an estimate ($\bar{x}_n$)of the mean μ of the population.

$$\bar{x}_n = 1/n \sum_{i=1}^{n} x_i \qquad \text{IV-1}$$

Together with the estimate var($\bar{x}_n$) a confidence interval of $\bar{x}_n$ can be obtained [WA75,M067,MU78] .

$$\mathrm{var}(\bar{x}_n) = \sigma_x^2/n \left[1 + 2\sum_{k=1}^{n} (1 - k/n)\phi(\tau) \right] \qquad \text{IV-2}$$

Of course, when the observations are independent, $\phi(\tau)=0$, and var($\bar{x}_n$) = σ_x^2/n with

$$\sigma_x^2 = 1/(n-1)\sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \qquad \text{IV-3}$$

Considering the batch means method, the means of the batches are equal to

$$y_{i,m} = 1/m \sum_{i=1}^{m} x_{m(j-1)+i} \qquad j = 1,\ldots k \qquad \text{IV-4}$$

where $k = n/m$

Averaging these batch means, another estimate ($\bar{y}_{k,m}$) of the mean μ is obtained.

$$\bar{y}_{k,m} = 1/k\sum_{j=1}^{k} y_{j,m} \qquad \text{IV-5}$$

Clearly, $\bar{y}_{k,m} = \bar{x}_n$.

When the autocorrelation ϕ_{xx} of x is monotonely decreasing, the batch means are not correlated, provided the batch size (m) is sufficiently large.
The variance of $\bar{y}_{k,m}$ can be estimated from

$$\text{var}(\bar{x}_n) = \text{var}(\bar{y}_{k,m}) = 1/k(k-1)\sum_{j=1}^{k} (y_{j,m} - \bar{x}_n)^2 \qquad \text{IV-6}$$

It remains to test the hypothesis that there is no correlation between neighbouring batch means. To this purpose, the method described by Fishman [FI73] was followed, using the statistic

$$C_k = 1 - \sum_{i=1}^{k-1} (y_{i,m} - y_{i+1,m})^2 / 2\sum_{i=1}^{k} (y_{i,m} - \bar{x}_n)^2 \qquad \text{IV-7}$$

For $k > 8$, the distribution of $y_{1,m}\ldots y_{k,m}$ is close to normal [FI78], and under H_0 that there is no correlation, C_k has a mean zero and variance $(k-2)/(k^2-1)$. For $C_k < u(P)\sqrt{(k-2)/(k^2-1)}$ the H_0 hypothesis is accepted, where $u(P)$ is the excentricity of a normal distribution with an accuracy of P%.
When the H_0 hypothesis is rejected two procedures can be followed: first, larger batches can be taken (increase of m). However, when the number of batches becomes less than 8, one should increase the number of observations n. Because the ultimate goal of simulation experiments is to compare some resporse parameters for different operating policies, a minimal confidence interval of the estimate ($\bar{x}_n$) of μ is desired. Therefore, the batch size (m) should be as small as possible, in order to have the maximal number of degrees of freedom (k) for the calculation of the variance. Moreover, variance reduction techniques are developed to reduce the sample size in simulation experiments (Mitchel [MI73], Naylor [NA66], Fishman [FI73]). Therefore, the simulation sequence

shown in Fig. IV-3 is modified in a sequence composed from two runs. In the first run n/2 delays are simulated, combined to k batches of m/2 samples. These simulations are executed using the sequence $\{u_i^{(k)}\}$ of independent random variables that generate the sequences of interarrival times and of analysis times. Thereafter a second simulation is run, that also combines n/2 delays in k batches. However, now the sequence $\{v_i^{(k)}\}$ of independent random variables generates the sequences of interarrival times and of analysis times. When $\{v_i^{(k)}\}=\{1 - u_i^{(k)}\}$ the two runs are called antithetic. This condition implies that the delays found in the k-th batch of the second and of the first run are negatively correlated. Mitchell [MI73] demonstrated that when the simulations are executed with two antithetic runs of n/2 delays, the standarddeviation is reduced, compared to a single run of n delays. This reduction amounts about 20% for an M/M/1 system.

3.4.2. The replication technique.

From Eqn. IV-6 the slowness of stochastic convergence appears. In order to halve the standarddeviation of a sample mean ($\sigma_{\bar{n}}$), one must quadruple the sample size. A demand for a small $\sigma_{\bar{n}}$ can easily lead to an unreasonably large sample size, associated with increased costs of computer time. Another way to diminish the variance of a performance characteristic in a simulation experiment is to include more controllable factors in the model. However this requires sometimes a rigorous change of the model. Therefore, the effects of all uncontrollable factors are absorbed in the random character of the input variables. Indeed, a variable is treated stochastically, by a lack of knowledge about the source of its variations. In computer simulation experiments one is usually interested in measuring differences in average responses for various combinations of factor levels. The variance of these differences is reduced by taking stochastic variates, generated from the same sequence of random numbers. For example in Ch. V the influence of the priority between various classes of samples is discussed by using the same 4000 samples for each run with different priority disciplines. In that case the input sequence of the samples to the laboratory is treated as a controlled variable, yielding measurements of differences between the runs having a reduced random error. This replication technique is based on a mathematical result that the standarddeviation of the difference between two sample averages $\bar{x}_1$ and $\bar{x}_2$ is reduced when $\bar{x}_1$ and $\bar{x}_2$ are positively correlated.

$$E[\bar{x}_1 - \bar{x}_2]^2 = E[\bar{x}_1^2] - 2E[\bar{x}_1.\bar{x}_2] + E[\bar{x}_2^2]$$

or: $\sigma^2_{(\bar{x}_1-\bar{x}_2)} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} - 2\sigma_{\bar{x}_1\bar{x}_2}$

$$= \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} - 2\psi_{\bar{x}_1\bar{x}_2}(0).\sigma_{\bar{x}_1}.\sigma_{\bar{x}_2}$$

Supposing that $\sigma_{\bar{x}_1} = \sigma_{\bar{x}_2}$, we find that

$$\sigma^2_{(\bar{x}_1-\bar{x}_2)} = 2\sigma^2_{\bar{x}_1} - 2\psi_{\bar{x}_1\bar{x}_2}(0).\sigma^2_{\bar{x}_1}$$

For $\psi_{\bar{x}_1\bar{x}_2} = 1$; $\sigma^2_{(\bar{x}_1-\bar{x}_2)} = 0$

The usefulness of the results obtained in this way depends on the planning horizon in the laboratory. At an infinite planning horizon the variance of a performance characteristic tends to zero ($\sigma_{\bar{x}} \rightarrow 0$), and each change of the output will be statistically significant. The optimal strategy found here, however,will also be the best one in a situation with a finite horizon. But, it becomes questionable whether a statistically significant better operation of the laboratory in reality will be observed, because in the real situation the system does not replicate. $\psi_{\bar{x}_1\bar{x}_2}(0)$ is near zero, and therefore $\sigma^2_{(\bar{x}_1-\bar{x}_2)} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}$, and it is more difficult to detect differences between $\bar{x}_1$ and $\bar{x}_2$.

3.4.3. The time series approach.

As outlined in Ch. II, an autoregressive first order model (AR(1)) of a time series is described by 2 parameters: ϕ_1 the autocorrelation at $\tau=1$, and the residual variance (σ^2_a). According to Box and Tiao [BO73] the means of two time series can be compared by applying the Student's t test for correlated time series. Therefore, first of all, the original correlated AR(1) time series (N_t) is transformed to independent normal variates (u_t) with a mean $\bar{u}$ and variance s^2_u [HS77] , by applying Eqn. IV-8

$$u_t = (N_t - \phi_1 N_{t-1})/(1 - \phi_1) \qquad \text{IV-8}$$

The variance (s^2_u) of this transformed AR(1) time series equals [HS77] (Appendix B)

$$s^2_u = s^2_N (1 + \phi_1)/(1 - \phi_1) \qquad \text{IV-9}$$

and the mean (Appendix B)

$$\bar{u} = \bar{N} \qquad \text{IV-10}$$

Thereafter the student's t test between the historical (N_t) and simulated data (N_t^*), which were respectively transformed to the independent normal variates u_t and u_t^*, with means $\bar{u}$ and $\bar{u}^*$, and standarddeviation s_u and s_u^*, was executed according to Eqn. IV-11:

$$t = (\bar{u} - \bar{u}^*)/\{s_u^2/(n-p) + s_u^{2*}/(m-p)\}^{\frac{1}{2}} \qquad \text{IV-11}$$

where n and m are the number of observations in both series, and p the number of degrees of freedom. Eqn. IV-9 can also be derived from the expression of the standarderror ($\sigma_{\bar{N}}^2$) of the estimate ($\bar{N}$) of the mean of a time series [RE70]

$$\sigma_{\bar{N}}^2 = \underset{N}{\sigma_N^2}\{1 + 2\sum_{k=1}^{N}(1 - k/N)\exp(-k/T_x)\} \qquad \text{IV-12}$$

For a first order correlated time series, the Eqn. IV-12 becomes [M067]

$$\sigma_{\bar{N}}^2 = \underset{N}{\sigma_N^2}\{1 + \frac{2\phi_1}{(1-\phi_1)}(1 - \frac{1-\phi_1^N}{N(1-\phi_1)})\} \qquad \text{IV-13}$$

for large values of n: $\phi_1^N \to 0$ and $\frac{1-\phi_1^N}{N(1-\phi_1)} \to 0$

Thus $\sigma_{\bar{N}}^2 = \underset{N}{\sigma_N^2}(1+\phi_1)/(1-\phi_1)$, which approximaltely equals $s_u^2/(n-p)$

The algorithm for u_t and s_u^2 can also be derived for higher order time series models [HS77] .

The estimated value of $\sigma_{\bar{N}}^2$ is strongly dependent on the accuracy of ϕ_1 and the exactitude of the order of the model. According to Bartlett [BA46] for a first order model:

$$\sigma^2(\phi_{xx}(1)) = (1 - \phi_{xx}^2(1))/(N-1) \qquad \text{IV-14}$$

The estimated autoregressive parameter (ϕ_1) and the residual variance (σ_a^2) of two time series can be compared simultaneously by using an inferential statistic $G(\Psi,\gamma)$, where $\Psi=\phi_1-\phi_1^*$, and $\gamma = \sigma_a^{2*}/\sigma_a^2$. Therefore two time series should be compared by establishing whether the inference $\Psi=0$, and $\gamma=1$ is

tenable or not. Hsu and Hunter [HS77] described a testing procedure for that purpose. This test is executed by the calculation of the value of $G(0,1)$. If the value of $G(0,1)$ is below a limiting value, Ψ and γ are not significantly different from respectively zero and one. On the other hand, if it is higher than this limit it should be further established whether this is due to either Ψ or γ or both. Therefore the value of $G(\Psi,1)$ and $G(0,\gamma)$ are calculated and compared with some critical point of $\frac{1}{2}\chi^2$. Details about the derivation of the joint posterior density function of γ and Ψ, denoted by $P(\Psi,\gamma|n_1,n_2)$ have been given by Hsu [HS77].

Calculation of the probability p(i,j) for a sample that the analytical procedure (i) will give the requested result with a probability (j).

The probability p(i,j) is calculated for a fixed routing procedure where a sample path is uniquely determined by the property of the sample itself. The choice between sections with the same probability for delivering the requested result is at random. Furthermore the assumption is made that j can take three values only: j=0; 0.5; and 1. Thus the probability (Pr(i)) that a sample, at its arrival in the laboratory, is routed to section(i) equals the sum of the probabilities that a sample is directed to section(i), with j=0 (Pr(i,0)); j=0.5 (Pr(i,0.5)) and j=1 (Pr(i,1)).

Thus: $Pr(i) = Pr(i,0) + Pr(i,0.5) + Pr(i,1)$ IV-A1

Assuming that indeed 50% of all $^{all}_{no}$ samples with $^{1}_{0}$ j=0.5, directed to the section i, are completed in that section, we find that the probability (Pc(i)) of completion in section i equals:

$$Pc(i) = (Pr(i,1) + 0.5Pr(i,0.5))/Pr(i) \qquad \text{IV-A2}$$

The probabilities Pr(i,0), Pr(i,0.5) and Pr(i,0) all are a function of the values of p(i,j). Having i=1,..4 and j=0, 0.5 and 1 with p(i,0)+p(i,0.5)+p(i,1)=1, Eqn. IV-A1 and Eqn. IV-A2 are two equations with two unknowns. Consideration of these equations for all sections, gives 8 equations with 8 unknowns, which is solvable. The values Pr(i) and Pc(i) for each section i= 1,..4 in the actual laboratory are tabulated in the first two rows of Table II-3 .

It remains to express Pr(i,j) as an explicit function of p(i,j).

Namely:

$$Pr(i,j)=p(i,j)\prod_{\substack{k=1\\k\neq i}}^{4}(1-p(k,j))+0.5\sum_{\substack{k=1\\k\neq i}}^{4}p(i,j)p(k,j)\prod_{\substack{l=1\\l\neq i\\l\neq k}}^{4}(1-p(l,j))+0.25\prod_{k=1}^{4}p(k,j) +$$

(A) (B) (C)

$$0.333\sum_{\substack{k=1\\ l\neq i}}^{4} p(i,j)(1-p(k,1))\prod_{\substack{l=1\\ l\neq i\\ l\neq k}}^{4} p(l,j) \qquad \text{IV-A3}$$

(D)

As an example Eqn. IV-A3 is discussed for j=1:

term (A): represents the probability to select method i, having a probability j=1 to give the requested information, while all other methods have a lower probability (j<1).

(B): is the probability to route the sample to section i, having a probability j=1 to give the requested information, while section k has the same probability j=1, and all other sections a lower one (j<1). The factor 0.5 appears because a random selection should be made between method i and k.

(C): cfr(B) but here all sections have the same probability to give the requested information. Here a random selection is made and the probability to select method i equals 0.25.

(D): cfr(B), however, here only one method has a probability j<1. A random selection should be made between three methods, having the same probability (j=1) to give the requested information.

Releasing the assumed correctness of the estimated probability (j) that a sample will be completed in a section, the probabilities p(i,j) can be recalculated for the case that :

e.g. only 90% of all samples with j=1 are completed

and 10% j=0

Then Eqn. IV-A2 becomes

$$Pr(i) = (0.9Pr(i,1) + 0.5Pr(i,0.5) + 0.1Pr(i,0))/Pr(i) \qquad \text{IV-A3}$$

The partition of the traffic of samples between the sections in the simulation model,is calculated from p(i,j) values for the partition of the arriving samples in the laboratory.

Consider a sequence of observations N_t, t=1...,n, described by a first order autoregressive model, transformed to a sequence $u_t = (N_t - \phi_1 N_{t-1})/(1 - \phi_1)$

$$\text{then } u_t = \frac{(N_t - \bar{N}) - \phi_1(N_{t-1} - \bar{N}) + \bar{N}(1 - \phi_1)}{(1 - \phi_1)}$$

Substituting $N_t - \bar{N} = n_t \quad t=1,\ldots,n$

then $u_t = (n_t - \phi_1 n_{t-1})/(1 - \phi_1) + \bar{N}$

For a first order autoregressive model $n_t = \phi_1 n_{t-1} + a_t$, where a_t (t=1,...n) represents independent and identically distributed normal random variables with mean zero and variance σ_a^2.

Thus $u_t = a_t/(1 - \phi_1) + \bar{N}$ and $\bar{u} = \bar{N}$

The variance of u_t equals:

$$s_u^2 = 1/(n-1)\sum_{t=1}^{n} (a_t/(1-\phi_1) + \bar{N} - \bar{N})^2 = [\sum_{t=1}^{n} a_t (1-\phi_1)]/[(1-\phi_1)(n-1)]$$

$$= s_a^2/(1 - \phi_1)$$

For a AR(1) model, the residual variance equals:

$$s_a^2 = s_N^2 (1 - \phi_1^2)$$

Thus: $s_u^2 = s_N^2 (1+\phi_1)/(1 - \phi_1)$

BA46 M.S. Bartlett, J. Royal Stat. Soc., B8 (1946), 27.

BO70 G.E.P. Box and G.M. Jenkins, Time series analysis. Forecasting and Control, Holden Day, San Francisco, 1970.

BO73 G.E.P. Box and G.C. Tiao, Bayesian inference in Statistical Analysis, Addison- Wesley, Publishing Co., Reading Mass., 1973.

CO63 R.W. Conway, Management Sci., 10 (1963), 47.

FI73 G.S. Fishman, Management Sci., 20 (1973), 363.

FI78 G.S. Fishman, Management Sci., 24 (1978), 510.

GP General Purpose Systems Simulator II, Reference Manual, IBM.

HO64 C.C. Holt, R.W. Shirey, D.V. Steward, J.L. Midler and A. Stroud, "Program Simulate, a users and programmers Manual", Social Systems Research Institute, University of Wisconsin, 1964.

HS77 D.A. Hsu, J.S. Hunter, Management Sci., 24 (1977), 181.

HU70 J.S. Hunter, T.H. Naylor, Management Sci., 16 (1970), 422.

KE73 J.A.G.M. Kerbosch, R.W. Sierenberg, Discrete Simulatie m.b.v. Algol, Fortran, PL/1, Samson, Brussel, 1973.

KI63 P.J. Kiviat, GASP-A general activity Simulation Program. Project 90 17-019(2), Applied Research Lab., U.S. Steel, Monroeville, Pa., 1963.

MA62 H.M. Markowitz, B. Hausner, H.W. Karr, Simscript, A Simulation Programming Language, The Rand Corporation RM-3310, 1962.

MI73 B. Mitchell, Opns. Res., 21 (1973), 988.

MO67 A.L. van der Mooren, Dutch Thesis, Delft, 1967.

MU78 P.J.W.M. Müskens, Dutch Thesis, Nijmegen, 1978.

NA66 T.H. Naylor, J.L.Balintfy, D.S. Burdick, K. Chu, Computer Simulation Techniques, New York, J. Wiley & Sons, 1966.

PU63 A.L. Pugh, Dynamo, User's Manual, Cambridge Mass., MIT press, 1963.

RE70 J.F. Reynolds, Opns. Res., (1972), 52.

SI62 Simpac User's Manual TM 602/00/00 Systems Development Corporation, Santa Monica, California, 1962.

ST77 H.J. Steudel, S.M. Pandit, S.M. Wu, Management Sci., 24 (1977), 456.

WA75 H.M. Wagner, Priciples of Operation Research, Prentice Hall Inc., London, 1975.

CHAPTER 5

SIMULATION OF A LABORATORY FOR STRUCTURAL ANALYSIS

In order to draw valid conclusions from experiments with the simulation model, the model should be a valid representation of the real system. According to Conway's [C059] opinion, some assurance of validity is provided when the model produces results that are not inconsistent with the known performance of the real system, for at least one alternative version of the simulated system and one set of conditions. This test is widely applied and is essentially a null test. A model which failed to pass is exceedingly suspect, but no strong statement can be made for a model which passed. The simulation of situations with known analytical results (e.g. M/M/n systems) can help to discover some programming defects.

In this section a statistical comparison is made of the output of our computer model with two actual situations of the laboratory, namely the situation during the period 6/1976 - 6/1977, described in Ch. II, and the period 6/77 - 6/78. The delays in the model were equalized to those of the period 76-77 by adjusting (i) the parameters defining the schedule of the other activities of the analysts, (ii) the priority differences between F(1) and F(2) samples, (iii) the priority difference between the samples which visited n and (n+1) sections, and (iv) the time spent to communicate the analytical result. Without changing the statistical parameters (means, variances, p.d.f.) of mentioned variables, and analysis times, and without a change of the kind

---

Published in part in: B.G.M. Vandeginste, Analyt. Chim. Acta CTO, accepted for publication.

B.G.M. Vandeginste, Communication presented at the IUPAC Congress, Helsinki (1979)

of distribution functions of the interarrival times, the output (mean delay's, mean number of waiting samples, correlations etc.) of the situation in 77-78 has been forecasted by substituting the observed traffic of F(1) and F(2) samples to the laboratory in the model and by adjusting the parameters defining the distribution of samples over the various sections. The statistics used are described in the preceeding chapter.

1. Validation of the model over the period 6/1976 - 7/1977

1.1 Validation of the parameters of the input and output density functions. The mean number of arrivals and departures per day of samples to respectively from the simulated laboratory are not significantly different from those in the actual laboratory (250 days of operation), as follows from the Student's t values listed in Table V-1. Kolmogorov-Smirnov (K-S) tests applied on the cumulative density functions of the input and output of the simulated and actual laboratory could not detect any significant differences. Although

Table V-1
Validation of the input and output density

Section	input (samples/day)			statistical test vs actual data compiled in Table II-1	
	mean	variance	$\phi(1)^3$	Student's t[1] t	K-S test[2] D_{max}
I.r.	2.7	5.5	-	0.46	0.039
P.m.r	7.4	12.1	-	0.80	0.072
M.s.	1.8	1.7	-	1.7	0.081
^{13}C-n.m.r.	2.7	3.0	-	1.1	0.095
Lab	11.2	21.2	0.40	1.2	0.097
	output (samples/day)				
I.r.	2.7	7.5	-	0.4	0.091
P.m.r.	7.4	44.7	-	0.5	0.099
M.s.	1.84	5.1	-	0.8	0.089
^{13}C-n.m.r.	2.65	10.1	-	0.4	0.090
Lab	11.13	64.0	-	0.9	0.160

[1] $t_{0.001}=2.6$

[2] $D_{0.05} = 0.119$

[3] only values significantly different from zero are tabulated

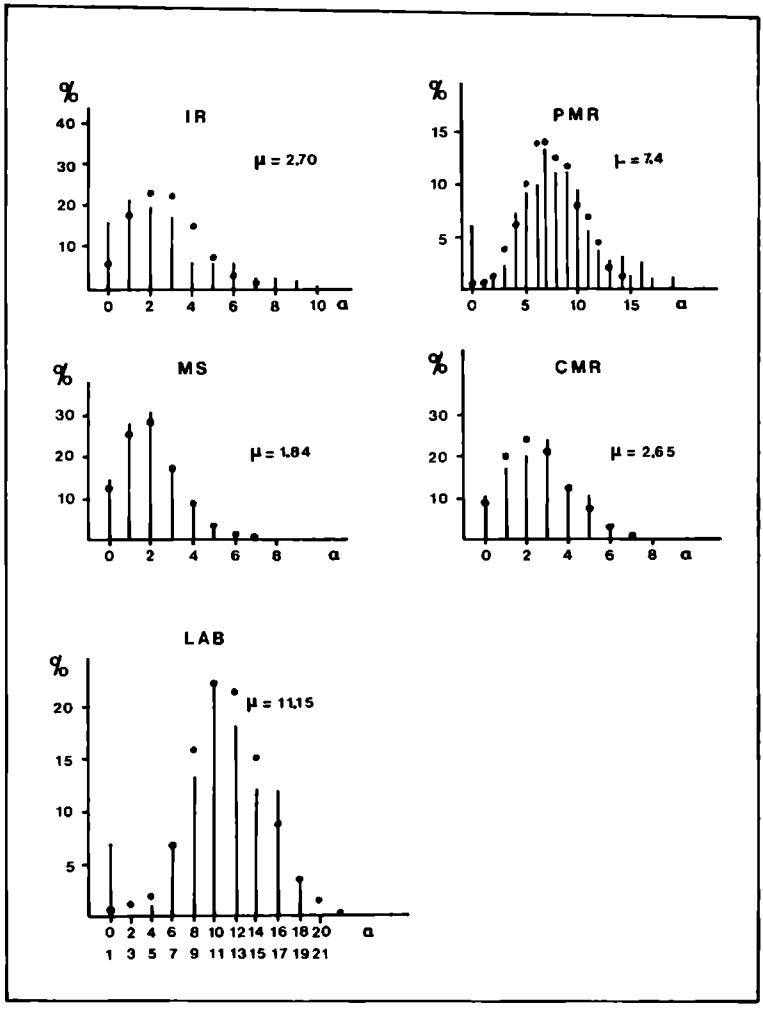


Fig. V-1: Histograms of the probability (%) of an input density α (samples/ day) to the sections and laboratory (model). — simulated values, ˙ Fitted Poisson distribution (same mean)

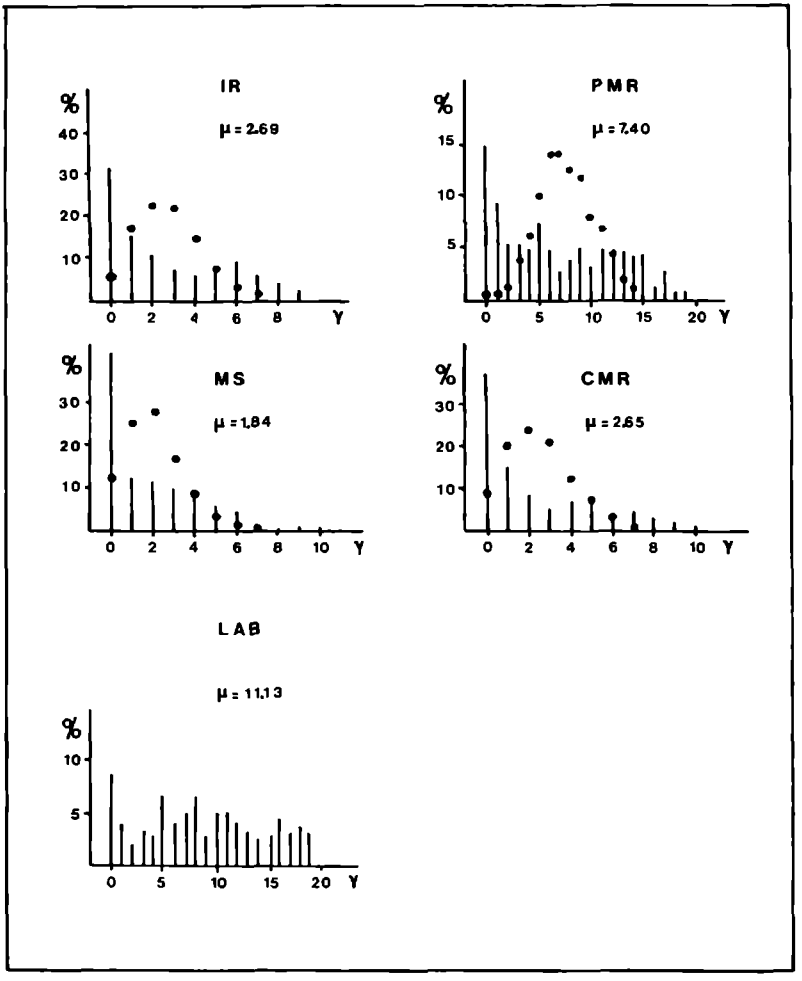


Fig. V-2: Histograms of the probability (%) of an output density γ (samples/ day) from the sections and laboratory (model). —— simulated values, . Fitted Poisson distribution (same mean)

these cumulative density functions do not meet the condition of continuity for application of the K-S test, this test may be used (De Jonge [J063]). However the tabulated value will be exceeded with a larger probability. These observations indicate that by the generation of two Poisson sample streams to the laboratory with different parameters (F(1): $\overline{IAT}$=0.93hrs and F(2): $\overline{IAT}$= 2.37hrs), which are distributed over the four sections, according to the

estimated probabilities that the various sections may give the requested analytical information, an input flow is obtained with the same statistical properties as observed in the actual situation. An important observation is that the cumulative density functions of the output flow of both, the simulated and actual situation, are not significantly different. That means that the mentioned effect in the laboratory, that the input and output rate density functions are completely different is also observed in the model (Fig. V-1 and V-2). A rather surprising result found in the model (as well as in the laboratory) is the memoryless property of the time series of the number of departing samples per day. There is no correlation between the number of finished samples at any day and the day before (Table V-1). Contrary to the actual laboratory (Fig. II-5) no significant periodicity could be detected in the autocorrelograms of the number of samples leaving the laboratory each day.

The conditional probabilities (p_{ij}) of transfer of samples between the sections (Table V-2) and the probabilities of each method to be selected first, secondly etc. (Table V-3), are in close agreement with the real situation (Tables II-3 and II-2), taking into account that the parameters regulating the sample flows through the network were calculated under the

Table V-2

Conditional probabilities (p_{ij}) for transfer of samples from one section to another in the model, and the probability (q_i) that a sample in node (i) leaves the system.

from \ to	OUT (q_i)	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.
OUT	-	0.17	0.57	0.10	0.16
I.r.	0.59	-	0.18	0.10	0.13
P.m.r.	0.85	0.06	-	0.04	0.05
M.s.	0.64	0.10	0.15	-	0.11
^{13}C-n.m.r.	0.78	0.07	0.09	0.06	-

assumption that the partition process over the sections was equal for samples arriving from outside and inside the laboratory. The percentage (% good) of samples that are successfully analyzed are indicated in Table V-3 for each section.

Table V-3

Probabilities for the selection of the methods in the model.

	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.	% of samples completed
first selection	0.17	0.57	0.10	0.16	80.2
% good	0.63	0.86	0.72	0.83	
second	0.22	0.39	0.17	0.21	93.0
% good	0.48	0.83	0.49	0.63	
third	0.28	0.19	0.26	0.27	95.7
% good	0.36	0.65	0.25	0.43	
fourth	0.20	0.01	0.35	0.44	100.0
% good	1.00	1.00	1.00	1.00	

1.2 Validation of the queue lengths.

In the proposed model, the queue length in an analytical section is defined as the sum of all waiting and unfinished samples. Student's t tests (Eqn. IV-11) did not detect significant differences between the mean queue lengths in the various analytical sections or in the total mean queue length in model and laboratory (Table V-4).

Table V-4

Validation of the queue lengths.

section	mean ($\bar{N}$)	var(N)	$\phi(1)$	var($\bar{N}$)	statistical tests vs actual data in Table II-5. Stud.t[1] (Eqn. IV-11)	$G(Q,1)^2$	$G(\Psi,1)^3$	$G(Q,\gamma)^4$	K-S[5]	K-S[6]
I.r	13.1	92.1	0.93	9.3	0.48	3.6	0.8	2.8	0.16	0.11
P.m.r.	30.1	219.9	0.87	12.4	1.9	1.9	1.7	0.2	0.23	0.09
M.s.	9.0	32.8	0.90	2.5	2.4	19.9	17.2	2.6	0.21	0.09
^{13}C-n.m.r.	19.4	73.1	0.90	5.3	0.98	1.2	0.002	1.2	0.12	0.05
Lab	71.6	514.4	0.90	37.4	1.5	6.8	6.4	0.44	0.22	0.06

[1] $t_{0.01}=2.58$ [4] $\frac{1}{2}\chi^2_{0.01}(2)=4.6$

[2] $\frac{1}{2}\chi^2_{0.01}(3)=5.7$ [5] test vs actual data: $D_{0.01}=0.145$

[3] $\frac{1}{2}\chi^2_{0.01}(1)=3.3$ [6] test vs Gaussian function: $D_{0.01}=0.096$

The fit of an exponential function through the autocorrelation functions of the queue lengths and a subsequently executed Bartlett test [BA46,MU78]

demonstrated that the time series of the queue lengths can be described with a first order autoregressive model. The comparison of the autoregressive parameter $\phi(1)$ and the residual variances (s_a^2) of the number of samples in the queues at the end of each day, in model and laboratory, by means of the G-statistics (Ch. IV) (Table V-4) demonstrated that the dynamic structure of the simulated queue lengths in all sections (except m.s.) is undistinguishable from that of the observed queues. Both time series are adequately described by an AR(1) model with equal parameters. This means that there are no significant differences in magnitude and velocity of queue fluctuations between model and laboratory. However, the autoregressive parameter $\phi(1)$ found in the m.s. section and total laboratory is somewhat too high. The Kolmogorov-Smirnov (K-S) test shows that the maximal difference (D) between the cumulative density functions of the number of waiting samples in the laboratory and model exceeds the value $D_{0.01}=0.145$. With such a result no strong statement can be made about the H_0 hypothesis that the two populations have the same distribution, since the data are not independent (high $\phi(1)$), causing a probability > 1% that D>0.145. Otherwise the calculated maximal differences (Table V-4) between the observed and Gaussian cumulative function demonstrate that the H_0 hypothesis that the number of samples in the model has a Gaussian shape cannot be rejected. Surprisingly, the from a theoretical point of view unexpected (Ch. III) Gaussian shape of the number of samples in the sections, observed in the laboratory (Ch. II) has been found again in all sections of the model. This Gaussian shape instead of the expected exponential shape will be explained by the simulation experiments presented in Ch. VI.

1.3 Validation of the delays.

Two problems were encountered validating the delays in the network. Firstly the same ratio between the delays of the samples which visited 1,2,3, and 4 sections should be obtained for the model and for the laboratory. This can be accomplished by adjusting the dynamic priority rule (Ch. III) between the samples which visited a different number of sections. Secondly, a good estimation of the variance of the mean delay is necessary.

(i) With the introduction of an urgency number that is dependent on the number of visited sections ($b_p = -nb_n$) Eqn. III-26 becomes $q_n = t_n - nb_n$. A variation of b_n between -100 and +100, varies the priority rule in the model from attributing absolute priority to samples which visited n sections over samples which visited already (n+1) sections, to the reversed situation. The effect

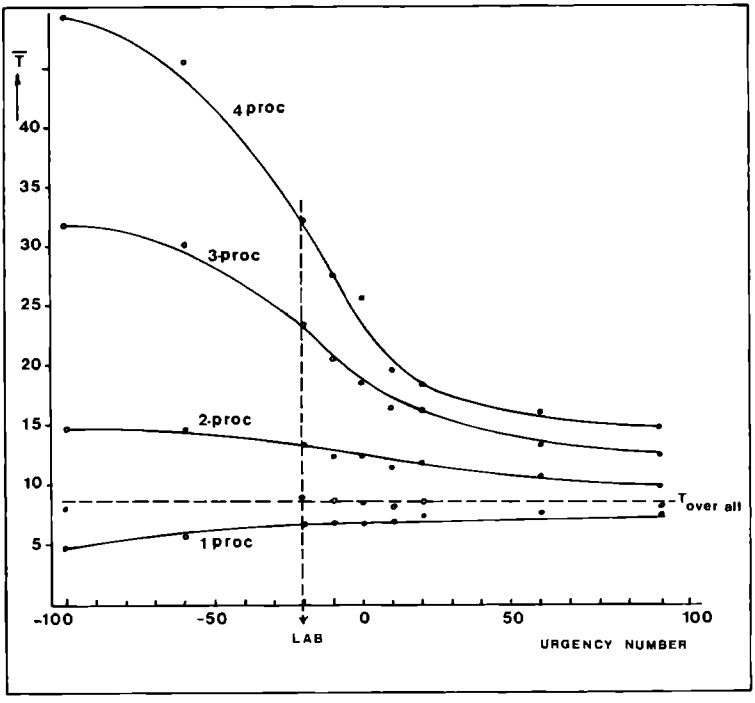


Fig. V-3: The average system time ($\bar{T}$) for samples that visited a varying number of sections as a function of the urgency number (b_n), defining the priority difference between samples that visited n and n+1 sections.

of the priority was investigated by means of the replication technique, described in Ch. IV. Fig. V-3 pictures the mean delays for samples subjected to 1, 2, 3, and 4 procedures as a function of the urgency number (b_n). These runs confirm the former statement based on Eqn. III-26 that the overall mean waiting time of all samples is not influenced by any priority rule, provided that the mean analysis times of all priority groups are equal. Fig. V-3 demonstrates clearly that the ratio of the delays of samples subjected to 1, 2, 3, and 4 methods, is strongly dependent on the priority difference attributed to these groups of samples. The cross sections of Fig. V-3 at different priority differences, presented in Fig. V-4, give a good indication of the relationship between delay and number of visited sections. Furthermore Fig. V-4 shows that in the laboratory, samples which visited the smallest number of sections have a lower priority over the other samples ($b_n \approx -10$). This implies that a small priority is given to samples which arrive from outside the

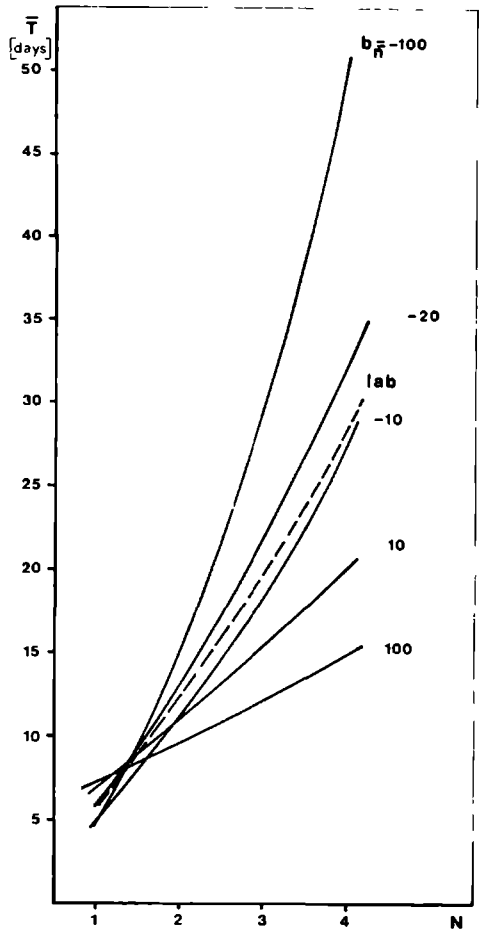


Fig. V-4: Simulated mean system time as a function of the number of visited sections and priority difference (b_n) between samples that visited n and n+1 sections. ---- :actual laboratory.

laboratory over those samples arriving from inside. This effect is very well demonstrated from the comparison of the delays of both kinds of samples, shown in Table V-5. The delay of the samples arriving from outside is significantly lower than the others. This is in agreement with the observations tabulated in Table II-8. From Fig. V-3 more general conclusions can be drawn. For $b_n = -100$, the differences between the mean delay times of the

Table V-5

Comparison of the delay times in the sections , of samples arriving from inside and outside the laboratory.

section	delay ($\bar{T}$) outside	var($\bar{T}$)	inside	var($\bar{T}$)	ratio	Student's t
I.r	4.12	0.10	6.55	0.36		2.1
P.m.r.	3.81	0.02	5.69	0.06	1.49	6.6
M.s.	4.56	0.23	5.54	0.42		1.2
^{13}C-n.m.r.	6.30	0.07	9.46	0.17	1.50	6.4
Lab	4.33	0.005	6.83	0.03	1.58	13.4

$t_{0.01}$ =2.58

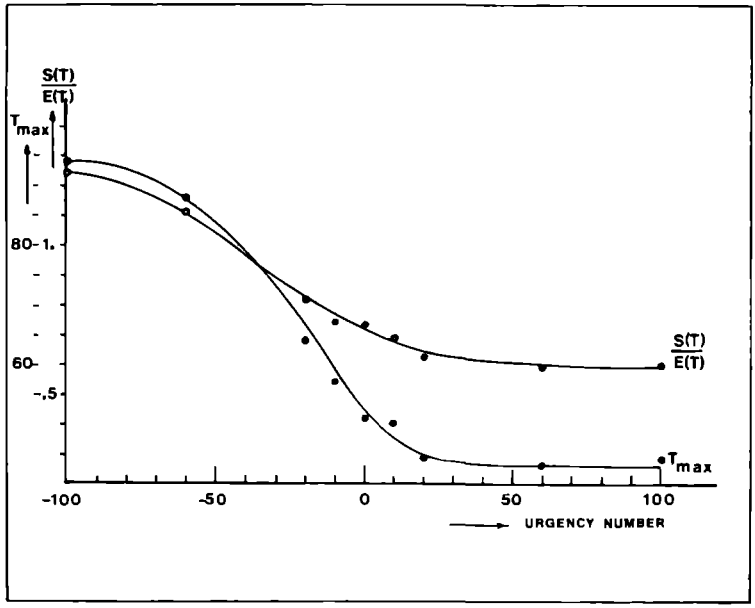


Fig. V-5: The maximal system time (T_{max}) and standarddeviation ($S_T/E(T)$) as a function of the urgency number (b_n) defining the priority difference between samples that visited n and n+1 sections.

four types of samples are much greater than for b_n=+100, whereas the overall mean delay time remains unaffected. As a consequence, the variation coefficient of the delay and the longest delay in the laboratory will be a function

of b_n, as demonstrates Fig. V-5. Here a minimal value for both performance variables is found for b_n=+100, under the assumption that there is no correlation between the analysis time and the number of visited sections.

(ii) The estimation of the standard error of the mean delay obtained by employing the batch means method and time series analysis method (Ch. IV) were tested against the estimation obtained by 12 replicated runs of the model. Therefore, the model was run under various conditions that should not introduce any effect, apart from producing fluctuating delays, caused by the statistic nature of the model. For example, the timing of the other activities is controlled by a random number generator, specifically dedicated to each analyst. By the exchange of analysts over the sections, a different delay should be found, from which an estimation of the standarderror can be obtained.
As a rule, the batch means method with 500 samples per batch gave an overestimation of the standarderror of the simulated mean delays (Table V-6). The estimated standarderror using the time series analysis approach, are close to the values obtained by the twelve replicated runs of the model (Table V-6).

Table V-6

Comparison of the standarderror of the mean delays (4000 samples-350 days), estimated according to the batch means method, and time series analysis method, with the standarderror obtained with 12 replica's of 4000 samples.

	section				samples with the same final method			
	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.
batch means[1]	0.73	0.81	0.97	0.34	1.04	0.89	1.24	0.62
time series	0.31	0.17	0.61	0.15	0.48	0.20	0.68	0.35
replica's	0.73	0.30	0.64	0.28	0.74	0.30	0.76	0.33

	samples with the same number of visited sections				overall
	1	2	3	4	
batch means[1]	0.66	1.18	2.00	2.80	
time series	0.15	0.48	1.40	1.60	0.16
replica's	0.19	0.45	1.20	1.20	0.23

[1] 8 batches of 500 samples

However, the standard errors of the delays in the sections were underestimated. This is probably due to the assumed first order autoregressive model that does not fit adequately the actual series of the delay. An indication that the

autoregressive model of the delay in the sections should not be of the first order is obtained by the discrepancies found between the value of the autocorrelation function at $\tau=1$, and the value $\exp(-1/T_x)$ calculated from a fit of an exponential function through the autocorrelograms (Table V-7). Apparently, the autocorrelation at $\tau=1$ of the delays in the sections is an underestimation of the correlation, resulting in an underestimated variance (Eqn. IV-13).

Table V-7

Comparison of the autocorrelation at $\tau=1$ ($\phi(1)$) and the estimated value of $\phi(1)$ from a fit of an exponential function through the autocorrelogram.

	section				samples with the same final method			
	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.
$\phi(1)$	0.60	0.55	0.33	0.35	0.28	0.41	0.32	0.19
$\exp(-1/T_x)$	0.90	0.98	0.85	<0.37	<0.37	<0.37	<0.37	<0.37
	samples with the same number of visited sections							overall
	1	2	3	4				
$\phi(1)$	0.55	0.33	0.20	0.22				0.30
$\exp(-1/T_x)$	<0.37	<0.37	<0.37	<0.37				<0.37

For the statistical comparisons, described in the next paragraphs, and the experimental design schemes described in Ch. VI, the standarderror was calculated by the time series approach (Eqn. IV-13). An exception is made for the delays in the sections where the values of the batch means method were used. The student's t tests executed on the mean delays of various categories of samples show that no significant differences could be found between the actual and simulated laboratory for 10 of the 13 calculated delays (Table V-8). Kolmogorov-Smirnov tests executed on the cumulative distribution functions show that the k=2 Erlang distribution provides a good fitting function (Fig. V-6). One should remark that this fit with a discrete function is allowed because, according to the real data, discrete delays (full days) can be obtained with the model.

Table V-8

Validation of the delays.

	Section			
	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.
mean (days)	4.8	4.1	5.0	7.4
variance	16.8	6.3	27.1	20.4
var($\bar{T}$)	0.25	0.42	2.1	0.15
student's t vs data in Table II-7	1.4	1.5	1.2	1.3
	samples with the same final method			
	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.
mean (days)	7.1	4.7	8.4	10.0
variance	51.8	13.7	71.6	73.1
var($\bar{T}$)	0.16	0.03	0.64	0.12
student's t vs data in Table II-7	1.35	4.8	1.4	3.1
	samples with the same number of analysis			
	1	2	3	4	overall
mean (days)	4.5	9.8	19.1	25.0	6.4
variance	9.4	26.5	69.9	107.7	52.8
var($\bar{T}$)	0.02	0.28	1.5	3.2	
student's t vs data in Table II-7	2.9	1.9	0.2	0.5	

$t_{0.01} = 2.58$

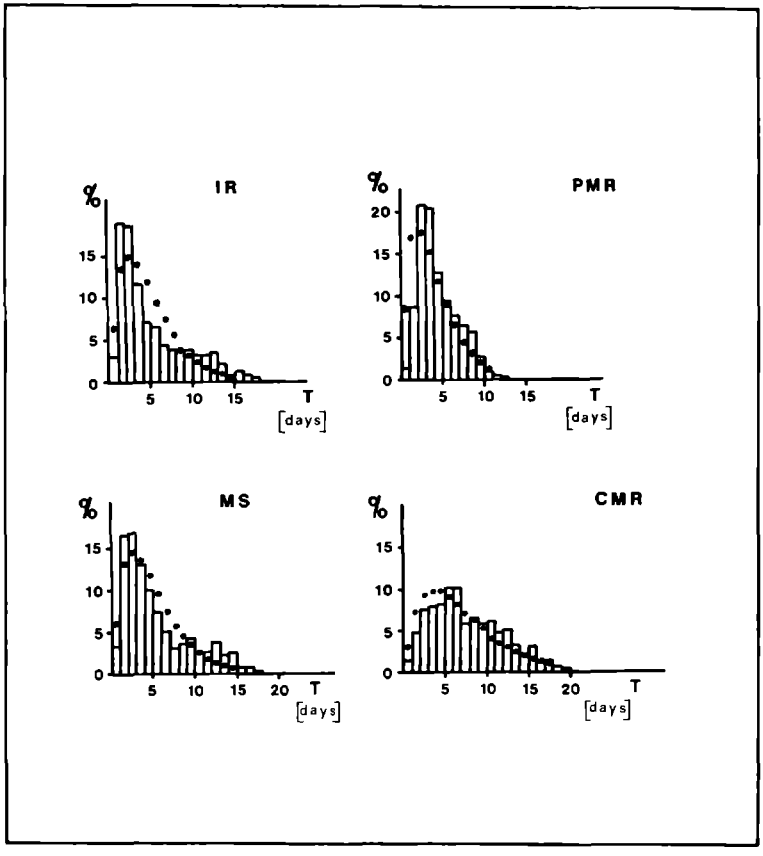


Fig. V-6: Histograms of the probability (%) of a delay (T) in the sections.
—— simulated data, . Fitted two stage Erlangian distribution.

1.4 Validation of the cross correlations in the system.

There is no mutual correlation between the number of samples in each section of the model, as shows the Table V-9. This agrees with the observations in the actual laboratory (Table II-10) and is a property of an open network [LE77].

For the simulation of the basic situation, the input flow of the laboratory and the sections was independent of the number of samples in the system. Cross correlation calculations, however, detected a small correlation between both variables (Table V-11). However, the calculated residual variances (%) are very high, which indicates that the major part of the fluctuations of

Table V-9

Maximal correlation between the number of samples in each section (99% confidence interval)

	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.
I.r.	1	+0.1 (0)[1] ±0.51[2]	+0.2 (+20) ±0.56	-0.4 (-25) ±0.56
P.m.r.		1	+0.2 (-15) ±0.48	+0.5 (+22) ±0.49
M.s.			1	+0.3 (+20) ±0.51
^{13}C-n.m.r.				1

[1]the time lag (τ) for maximal correlation

[2]99% confidence interval

the fluctuations of both variables are mutually independent. As approximately similar crosscorrelation values were calculated for the actual laboratory, the conclusion in Ch. II, that the samples are not preferably moved to the section with the lowest saturation degree, is supported. The correlations in the model between the number of samples in the sections and the delay of the samples arriving at the section (Table V-10) is higher than in the actual laboratory (Table II-12).

Table V-10

Maximal correlation between the number of samples (x) in the section and the delay (y) of the samples arriving at the laboratory (model)

	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.
ϕ_{xy}	0.62 (-5)[1]	0.70 (-4)	0.63 (-5)	0.43 (-5)
99% conf. interval	±0.30	±0.32	±0.28	±0.21
residual variance	0.62	0.50	0.60	0.81

[1]the time lag (τ) for maximal correlation

The maximal correlation at τ=-5 (Fig. V-7) can be explained by the departure of the samples ± 5 days after their arrival at a section. Thus, the delay of

Table V-11

Maximal correlation between the input flow (x) and the number of samples in the system (y)

	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.	Lab
ϕ_{xy}	+0.23 (+1)[1]	+0.41(0)	+0.10 (0)	+0.28 (?)	+0.49 (1)
99% conf. interval	±0.19	±0.20	±0.17	± 0.17	±0.17
residual variance	0.94	0.83	0.99	0.92	0.76

[1]the time lag (τ) for maximal correlation

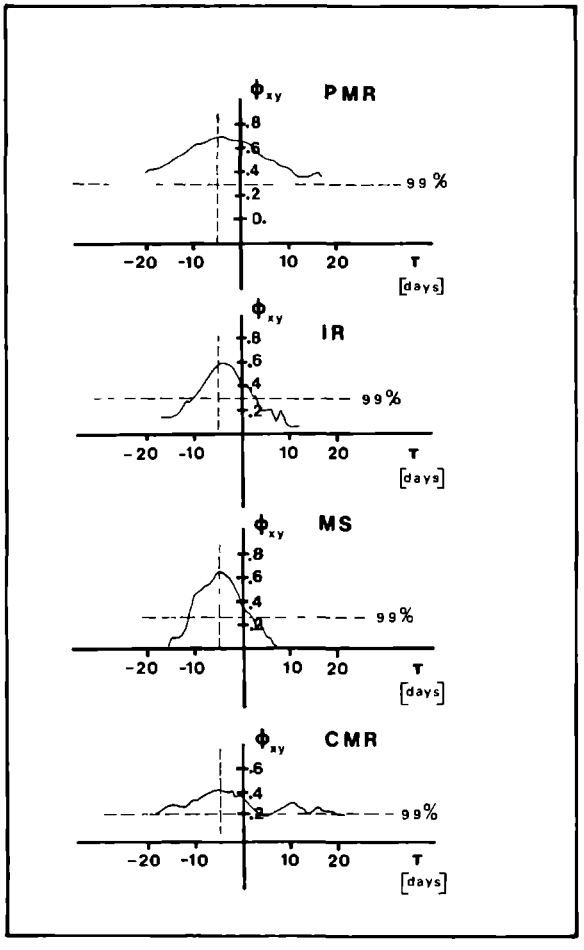


Fig. V-7: Crosscorrelograms (ϕ_{xy}) between the number of samples (x) in a section and the delay (y) of the samples. ---- 99% level for significancy from 0.

the samples, leaving a section at a time t has the greatest correlation with the number of samples in the system at the moment of their arrival, namely $\tau=-5$.

By the application of dynamic priorities, different priorities can be attributed to the F(1) or F(2) samples. This priority can be varied from FIFO (equal priority) to absolute priority. The effects on the delays for samples leaving the system through the same exit node are shown in Fig. V-8.

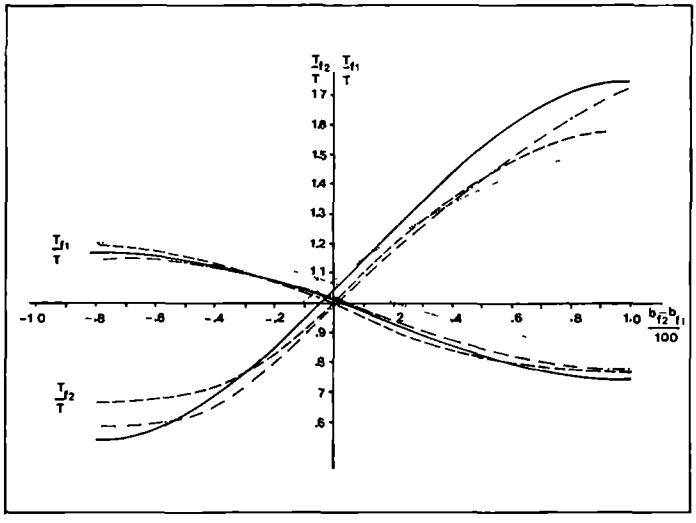


Fig. V-8: Simulated mean delay ($\bar{T}_{f1}/\bar{T}$ and $\bar{T}_{f2}/\bar{T}$) of a system with two groups of samples (F_1 and F_2), analyzed by the same final method, as a function of the priority difference ($b_{f2}-b_{f1}$) between both groups of samples.
—— final method is p.m.r.; -.-.- ^{13}C-n.m.r.; --- i.r.; .... m.s.

Although in the actual laboratory, significant differences could be detected between overall delays of both groups of samples only, and also between the delays of samples leaving the system through the i.r. section (Table II-9), it can be concluded that in the laboratory a small priority is attributed to the F(1) samples. Fig. V-8 indicates moreover that the urgency number (b_p) for the F(1) samples in the p.m.r., ^{13}C-n.m.r. and m.s. sections does exceed the urgency number of the F(2) samples with less than 10 to 20 time units. This means that a F(1) sample arriving at the laboratory has priority over all F(2) samples having a smaller delay than 10 to 20 hours (1 to 2 days) (the i.r. section: 30-50 hrs). Typically although the various sections have a

different (i) number of facilities, (ii) utilization factor, (iii) variation-coefficient of the analysis times and (iv) amount of other activities, a very similar effect is observed on the delays of the samples leaving the system through the various sections. According to the theoretical outline given in Ch. III, the priority difference has the greatest effect on the F(?) samples, with the smallest input flow to the laboratory (Fig. V-8).

2. *Forecast for period 77-78*

2.1 Adjustment of the flow.

The flows through the real laboratory during the period 77-78 and the period 76-77 are compared in Table V-12.

Table V-12

Comparison of the input flows (α) of the period 77-78 and 76-77 (Table II-1)

section	mean (α)	var($\bar{\alpha}$)	Student's t test[1] vs actual data in Table II-1
I.r.	1.6	0.02	4.1
P.m.r.	7.57	0.17	0.3
M.s.	1.93	0.03	0.6
^{13}C-n.m.r.	1.42	0.02	5.6
total intern flow	11.08	-	
Lab	10.28	0.3	2.1

[1] $t_{0.01} = 2.58$

This comparison indicates that the traffic to the i.r. and ^{13}C-n.m.r. sections did significantly decrease, while the total sample flow to the laboratory remained unchanged. This means that the mean number of visited nodes decreased also (total internal flow ($\Sigma\alpha_i$) decreased from 15.1 to 11.1 samples per day), which can only be caused by a change of the estimated probabilities (p_{ij}) that the various sections will solve the submitted analytical problem. The conditional probabilities of transfer of samples from one to another section and the probabilities (q_i) that a sample leaves the system through node (i) are shown in Table V-13. The probabilities that a method will be selected are shown in Table V-14.

Table V-13

Conditional probabilities p_{ij} of transfer of samples from one to another section, and the probability q_i that a sample in node i leaves the system.

from to	Out (q_i)	I.r.	P.m.r.	M.s.	^{13}C-n.m.r.
Out	-	0.11	0.67	0.12	0.11
I.r.	0.53	-	0.25	0.16	0.06
P.m.r.	0.90	0.03	-	0.05	0.02
M.s.	0.75	0.08	0.13	-	0.04
^{13}C-n.m.r.	0.82	0.06	0.05	0.06	-

Table V-14

Probabilities that the methods will be selected.

section	first	%correct	sec.	% correct	third	% correct	fourth	% correct
I.r	0.11	0.46	0.24	0.67	0.15	0.73	0.24	1
P.m.r.	0.67	0.90	0.33	0.84	0.28	0.96	0.24	1
M.s.	0.12	0.73	0.31	0.80	0.36	0.77	0.12	1
^{13}C-m.n.r.	0.11	0.83	0.11	0.78	0.21	0.80	0.41	1

2.2 Comparison between the forecasted and actual delays in the laboratory: period 77-78.

Tables V-15 and V-16 show that the mean number of waiting samples in 77 is decreased compared to the 76-77 period (except for the ^{13}C-n.m.r. section). As a consequence, a significant decrease of the delays is observed (except the i.r. and m.s. sections). The simulation of this situation reveals that, effectively, the observed decrease is also forecasted by the model. Anyhow, only significant differences between model and laboratory are found for some delays in the sections, and for the samples that visited two sections. The mean queue lengths and the dynamic behaviour (G(0,1) test) of the queues are not significantly different. However, the forecasted decay of the delay is too high for the i.r. and m.s. section. An argument for this discrepancy is that some modifications were introduced in these sections.

In the i.r. section some investigators (not member of the analytical staff) are permitted to measure their own spectra. If the spectrum is too complex,

Table V-15

Comparison of forecasted and actual mean queue sizes, and comparison of the parameters of the fitted AR(1) models.

section	actual						
	mean ($\bar{N}$)	var(N)	var($\bar{N}$)	$\phi(1)$	Stud. t^1 vs 76-77	s_a^2	
I.r.	6.6	22.5	1.1	0.85	2.8	6.24	
P.m.r.	27.3	209.5	8.4	0.82	2.9	68.6	
M.s.	8.3	24.8	1.2	0.86	3.3	6.5	
^{13}C-n.m.r.	11.4	44.4	4.0	0.92	2.2	6.8	
Lab	53.7	539.6	33.6	0.88	3.6	121.7	
	forecasted						
	mean ($\bar{N}$)	var(N)	var($\bar{N}$)	$\phi(1)$	Stud. t^1 vs 77-78	s_a^2	$G(0,1)^2$
I.r.	3.8	7.84	0.14	0.63	2.5	4.7	9.4
P.m.r.	28.8	187.7	9.1	0.85	0.4	52.1	2.8
M.s.	5.0	11.4	0.31	0.75	2.7	5.0	4.5
^{13}C-n.m.r.	12.4	19.3	0.6	0.77	0.5	7.9	5.4
Lab	50.0	282.2	16.1	0.87	0.5	68.6	10.7

$^1 t_{0.01} = 2.58$

$^2 \frac{1}{2}\chi^2_{0.01}(3) = 5.7$

the analyst helps solving the structure. However these samples were not included in the computed data, but influence certainly the waiting time of the other ones. The model did not account for this additional workload. In accordance with the real laboratory, the availability of the m.s. instrument is increased in the model. Apparently, the sensitivity of the model for this fact is too high, since the delay at the m.s. section decreases too much.
The cross correlograms and cross correlations in model and laboratory did retain the same behaviour as pictured in Tables V-9 and V-11.

Table V-16

Comparison of forecasted and actual delays in the laboratory: period 77-78.

section	actual			forecasted		
	mean (days)	var($\bar{T}$)	Stud. t^1 vs 76-77	mean (days)	var($\bar{T}$)	Stud. t^1 vs 77-78
I.r.	4.2	0.05	0.3	2.3	0.004	6.9
P.m.r.	3.6	0.01	10.0	3.9	0.49	0.4
M.s.	4.3	0.06	5.9	2.8	0.06	4.4
^{13}C-n.m.r.	8.1	0.19	2.6	6.6	0.06	2.3
Lab	5.2	0.02	7.9	5.1	0.01	0.7
samples with the same final method						
I.r.	5.9	0.14	0.5	5.4	0.23	0.9
P.m.r.	4.0	0.02	8.3	4.4	0.01	2.0
M.s.	6.7	0.14	4.3	5.7	0.13	1.9
^{13}C-n.m.r.	9.8	0.35	2.0	8.1	0.12	2.5
samples with the same number of analysis						
1	4.0	0.01	6.5	3.8	0.005	1.1
2	9.1	0.13	3.5	7.7	0.06	3.5
3	16.3	1.05	2.2	12.9	0.79	2.6
4	25.3	9.5	0.35	17.8	0.51	0.6

$^1 t_{0.01} = 2.58$

3. Conclusions

In this section the possibility is demonstrated to model an analytical laboratory on the basis of data collected during one year operation of the laboratory. With this model, the output of the laboratory during that year has been simulated. The sample input to the laboratory and the various sections has been adequately described by the generation of two Poisson sample streams to the laboratory, which are distributed over the four sections, according to the estimated probabilities that the various sections will give the requested

information. An even more important conclusion is,that the model correctly forecasts the effect of changes which occurred in a later period of the real laboratory. The model was applied successfully and can quantify some unmeasured parameters, such as the attributed priorities to the various groups of samples. As a result, the effect of the variation of those parameters can be calculated, or the value of these parameters can be calculated in order to obtain a desired behaviour of the laboratory. Some experiments with the model, in order to quantify those unmeasured parameters,have demonstrated that the variation coefficient of the delay and the longest delay in the laboratory is minimal when in the sections absolute priority is attributed to the samples that visited already the most sections, under the assumption that there is no correlation between the analysis time and the number of visited sections.
In accordance with the theoretical calculations on priority queueing in M/M/1 systems, the model proves that the attribution of a priority difference between the F(1) and F(2) samples, has the greatest effect on the F(2) samples, having the smallest input stream to the laboratory.

References

BA46 M.S. Bartlett, J. Royal Stat. Soc., B8 (1946),27.

CO59 R.W. Conway, B.M. Johnson and W.L. Maxwell, Management Sci.,6 (1959), 92.

JO63 H. de Jonge, Inleiding tot de Medische Statistiek, Walters Noordhoff, Groningen, 1963.

LE77 A.J. Lemoine, Management Sci., 24 (1977), 464.

MU78 P.J.W.M. Müskens, Dutch Thesis, Nijmegen, 1978.

CHAPTER 6

FORECAST OF THE EFFECT OF MODIFICATIONS OF THE LABORATORY ORGANISATION ON THE DELAY BY DIGITAL SIMULATION

The principal objective of the design of a simulation model is to conduct simulation experiments in order to learn more about the system under investigation. The effects of some variations of parameters, variables or operating characteristics can be estimated. For example, the sensitivity of the system for the value of the mean interarrival time can be estimated. The aim of a simulation experiment may be twofold: the exploration and description of the response surface of the system over some region of interest in the factor space, or the optimization of this response in the presence of a large amount of variables and parameters. Very often, the influence of some variables is dependent on the level of the other variables. As a result an interaction can be found between the variables. For example, the effect of a decrease of the interarrival time will be dependent on the priority of the considered group of samples. In order to minimize the number of requested experiments, exploratory experiments should be conducted by means of experimental designs [DA75] and optimization experiments by means of experimental optimization techniques, such as the steepest ascent method [BR59] and Simplex method [DE73]. In this section, the results of some experimental designs and studies on functional relationships between some independent variables are presented. An extensive discussion of factorial designs, along with methods for constructing and analyzing the designs is given by Davies [DA71].

B.G.M. Vandeginste, submitted for publication.

1. Strategies concerning priorities.

1.1. Analysis dependent priority.

In Chapter III, it was demonstrated that the 'shortest analysis time first'(SAT) priority gives the smallest delays in queueing systems. The effect of the application of this priority rule, along with the introduction of estimated analysis times and a 'shortest expected analysis time first (SEAT)' discipline was investigated starting from the actual situation in the laboratory. Applying the SEAT discipline in the model, the samples are scheduled acccording to an estimated interpretation time, according to Eqn. VI-1

$$(IT)_{expected} = (IT)_{real} + s_{IT} * r \qquad \text{VI-1}$$

where s_{IT} is the standard error of the estimation of the analysis time, and r is a Gaussian distributed random number with zero mean and a standard error equal to one. The graph (Fig. VI-1) of the overall delay as a function of the precision of the estimation of the interpretation time confirms the expectation that $\bar{T}_{SAT} < \bar{T}_{SEAT} < \bar{T}_{Random}$.

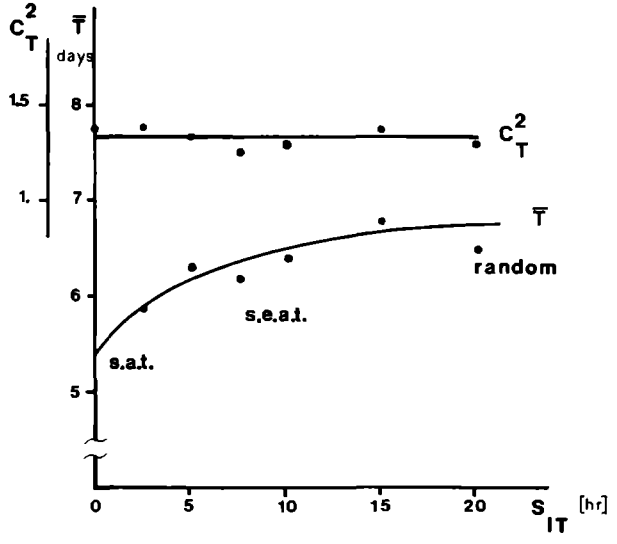


Fig. VI-1: The 'shortest expected analysis time first' (SEAT) priority rule. Simulated effect on the mean system time ($\bar{T}$)and variation coefficient of the delay (C_T^2) as a function of the standard error of the estimation of the analysis time.

The relative superiority of the shortest analysis time first operation rule is consistent with previous research from Conway et al. [C067]. When the interpretation time is exactly estimated, the reduction of the overall delay is about 20%, while the variation coefficient (C_T^2) of the delay is not affected by the application of this priority rule. The accuracy of the estimation of the interpretation time, necessary to obtain mentioned reduction is very low (s_{IT}=2.5 hrs). An alternative procedure is the separation of the samples into two groups: the so called 'easy' and 'difficult' samples with respectively 'small' and 'large' analysis times. The validity of the results obtained in Ch. III 2.2 was checked against simulations with the model. The strategy applied in the model, concerning the 'easy' samples was as follows:
(i) 'easy' samples visit one section of the laboratory only, and are selected on the basis of their interpretation time. i.e. there should be no doubt that the section will give the requested information.
(ii) the measurements of 'easy' samples are started, even when the minimal batchsize required for the measurements is not present. The results of 'easy' samples are immediately communicated to the client. The presence of 'easy' samples does not affect the other activities.

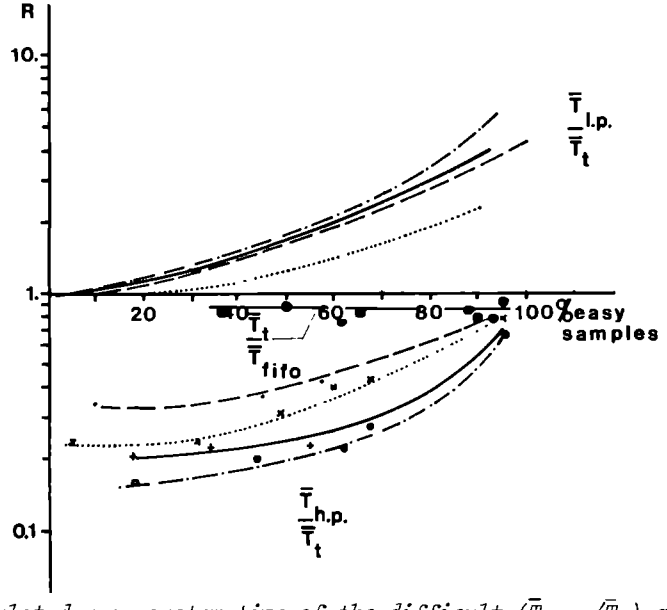


Fig. VI-2: Simulated mean system time of the difficult ($\bar{T}_{l.p.}/\bar{T}_t$) and easy ($\bar{T}_{h.p.}/\bar{T}_t$) as a function of the fraction of easy samples. Absolute priority is assigned to the easy samples. (——) final method is p.m.r., (-.-.) ^{13}C-n.m.r. (---) i.r., (....) m.s.
● *Reduction of the overall maen delay ($\bar{T}_t/\bar{T}_{fifo}$)*

The comparison of Figures III-16 and VI-2 demonstrates that the separation of the samples into two categories has the same effect on the difficult samples as forecasted by theoretical calculations presented in Ch. III.2.2. When 70% of the samples belongs to the category of 'easy' samples, the delay of the 'difficult' samples is increased by a factor 2 to 2.5, while the delay of the other samples is approximately reduced with a factor 2.5 to 5. As a result, the overall delay is reduced with 20%. A $4^1x4^1x2^1$ factorial experiment demonstrated that the accuracy of the estimation of the interpretation time (standard error between 0 and 40%) did not influence significantly the delay of the samples grouped in the category 'easy' and 'difficult' samples.

Table VI-1

Effect of the accuracy of the estimated interpretation time on the delay of 'easy' and 'difficult' samples, when 'easy' samples have absolute priority. Analysis of variance of a $2^1x4^1x4^1$ design.

factor levels:
(A) sections (i.r.,....m.s.)
(B) standard error (% of interpretation time): 0, 10, 20, 40
(C) % easy samples: 35, 56

	easy samples				difficult samples			
source of variation	sum of squares	degrees of freedom	mean square	var ratio	sum of squares	degrees of freedom	mean square	var ratio
A	7.84	3	2.61	3.0†	326.0	3	108.7	83.8†
B	0.058	3	0.019	2.18	4.42	3	1.47	1.1
C	0.014	1	0.014	1.61	13.9	1	13.9	10.7
residu	0.208	24	0.0087		31.10	24		
total	8.12	31			375.41	31		
	1- procedure samples							
A	0.10	1	0.10	40.				
C	1.3	3	0.43	43.†				
residu	0.05	3	0.01					
total	1.45	7						

† highly significant (99%)

When the goal of attributing priority to 'easy' samples is to minimize the overall delay of all samples, a small fraction (~10%) of the samples with long interpretation times should be designated as 'difficult' samples and give absolute priority to all other samples.

1.2. Priority based on the number of visited sections.

In Chapter V, it was demonstrated that the functional relationship between the delay and the number of visited sections is highly dependent on the priority difference between the samples that visited a different number of sections in the laboratory (Fig. V-4).

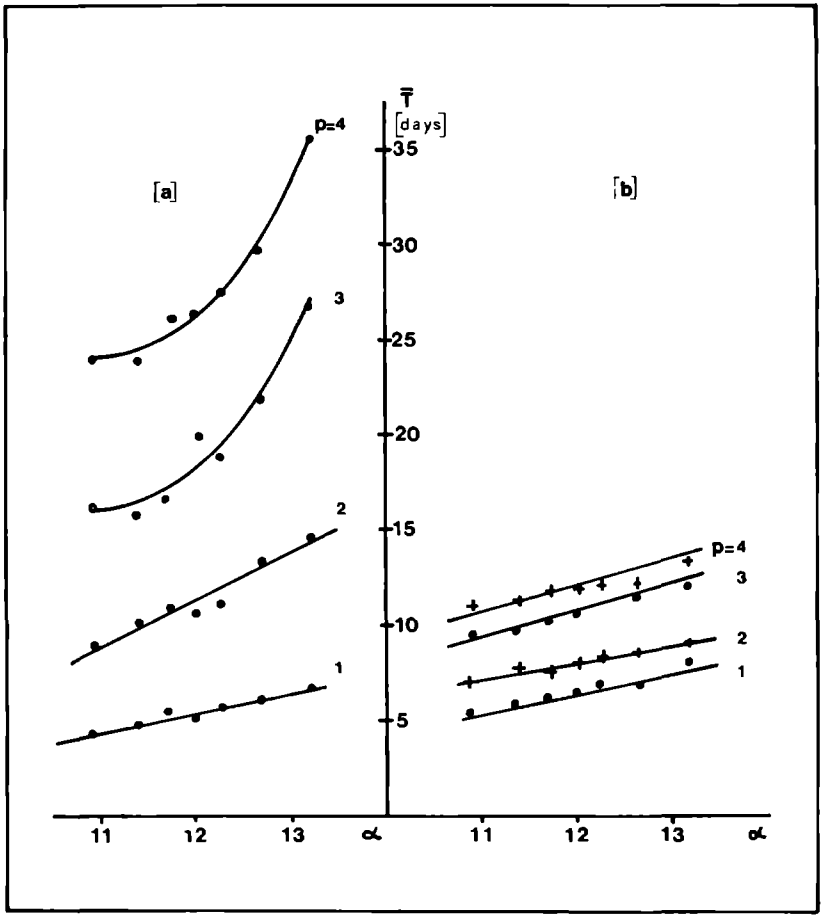


Fig. VI-3: Flow (α) (samples per day) dependency of the mean delay ($\bar{T}$) of samples as a function of the number of visited sections.
(a) samples which visited the less sections have absolute priority.
(b) reversed situation.

When samples, originating from outside the laboratory have absolute priority on the samples arriving from inside the laboratory, a strong dependence is found between the delay and the number of visited sections

By inversion of this priority rule an appreciable loss of this dependency is observed (Fig. V-4). An investigation on the sensitivity of these groups of samples for an increase of sample flow to the laboratory resulted in the following conclusions (Fig. VI-3). The sensitivity of the delay of the four groups of samples for an augmentation of the flow to the laboratory is highly dependent on their attributed priority. If samples arriving from outside the laboratory have priority, the delay of the other samples is appreciably sensitive for the total flow. In the opposite situation, where the priority of the samples increases with the number of visited sections, the dependency of the delay on the flow becomes quite similar for all groups. Furthermore, it can be remarked that, apparently, the delays of the smallest groups of samples (group 2 to 4) are very dependent on the attributed priority. Evidently, the relationship between the overall delay and input flow is independent of the applied priority rule (Fig. VI-4b) between samples that visited a different number of sections. Similarly, the variation coefficient of the overall delay is not dependent on the input flow (Fig. VI-4a).

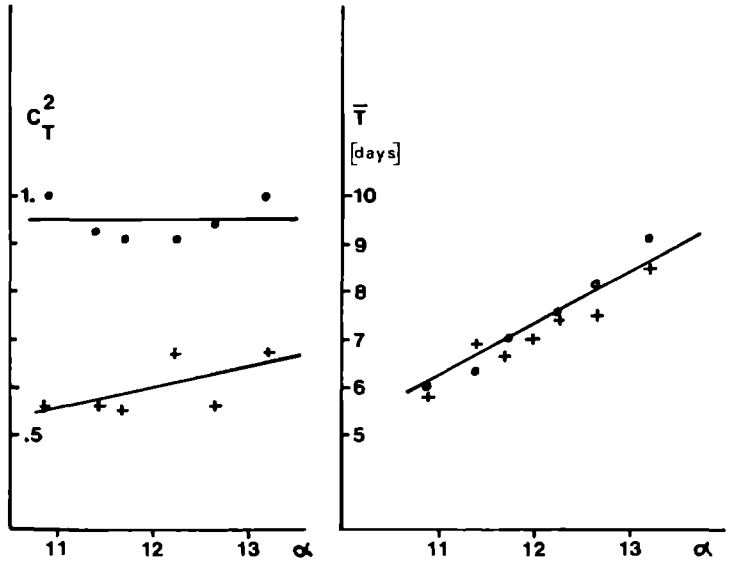


Fig. VI-4: Flow(α) (samples per day) dependency of the overall mean system time($\bar{T}$)and variation coefficient (C_T^2) of the delay, simulated for two priority rules: (+) samples that visited the most sections have absolute priority, (o) reversed situation.

1.2. Priority dependency on sample source.

In Chapter III, it was demonstrated that the sample group with the smallest input flow to the laboratory has the highest sensitivity for its priority compared to the other group of samples. Furthermore, from theoretical calculations on a M/M/1 system (Ch. III), it is expected that the level of the input flow to the laboratory will influence mostly the group of samples with lowest priority. This sensitivity was simulated for two extreme situations where one of the groups get absolute priority over the other one. The ratio between the sensitivities of low priority samples and high priority samples for the two extreme situations are tabulated in Table VI-2 (R_1 and R_2).

Table VI-2

Sensitivity (s) of the delay for a variation of the input flow ($\Delta\alpha/\alpha$)=0.20) of high and low priority samples ($s=\Delta T/\Delta\alpha/\alpha$); flow ratio α_1/α_2 =2.6

Group with absolute priority	F-1 (α_1=8.6)			F-2 (α_2=3.4)		
final method	s_{F1}	s_{F2}	$R_2=s_{F2}/s_{F1}$	s_{F1}	s_{F2}	$R_1=s_{F1}/s_{F2}$
I.r.	0.2	0.4	2.0	0.28	0.06	4.7
P.m.r.	0.11	0.24	2.2	0.16	0.05	3.2
M.s.	0.13	0.16	1.2	0.14	0.05	2.8
^{13}C-n.m.r.	0.17	0.21	1.2	0.24	0.08	3.0

The values of R_1 and R_2 are all greater than one. This indicates that the high priority samples are less sensitive to a variation of the flow, than the low priority samples, irrespective of the magnitude of this sample group. Moreover, a comparison of the R_1 and R_2 values in that table, indicates that for each section $R_1>R_2$: i.e. the ratio between the sensitivities of low priority and high priority samples is the greatest when the largest sample group F(1) has absolute priority. The preceeding Chapter indicated that the mean delay of this greatest group of samples (F1) in the laboratory is hardly dependent on the priority difference with the other group. Now from the comparison of the effect of the flow on the delay of the greatest group of samples, when having absolute priority and not (columns 1 and 4 in Table VI-2), it appears that this effect is hardly dependent on the attributed priority also.

2. Dispatching decisions.

The effect of the introduction of an adaptive routing procedure in the model has been investigated. The routing algorithm takes into account the probability (j(i)) that a section (i) will give the requested information, and the queue length in each section, normalized on the total workload of the laboratory (Eqn. VI-2)

$$R(i) = f(1-j(i))+(1-f)N_i/\sum_{i=1}^{L} N_i \qquad \text{with } 0<f<1 \qquad \text{VI-2}$$

The sample is firstly routed to the section with the smallest R value. Following example demonstrates the consequence of the application of Eqn. VI-2: For f=0.4 a sample is routed to a section with j=0.5 instead of j=1, provided the number of waiting samples in the former section is 10 units smaller, having totally 30 samples waiting in the laboratory. The effect of the algorithm was calculated considering two starting points: i.e. assuming that all (respectively no) samples are successfully analyzed in a section with j=1 (respectively: j=0), and secondly: assuming that 16% of the samples which are directed to a section with j=1 are not successfully analyzed, and consequently are routed to another section; in addition, 6% of samples submitted to a section with j=0 are completed in that section (Table IV-2).i.e. the probability of finding the structure with a given method is estimated less accurately. The model demonstrates that balancing the probability of obtaining the requested information, against the number of samples in the sections decreases the delay. However, the observed effect is relatively small (≈12%) (Fig VI-5). In terms of variation coefficients of the overall delay the model is insensitive to that strategy. By attributing a too large importance to the number of waiting samples (f<0.2), the mean number of visited sections increases from 1.26 to 1.40, resulting in an increase of the delay, which is very sensitive to that number. The more inaccurate the estimated probability is that some section will give the requested analytical information, the more useful it is to balance this probability against the number of samples in each section. Even, when the state of the laboratory is considered exclusively, no increase of the mean number of visited departments is observed, and the mean delay has diminished. Obviously, an augmentation of the input flow to the laboratory has the same effect as an increase of the mean number of visited sections per sample (Fig. VI-6)

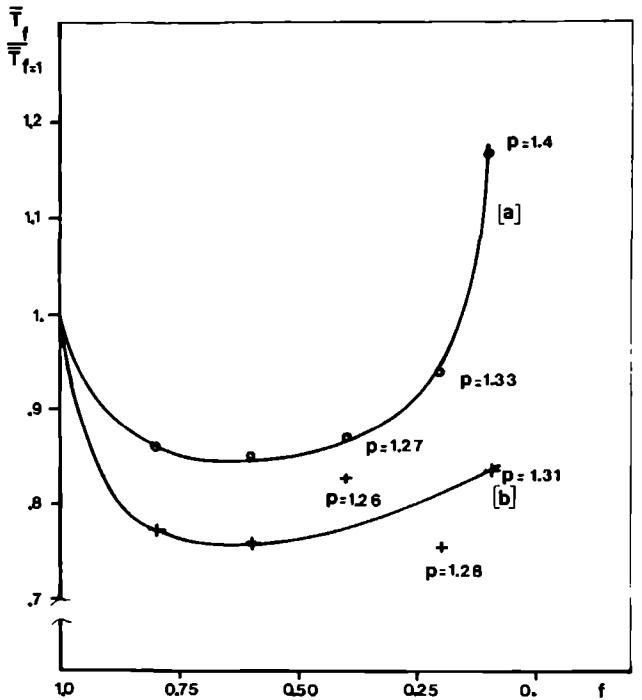


Fig. VI-5: Simulated mean system time ($\bar{T}_f / \bar{T}_{f=1}$) of all samples as a function of the weighting factor (f) balancing the queue sizes in the sections versus the probability for a section to give the requested information.
f=0: sample routing based on the queue sizes exclusively.
f=1: sample routing based on the probabilities exclusively.
o the estimated probability that a section will give the requested information is exactly known.
+ this probability is incorrectly estimated for 16% of the samples. p is the mean number of visited sections.

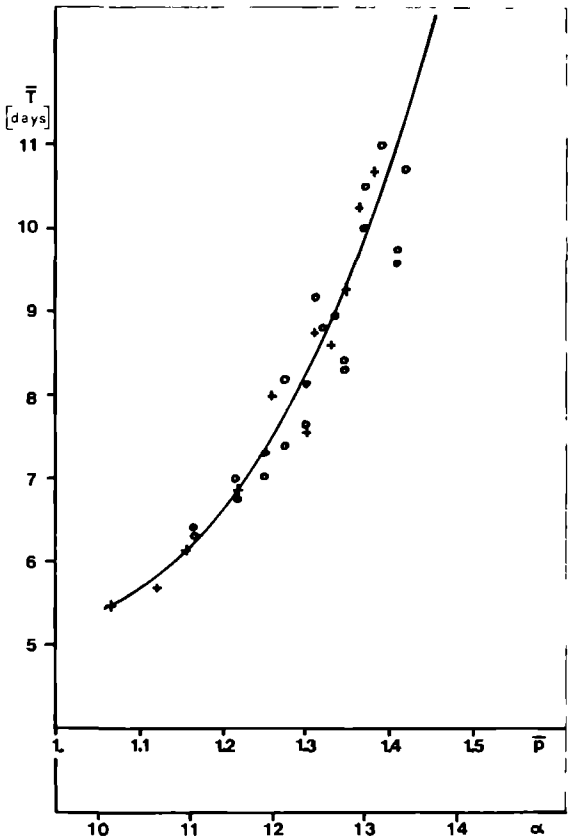


Fig VI-6: o Simulated effect of the input flow (α) to the laboratory (samples per day) on the overall system time ($\bar{T}$)
+ Simulated effect of the mean number ($\bar{p}$) of visited sections on the mean system time ($\bar{T}$)

3. Analyst assignment decisions.

A completely decentralized organization was assumed for the simulation of the actual laboratory. The analyst is always assigned to the same analytical section, irrespective the state of the other sections. The effect of the work selection rules, described in section IV, in a centralized laboratory organization was investigated. In the centralized organization the experience of the analyst for the different analyses is balanced against the state (queue lengths) in the sections. In the model a relationship was assumed between the experience of the analyst (j) and his mean analysis time in section (i) (Eqn. VI-3)

$$\overline{AT}_j = \overline{AT}/\exp(j,i) \qquad \text{VI-3}$$

The results of the simulations with a centralized organization where all analysts have experience with all methods, and where the selection of the sample is independent from the state of the laboratory, are shown in Fig. VI-7. In these runs the analyst selects the sample of his greatest experience, for which an unmanned instrument is available, without regarding whether eventually a more experienced analyst is idle (rule 1). Fig VI-7 demonstrates the disastrous effect of allowing analysts to analyze samples without sufficient experience (<0.8). A reduction of the overall delay is only achieved when all analysts are fully qualified for all methods. The very small effect of the extension of the number of analysts in a completely centralized organization (all analysts are fully qualified) indicates that the instruments are the bottleneck of the system and not the number of analysts. As a result, a temporarily admission of analysts will not influence the delay significantly. When an analyst is authorized to do ar analysis for which he is not fully qualified, provided no fully qualified collegue is idle (rule 2), a somewhat smaller effect on the delay is observed (Fig. VI-7). However, the conclusion remains valid that under the conditions of the laboratory, the introduction of a decentralized organization has only sense if the analysts are allmost fully qualified for the other methods (mean analysis time exceeding the mean analysis time of a specialist with less than 10%). Clearly, the inclusion of the lengths of the various queues in the decision, which section the analyst will select next, will not influence the effect of centralizing the organization, when all analysts are fully qualified for all analytical procedures (Fig. VI-8).
On the contrary, when the experience of the analyst for methods beyond his

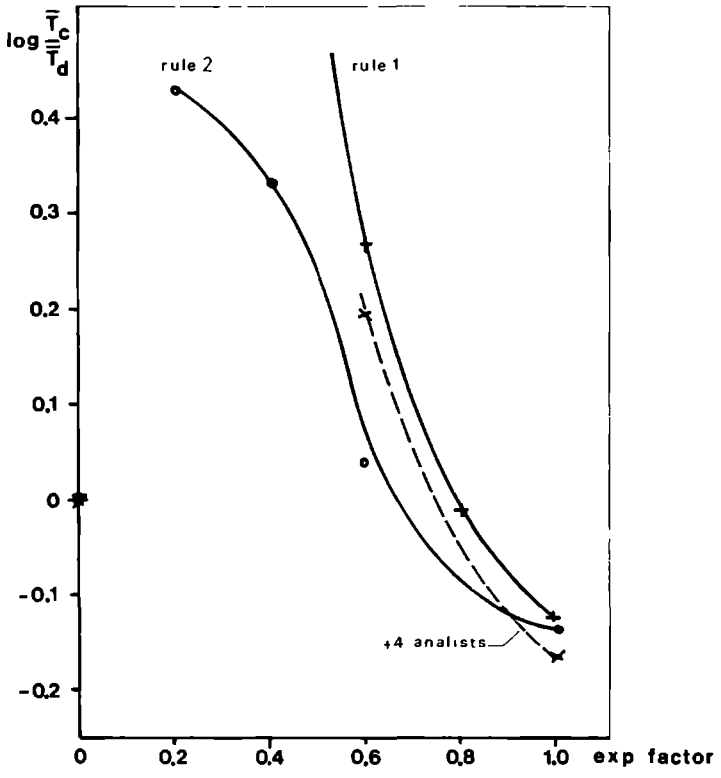


Fig. VI-7: Simulated effect of the analyst assignment decisions on the mean system time , employing priority rule 1 and 2 (see text).
exp. factor=0 : completely decentralized organization (mean system time: $\bar{T}_d$)
=1 : completely centralized organization (mean system time: $\bar{T}_c$)
(——) the system is extended with 4 analysts (from 9 to 13) and assignment rule 1 is used.

own specialism is small (exp<0.6), the consideration of the queue lengths according to algorithm VI-4, amplifies the bad influence of centralization (Fig. VI-8)

$$\text{Max}[\text{exp}(i,j) + T_{first}(i)] \qquad \text{VI-4}$$

where $T_{first}(i)$ is the total delay of the sample in front of queue(i)

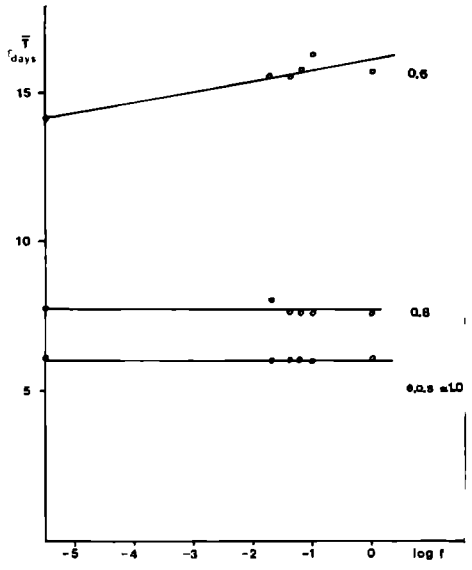


Fig. VI-8: Simulated effect on the mean system time ($\bar{T}$) of weighting the experience of the analyst against the delays of the samples in front of the queues in a centralized laboratory organization. The experience with the other methods than the own specialism (e.o.s. varies from 0.6 to 1.0) logf=-5: no state dependency.

4. Strategies on the termination of the analysis.

In the model a maximal allowed analysis time could be selected according to Eqn. VI-5:

$$AT_{max} = \overline{AT}(1+C^2_{AT})f \qquad \text{VI-5}$$

where the value of f depends on the simulated strategy:

(a) the maximal analysis time is independent of the properties of the sample or state of the laboratory (f=1)

(b) the maximal analysis time increases with the number (N) of visited sections (f=1+N)

(c) the maximal analysis time decreases with the queue length in section (i) ($f=\sum_{i=1}^{4} N_i/N_i$)

(d) the maximal analysis time decreases with decreasing probability that the method will solve the analytical problem (f=1+j(i)/0.5)

The effects of these strategies were compared to the actual situation where

no limitations on the interpretation time are imposed on the analyst. In contrast to the expectation based on the study of a M/M/1 system, which is not imbedded in a network, no improvement of the delay could be found. Strategies (a) and (b) cause an augmentation of the delay, even with 50%. This is caused by the increase of the mean number of visited departments from 1.27 to 1.47, respectively to 1.52. Apparently, the effect of an augmentation of the mean number of visited sections (20% for strategy (a)) surpasses the effect of a smaller mean analysis time (13.5% for strategy (a)). It causes an augmentation of the utilization factors and consequently, the delay. The application of strategies (c) and (d) gave no improvement of the system performance as compared to the actual situation.

5. Batch analysis of samples.

5.1 Effect on the mean delay.

By means of a 2^1x3^2 factorial design (Table VI-3), the effects on the mean delay and the interactions between the minimal and maximal batchsize, and overhead factor were determined. The analysts start the measurements of the samples when the minimal batch size is present in the section. The overhead factor is a reduction factor of the measurement time, taking into account that the treatment of samples can be executed simultaneously. No overhead was attributed to the interpretation time.

The factor levels used in the factorial design were as follows:

factor A: minimal batch size: 1 (loose rule), maximal batch size (tight rule)

B: overhead: 1.0, 0.9, 0.8

C: maximal batchsize: 0.5λ, λ, 1.5λ (λ: input density to the section)

From the results tabulated in Table VI-3, it is seen that the overhead and minimal batch size have a pronounced effect on the mean delay, along with a small interaction between the overhead and the maximal batch size. The loose rule, that the analyst should not wait until a sufficiently large batch of samples is present, performs better than the tight rule, where the analyst should wait. An interesting observation from Fig. VI-9, showing the response plane T=f(overhead,min. batch size), is that even for large overhead factors (20%) it is advantageous to start the measurement of a sample without delay. However, it should be stressed that the schedule of the other activities of the analysts in the model is independent of the state of the laboratory. This means that the schedule of these activities is not altered when the maximal batch size is not present. In the opposite situation a smaller effect of the

minimal batch size can be expected, because here other activities are preferably executed during the time that the batch size is not reached.

Table VI-3

Effect of the minimal batch size (A), overhead (B) and maximal batchsize (C). Analysis of variance of a 2^1x3^2 design.

source of variation	sum of squares	degrees of freedom	mean square	variance ratio	effect
main effects					
A	2.6	1	2.6	54.1+	+ 0.8
B	5.24	2	2.6	54.1+	- 1.3
C	0.32	2	0.16	3.3	
two factor interactions					
AxB	0.07	2	0.034	0.7	
AxC	0.25	2	0.125	2.6	
BxC	1.43	4	0.36	7.5*	
three factor interactions					
AxBxC	0.01	4	0.0025	-	
var($\bar{T}$)			0.048		

+ significant 1%<P<5%
* significant P<1%

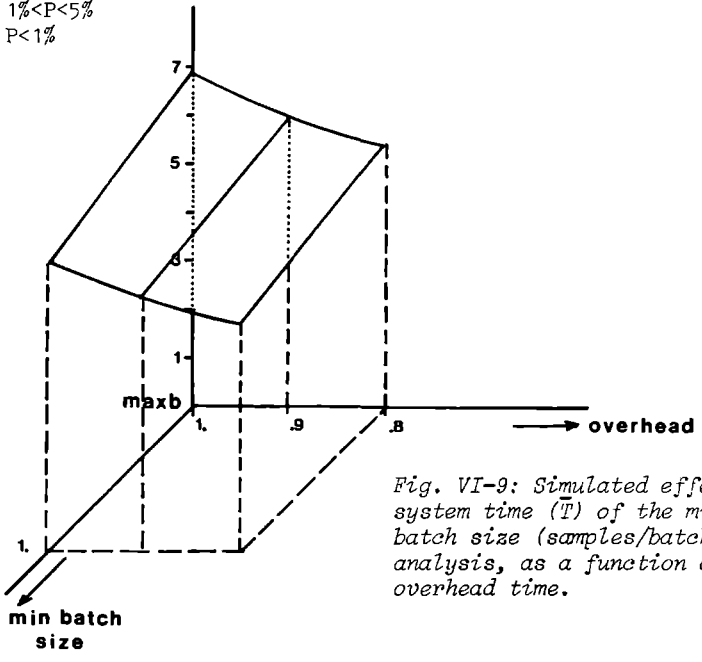


Fig. VI-9: Simulated effect on the system time ($\bar{T}$) of the minimal batch size (samples/batch) for analysis, as a function of the overhead time.

5.2. Effect on the distribution of the number of samples in the system.
In the preceeding Chapter, the unexpected Gaussian shape of the probability density function of the number of samples in the sections in model and laboratory was noticed. An investigation on the possible sources for the discrepancy between the observations and general queueing theory, revealed that with the introduction of batches, the Gaussian shape of the distribution function of the number of samples in the system is obtained. In Fig. VI-10 the histograms are presented of the number of samples in a M/M/1 system, where the analyst waits, respectively does not wait for starting the measurements until a minimal batch size has been reached. Fig. VI-10a demonstrates that the probability function of k samples in the system $p(k)=z^k(1-z)$ fits the simulated histogram of a pure M/M/1 system well. Likewise that function fits well the histograms of the number of waiting samples in the laboratory model, run without restrictions for starting the measurements (Fig. VI-10c,d). The differences between these histograms and those obtained for the actual laboratory, with restrictions for starting measurements (Fig. VI-10 c,d and Fig. II-6) are apparent. The results indicate that the minimal batch size affects the variation coefficients of the number of waiting samples considerably (Table VI-4). The tight rule perform better than the loose rule in terms of the variance of the number of waiting samples.

Table VI-4

Effect of the minimal batch size on the variation coefficient of the delay and of the number of waiting samples.

	variation coefficient						
	number of waiting samples						overall delay
section	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.	Lab	M/M/1	lab
min. batch size							
1	0.71	0.63	0.72	0.52	0.18	0.79	1.24
max. batch	0.54	0.24	0.40	0.19	0.10	0.41	1.05

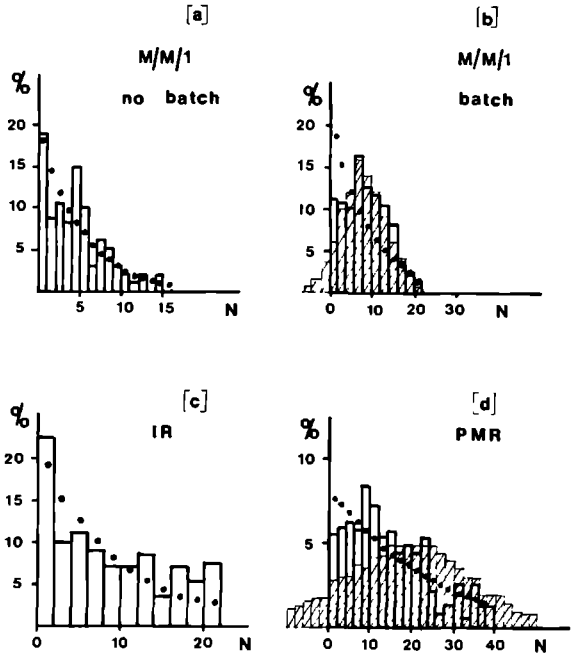


Fig VI-10: Simulated effect of batch analysis on the probability (%) of n samples in the system.
(a) M/M/1 system without batch analysis
0 best fitting exponential function $f(z)=z^n(1-z)$ with z=0.82 (D_{max}=0.066)
(b) M/M/1 system with batch analysis
0 best fitting exponential function $f(z)=z^n(1-z)$ with z=0.81 (D_{max}=0.148)
shaded figure: Gaussian distribution with the same parameters as the histogram (m=8.05, s=5.16) (D_{max}=0.060)
(c) model: simulation of the i.r. section without batch analysis
0 best fitting exponential function $f(z)=z^n(1-z)$ with z=0.81 (D_{max}=0.062)
(d) model: simulation of the p.m.r. section without batch analysis
0 best fitting exponential function $f(z)=z^n(1-z)$ with z=0.92 (D_{max}=0.074)
shaded figure: Gaussian function with the same parameters as the histogram (m=20.1, s=16.0) (D_{max}=0.100)

6. Strategies concerning the arrival of samples

In Chapter II it was demonstrated that the collection of samples into batches effects the delay of a M/M/1 system considerably, except when the batches arrive equidistantly at the laboaratory (e.g. once or twice per day) and the batches are Gaussian distributed with a relatively small variation coefficient [$C_r^2=\sigma_r^2/\bar{r}^2<1-\bar{r}(1-\rho)$]. For example, the equations presented in Table III-3 forecast an increase of the delay with a factor 3.1, resp. 2.6 and 5.4 for resp.

Poisson, constant and exponentially distributed batches of size $\bar{r}=7$ in a M/M/1 system with an utilization factor $\rho=0.6$ and batches arriving at constant intervals. The results of the experimental design presented in Table VI-5 indicate that the observed effects in the laboratory model are considerably smaller than calculated for a single M/M/1 system. Moreover the observed effect of the batchsizes of the input is much smaller than the effect of the minimal batchsize required for starting the measurements. Furthermore the probability density function of the batchsize of the input does not affect the mean delay significantly. This discrepancy between the behaviour of a single

Table VI-5

Effect of the mean batchsize ($\bar{r}$), probability density function of the batchsizes, and the minimal required batchsize before starting the measurements. Analysis of variance of a $2x3^2$ design.

factor levels:

(A) minimal batchsize: 1, maximal batchsize

(B) p.d.f. of the batchsize: constant, Poisson, exponential

(C) batchsize: 1, 2 batches/day, 1 batch/day

p.d.f. of the interarrival times of the batches: constant

source of variation	sum of squares	degrees of freedom	mean square	effect (days)
main effects				
A	7.16	1	7.16†	+1.26
B	0.55	2	0.27	
C	1.08	2	0.54*	+0.58
two factor interactions				
AxB	0.24	2	0.12	
AxC	1.25	2	0.62*	
BxC	2.30	4	0.57*	
residual	0.21	4	0.05	

*significant $1\%<P<5\%$

†significant $P<1\%$

M/M/1 system and that of such a system imbedded in the described network, is probably due to the fact that 25% of the traffic generated in the model is originating from inside the network (samples analyzed by several methods).

Because in the model no transfer time between the nodes is included, these samples arrive separately at the sections. Another aspect explaining this discrepancy is the fact that the samples of the batch arriving at the laboratory are divided over the sections according to the described decision rules. Consequently the mean batch sizes to the sections are relatively small (Table VI-6) and the probability density function of the batch size is disturbed.

Table VI-6
Mean batch size to the sections in the model (except inner transfers)

section	1 batch/day	2 batches/day
I.r.	2.	1.
P.m.r.	6.7	3.3
M.s.	1.1	0.6
^{13}C-n.m.r.	1.9	0.9

Arrivals of batches of constant size to the network do not lead to constant batches arriving at the sections. Other measures for laboratory performance, such as variance of flow time and longest flow time are practically not affected by the introduction of batch input. In addition, the findings that the tight rule, obliging an analyst to wait until a minimal batchsize is present in the laboratory, performs best in terms of the variation coefficient of the number of waiting samples are confirmed in terms of the variation coefficient of the delay (Table VI-4).

7. *The effect of the means and variation coefficients of the measurement- and interpretation time on the delay*

From theoretical considerations on a M/M/1 system and on a Er4/M/1 system in Chapter III, it was concluded that the effect of the means of the measurement and interpretation times should be much greater than the effect of their variation coefficients. An analysis of variance on a 2^4 design (four factors on two levels), executed on the i.r. section of the model confirmed this conclusion for more complex sytems. As the mean interpretation time of the i.r. spectra exceeds their measurement time by a factor 3 (Table IV-1), a greater effect is found for the mean interpretation time. A variation of the measurement time from $1.2\overline{MT}$ to $0.8\overline{MT}$, reduces the total analysis time with 10% while the same variation of the interpretation time reduces the total analysis time with 30%.

Table VI-7

Effect of the means and variation coefficients of the measurement (MT) and interpretation time (IT) on the delay in the i.r. section. Analysis of variance of a 2^4 design.

factor levels

(A) MT: $0.8\overline{MT}$, $1.2\overline{MT}$ ($\overline{MT}$ and $\overline{IT}$ are the values tabulated in Table IV-1)

(B) IT: $0.8\overline{IT}$, $1.2\overline{IT}$

(C) C^2_{MT}: 0.5, 2

(D) C^2_{IT}: 0.5, 2

source of variation	mean square	variance ratio	effect
A	19	6.7*	+2.2
B	239	85.3†	+7.7
C	16	5.7*	-2.0
D	0	-	
two factor interactions			
AxB	6	2.1	
BxC	9	3.2	
CxD	5	1.7	
AxD	0	-	
BxD	0	-	
residual	4	2.8	

†highly significant P<1%

*significant 1%<P<5%

The variance ratios tabulated in Table VI-7 show clearly that a considerable variation of the variation coefficient i.e. a reduction to 25% of the original value, affects the mean delay to the same extend as a reduction of the analysis time with 10% only. This demonstrates the greater sensitivity of the delay for the mean analysis time than for the variation coefficient of the analysis time. Since no interaction is found between the measurement time and interpetation time, the effect of both parameters in all sections can be studied separately. With this experiment the bottleneck of the system can be determined (Table VI-8). Obviously, the measurement time of the p.m.r. spectra is the greatest source of variation of the overall delay. This indicates that the availability of the p.m.r. instrument forms the bottleneck of the system. Although the dependencies of the delays on the sample flow to the laboratory are allmost equal in all sections (except m.s. section)

(Fig. VI-11), the delay in the p.m.r. section exhibits the strongest effect on the overall delay, because 80% of the samples passes through that section.

Table VI-8

The effect of the measurement time (MT) and interpretation time (IT) in the sections on the overall delay
Analysis of variance of a 2 x 2^4 design

factor levels
(A) i.r.: 0.9$\overline{MT}$(ir), 1.1$\overline{MT}$(ir)
(B) p.m.r.: 0.9$\overline{MT}$(pmr), 1.1$\overline{MT}$(pmr)
(C) m.s. : 0.9$\overline{MT}$(ms), 1.1$\overline{MT}$(ms)
(D) c.m.r.: 0.9$\overline{MT}$(cmr), 1.1$\overline{MT}$(cmr)
(E) 0.9$\overline{IT}$(ir), 1.1$\overline{IT}$(ir)
(F) 0.9$\overline{IT}$(pmr), 1.1$\overline{IT}$(pmr)
(G) 0.9$\overline{IT}$(ms), 1.1$\overline{IT}$(ms)
(H) 0.9$\overline{IT}$(cmr), 1.1$\overline{IT}$(cmr)

source of variation	mean square	variance ratio	effect	source of variation	mean square	variance ratio	effect
A	0.11	1.9		E	1.22	22.6†	0.55
B	5.11	86.6†	1.13	F	1.93	35.7†	0.70
C	0	-		G	0.47	8.7*	0.47
D	1.03	7.4†	0.51	H	1.05	19.5†	0.51
two factor interactions				two factor interactions			
AxB	0	-		ExF	0.31	5.7	
BxC	0.07	1.2		FxG	0	-	
CxD	0.01	0.17		GxH	0	-	
AxC	0.01	0.17		ExG	0.12	2.2	
AxD	0	-		ExH	0.03	0.5	
variance	0.059			variance	0.054		

† highly significant P<1%
* significant 1%<P<5%

8. Sensitivity for other activities

As expected from the theoretical considerations outlined in Chapter II, the 'overhead' of the analysts and failures of the instruments influence the overall mean delay considerably. In the model other activities are started and executed also when samples or spectra are present in the laboratory. The delay in the laboratory without other activities and without failures of the instruments was 3.9 days, with a maximum delay of 24 days. Evidently, the utilization factors of the analysts in the laboratory model remained unchanged. A subsequent introduction of a minimal batchsize of one sample per analysis enhances the delay further with 0.9 day, with a maximal delay of 10 days. This result confirms the conclusion from queueing theory (Table III-2)

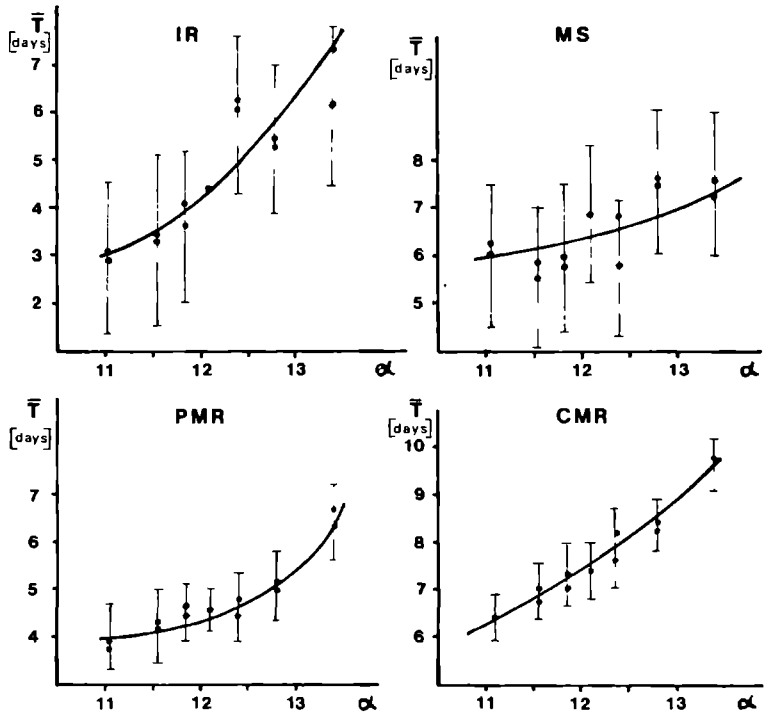


Fig. VI-11: Mean delay ($\bar{T}$) as a function of the input density to the laboratory (α) (samples/day) for the various sections.
I 95% confidence interval.

i.e. the delay in the laboratory under investigation cannot be explained by the statistical parameters of the measurement- and interpretation time. Clearly, the delay is strongly influenced by the schedule of the other activities and the minimal batch size per analysis. One should conclude that the control of the other activities would be the key to reduce the delay. As mentioned before, three categories of other activities are distinguished in the model: absence of the analysts (5%), coffee breaks (6%) and non-analyzing activities, such as administration, research etc.. However, permitting the non-analyzing activities only when less than 10 samples wait for analysis in the laboratory, the overall delay is reduced with 20% only.

9. Sample sequence within a group of samples with the same priority.

In Ch. V, a higher correlation was mentioned between the number of samples in the section at the arrival of a sample and its delay in the model as compared to the actual situation. This discrepancy seems an indication for an invalid model assumption that the samples of the various priority groups are analyzed in a FIFO sequence. Therefore a simulation experiment was executed with a random analysis sequence within the various priority groups. Comparison of Tables V-10 and VI-9 reveals that with the introduction of a random sequence, the correlation between the delay and number of waiting samples has decreased, whereas in the p.m.r. section (receiving 80% of all samples), the correlation is not significantly different from zero. The effect of randomization of the sequence is visualized in Fig VI-12. The longest delay before a sample will be analyzed (95% probability) is hardly correlated (ρ=0.23) with the number of waiting samples in the system at its arrival (Fig II-10). Similarly, the model run with a random sequence within the groups shows no correlation (Fig. VI-12a). In contrast, a FIFO sequence in the model increases the correlation considerably (ρ=0.73) allowing a reasonable forecast of the maximal delay at the moment of arrival of the sample at the laboratory (Fig. VI-12b). The fact that the laboratory uses a random sequence instead of the presumed FIFO sequence has no consequences for the validated results of the model, as both sequences yield the same delay [KL75].

Table VI-9
Maximal correlation between the number of samples (x) in a section and the delay (y) of the samples arriving at the laboratory (model): Random sequence

	i.r.	p.m.r.	m.s.	^{13}C-n.m.r.
ϕ_{xy}	0.43 (-1)	0	0.49 (-6)	0.48 (-30)
95% conf. interval	0.37	-	0.25	0.26
residual variance	0.81	1	0.76	0.77

When the maximal delay of a sample can be forecasted with 95% certainity from the number of samples present in the system at the moment of its arrival, then, the delay can be kept within certain limits with a given probability, by applying a threshold control of the number of samples in the system. Müskens [MU78] described a threshold control system, where the time lag is calculated

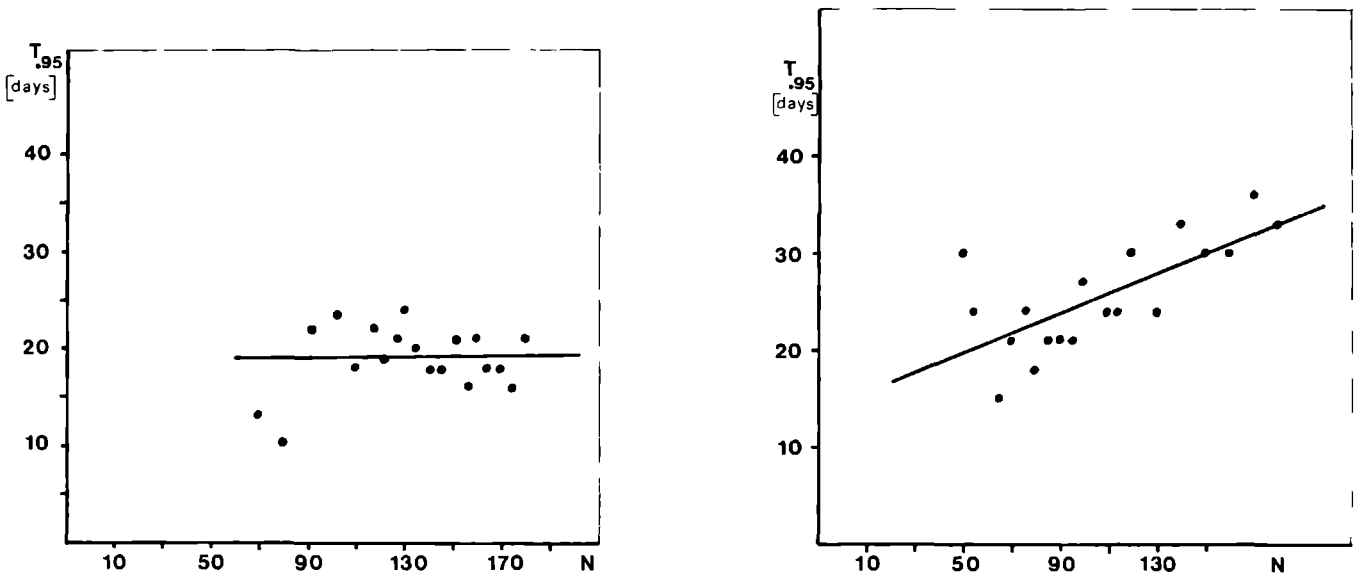


Fig. VI-12 a. Simulated conditional 0.95 probability limit ($T_{.95}$) for a sample to be analyzed within a delay (T) as a function of the number (N) of samples in the laboratory: Random sequence (ρ=0.10)
b. FIFO discipline (ρ=0.72)

in dependence on the observed value of the time series (e.g. the number of samples in the system), after which the system should be measured again, in order to obtain a certainty of P% that a given threshold value will not be exceeded. This statistical forecast is based on the probability density function of the queue size, and on an AR(1) model, describing the underlying structure of the queue size. At any time t, the queue level at a time τ later can be forecasted by applying Eqn. VI-6.

$$N_f(t+\tau) = N(t)\phi_1^{\tau} \qquad \text{VI-6}$$

where, $N_f(t+\tau)$ is the forecasted queue level from the mean level at a time τ later, $N(t)$ is the queue level from the mean level at time t and ϕ_1 is the autocorrelation at $\tau=1$ of the number of waiting samples. According to Müskens [MU78], the prediction error, using the autocorrelation function as predictor equals: $\sigma(t+\tau)=\sigma_N\sqrt{(1-\phi_1^2)}$. The probability ($\alpha$) at a time t that a threshold value (N_{th}) will be exceeded at a time τ later equals:
$u(\alpha) = [N_{th} - N(t)\phi_1^{\tau}]/[\sigma_N\sqrt{(1-\phi_1^{2\tau})}]$ with $u(\alpha)$ the excentricity of the normal distribution, giving the requested probability (α). Fig VI-13 depicts the time interval after which the number of samples should be evaluated again in the p.m.r. section, in dependence of the threshold value, the accepted risk (α) to exceed that value, and the actually observed number of samples. From the simulated relationship between the number of samples in the p.m.r. section and the maximal delay (95% probability), plotted in Fig. VI-14, it is clear that when a maximal delay of 15 days in that section is desired, the number of samples in that section should not exceed 50 (units). From Fig. VI-13 it follows that an accepted risk of 5% for the threshold value N_1=50 to be exceeded, the state of the p.m.r. section should be evaluated within 1 to 2 weeks, when 25 to 10 samples are waiting. When over 30 samples are waiting the state of the section should be surveyed every day.

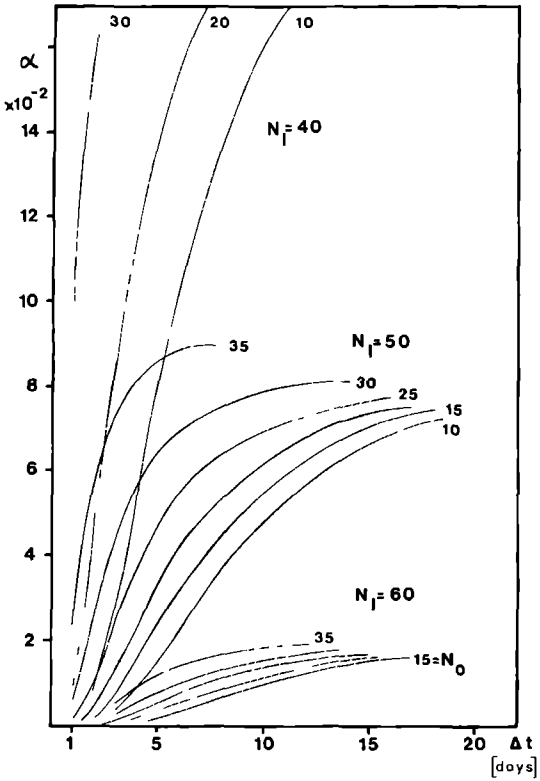


Fig. VI-13: Threshold control of the queue sizes in the p.m.r. section: the time lag (Δt) (days) after which a risk (α) of the exceeding of a threshold value N_I is obtained, as a function of the actual number of waiting samples (N_o).

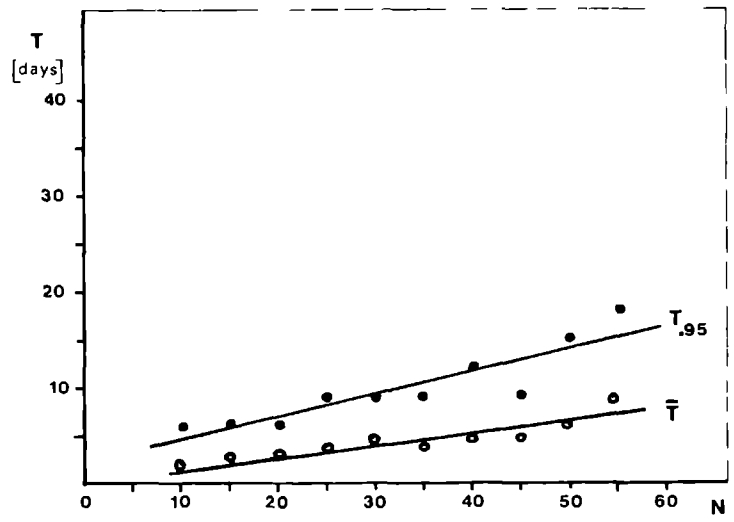


Fig. VI-14: P.m.r. section:
O simulated conditional mean delay ($\bar{T}$) as a function of the number (N) of samples in that section.
● Conditional 0.95 probability limit ($T_{.95}$) for a sample to be analyzed within a delay (T) as a function of the number (N) of samples in that section (FIFO sequence)

Conclusions.

The relative impact of the decision rules was found to be dependent on the measure of performance considered: i.e. mean delay or variation coefficient of the delay. One may generally conclude that the simulated effects with the laboratory model are less pronounced than predicted for single queueing systems. For example, the effect of batch input is considerably smaller than expected from calculations on M/M/1 systems. As expected, other activities, executed while samples are waiting, determine largely the observed delays in the laboratory. Likely, the minimal batch size of samples, which are measured simultaneously, is a relatively important factor for the variance of the delay, and had pronounced effects on the mean delay and probability density function of the number of waiting samples in the system. The Gaussian shape of that number, as observed in the laboratory can be explained by the introduction of batch analysis in the model. Operation without batches performs better, even when the measurement times can be reduced with 20% by batch analysis. Evidently, the workload of the laboratory, which is the product of the mean number of visited sections and the number of arrivals per day, affects the delay considerably. All decision rules that increase the number of visited sections affect the delay negatively. An illustrative example is the effect of the reduction of the mean analysis time by the introduction of a maximal analysis time which is completely surpassed by the increased number of visited sections. The effect of an increase of the workload is different for the various groups of samples with different flow and priority. The total delay is reduced when absolute priority is attributed to easy samples. This reduction is relatively insensitive for the limiting analysis time of 'easy samples'. Likely, this reduction is insensitive for the estimation error of the analysis time of the samples. The performance of the laboratory model is enhanced if the probabilities that the various sections will give the requested information are considered along with the state of the laboratory in order to route the sample to some section. The effect is more pronounced when these probabilities can be estimated less accurate. As expected from the theoretical outline of M/M/1 systems, the system is more sensitive for the mean analysis time than for the variation coefficient of the analysis time. The transition from a centralized to a decentralized organization is only advantageous when all analysts are fully qualified for all methods.

The ultimate delay for an analytical result to be available, is hardly correlated with the number of samples present in the actual laboratory. This is probably due to a random sequence dispach of the samples with the same priority. By the change of the random sequence to a FIFO sequence, however, a reasonable correlation is obtained, which agrees with the conditional probability density function of a M/M/1 system. Combining the AR(1) model of the number of samples in the system, with the conditional probability function of the delay, a threshold control system can be created.

BR59 S.H. Brooks, Opns. Res. 7 (1959), 430.

CO67 R.W. Conway, W.L. Maxwell, L.W. Miller, Theory of Scheduling, Addison Wesley Pu. Comp., Reading Mass. 1967.

DA71 O.L. Davies, ed. Design and Analysis of Industrial Experiments, Oliver and Boyd, Edinburgh, 1971.

DA75 O.L. Davies, Talanta, 22 (1975) 371.

DE73 S.N. Deming, S.L. Morgan, Anal. Chem., 45 (1973) 278A.

KL75 L. Kleinrock, Queueing Systems, vol. I., Theory, J. Wiley & Sons, New York, 1975.

MU78 P.J.W.M. Müskens, Dutch Thesis, Nijmegen, 1978.

GENERAL REMARKS AND CONCLUSIONS

Apart from the technical description of the model, some general considerations are given on mathematical modelling and decision making in analytical chemistry. For the planning of the simulation experiments a very time-consuming process of formulation of the problem up to the design of the simulation experiments and analysis of the simulation data is always necessary. One may reasonably doubt whether the efforts for building the model are worthwile in view of beneficial results which are expected.

Biswas [BI75] gave an excellent view on this topic. The primary role of a decision-maker is to make right decisions, which may be defined as decisions made on the basis of perfect knowledge (Churchman 1961, Biswas 1971). Since knowledge is always imperfect, the best decision should be aimed at, based on the available information. However, in order to know which information is needed for decision-making, some insight in the (laboratory)system should be available. Therefore, the process of information gathering is imbedded in a vicious circle with the outcome of past decisions. With a better understanding of the system, more relevant data can be collected. As a result, frequently, (also in presented research) the modelling and data collection process proceeds in parallel. During this modelling-cycle, the output of the crude model is checked against observations in the laboratory. As long as the output of the model does not match the observations, the model is refined. During that refinement process, new observations may be necessary, but in the mean

time partial results become already available.
In our particular case, the laboratory under investigation was informed during that period, on their applied priority rule between the two principal sources of samples that was not in accordance with their aims. Moreover, their attention was turned to the fact that an augmentation of priority of the samples that visited more sections, should avoid the very long delays and decrease the variation coefficient of the delay.
The ultime question whether the efforts to build the model are useful, depends on the profit obtained by avoiding wrong decisions. A necessary condition, however, for a useful model is that it has enough credibility with the policy makers. A major reason for a lack of credibility can be the lack of user involvement in the model development process [BI75].
Therefore, during the modelling period, intermediate results were communicated and discussed in several plenary meetings with all laboratory personnel and staff. Such interactions proved to be mutually beneficial.
One of the features of the decision-making process is, that only a limited number of policy alternatives are considered for any decision. These alternatives generally differ incrementally from existing policies [BI75] which means that the advances are made in small steps. Therefore, only strategies and policies were simulated which did not need a drastic change or reorganization of the current policies of the laboratory. As a result, solutions of the model remain acceptable to policy makers.
The follow up of the research presented here, is a half-yearly updating of the model with the most recent observations, combined with a control of the main characteristics of the laboratory. Voluntarily, the model can be used to forecast the effect of alternatives, proposed by the decision-maker(s). This updating process is necessary as understanding of the process being modelled improves, and otherwise the model tends to become out-of-date.
The flexibility of the model, presented here, is such that the number of facilities (instruments and personnel) are very easily adapted. Likewise, the fixed characteristics defining the statistical properties of the sample flow, the analytical procedures and of the non-analyzing activities may easily be varied.
In general, it is worthwhile not to build generalized all-purpose models. These models are expensive to develop, difficult to control, and have large data requirements [BI75].

Based on this point of view, for every laboratory with another structure than the laboratory model presented here, a new model should be built.
Finally, as a general conclusion, I subscribe the statement of Biswas [BI75] that:
"The issue is very definitely on the side of having a model, even a crude one, against having no model at all".

References

BI75 A.K. Biswas, Ecol. Modelling, 1, (1975) 31.

SUMMARY

In Analytical Chemistry a research topic is the development of optimal strategies for obtaining analytical information.
These optimal strategies can be derived by the application of mathematical models, a commonly accepted method in operations research, a branch of applied mathematics.
This thesis describes the results of an investigation of the delay of samples in a laboratory for structural analysis. These results have been obtained by the application of queueing theory and digital simulation. Until now little attention was paid to the quantification of the effects of various factors on the delay in an analytical laboratory. The introductory chapter illustrates the importance of the study of delays with a discussion of the interactions between the analytical laboratory and the processes the samples originate from.
In Chapter II, it is demonstrated that a laboratory for structural analysis (Philips Duphar B.V., Weesp) can be represented by a network of queues of samples, spectra and results. The 4 sections (i.r., p.m.r., m.s. and C^{13}-n.m.r.) are represented as the 4 nodes of that network.
Many properties of the laboratory under investigation are characteristic for 'open' networks, i.e. networks where (i) no correlation exists between the number of samples in the various nodes (ii) the sample stream towards a node is independent of the state in the node and (iii) the input- and output sample

flows are mutually independent and not autocorrelated. The histograms of the number of arrivals and departures per day are completely different. The histograms of the delays in the sections were fitted to a K = ? Erlangian probability density function; the histograms of the number of arrivals to a Poisson distribution. The parameters of the distributions describing the analysis time are determined. The results of calculations on various theoretical models based upon queueing theory are presented in Chapter III. Although these models are too simple to provide exact results for the complex laboratory under investigation, a reasonably good forecast of the effect of some variables can be obtained. Clearly, the utilization factors of personnel and instruments have a strong effect on the delay. The delay tends to infinity when the utilization factor approaches unity.

The 'overhead' of the personnel has the same effect. When the sum of overhead and utilization factor approaches to unity, the delay becomes infinite, even for low utilization factors.

Various groups of samples, with a different priority have a different delay. The inclusion of a cost factor attributed to waiting periods, enables to optimize the system.

A digital simulation model of the investigated laboratory is presented in Chapter IV. With this model the forecasted effects, presented in Chapter III, are verified and quantified. Moreover, laboratory systems can be processed for which no simple theoretical models exist e.g. systems with state dependent decisions.

Different strategies for sample priorities, sample routing, allocation of personnel and termination of the analysis are described. The usefulness of various statistical methods for model validation is examined.

The simulation of the actual situation in the Philips Duphar laboratory is presented in Chapter V. The actual number of arrivals per day to the laboratory and to the individual sections could be adequately described by generating a Poisson distributed sample stream for each sample origin.

The model not only describes the actual situation, but forecasts correctly the effect of modifications in the operation and organization which were implemented in a later period of the real laboratory.

No significant differences were observed between the model and the actual laboratory as far as correlations between various variables, the mean delays and the mean number of samples in the various section are concerned. The time series of the number of waiting samples in the model can be described by a

first order autoregressive model. The frequency and magnitude of the fluctuations of the number of samples in model and reality are not significantly different.

The variation coefficient of the delay could be minimized by assigning absolute priority to the samples that visited the most sections in the laboratory.

The application of 'experimental design' techniques, described in Chapter VI, indicated that the delay of the samples is more sensitive to the mean value of the analysis time than to its variation coefficient. Consequently, it is better to modify an analytical procedure in such a way that the mean analysis time is reduced, rather than the variation coefficient of the analysis time.

A transformation of a decentralized organization (in which the personnel has experience with one analytical method only) to a centralized organization, where each analyst can operate all methods is only of advantage when the instrumentation is not a bottleneck, and the analysts have enough expertise to do the analysis beyond their own specialism. The mean delay of the laboratory can be decreased with 20% by assigning absolute priority to the samples with an estimated high interpretation time (> 3 times the mean interpretation time). Furthermore, a 24 hour service can be established for some groups of samples (e.g. samples with very short interpretation times). The delay is very sensitive for variations in the density of the sample stream when no appropriate organizational measures are taken. Especially the p.m.r. section has a high saturation degree. Balancing the probabilities that the various sections can furnish the requested structure against the workload reduces the mean delay with approximately 15%.

A batch sample imput to the system and a batch measurement of the samples influence the delay adversely (factor 1,5 to 2), when no overhead reduction is obtained.

When the non-analyzing activities of the personnel are limited to the non-busy periods of the laboratory ($N < 20$), only a slight reduction (< 10%) of the delay will be obtained.

An extension of the personnel without an extension of the instruments will have no effect.

The priority between samples of various origins affects mainly the smallest group of samples. Therefore, a periodic control of the delays of the various groups of samples is advocated.

SAMENVATTING

Naast de ontwikkeling en verfijning van analysemethoden, is het tevens van belang strategieën te ontwerpen om analytische informatie op een optimale wijze te verkrijgen. Hiertoe zijn door enkele onderzoekers modellen gebruikt uit het vakgebied der Operations Research. Dit proefschrift omvat het onderzoek van de doorlooptijd van monsters in een laboratorium voor struktuuranalyse met behulp van wachttijdentheorie en digitale simulatie. Binnen het vakgebied der analytische chemie hebben de faktoren die deze wachttijd beïnvloeden tot nog toe weinig aandacht gekregen.
In het inleidende hoofdstuk wordt het belang van de studie van doorlooptijden van monsters aangetoond, door in te gaan op het systeem opdrachtgever-analytisch laboratorium. Daarnaast wordt verduidelijkt waarom modelvorming de enige mogelijkheid is om laboratoriumsystemen te onderzoeken.
Hoofdstuk II toont aan dat een laboratorium voor struktuuranalyse (Philips Duphar B.V., Weesp), voorgesteld kan worden als een netwerk van wachtrijen van monsters, spektra en analyseresultaten. De 4 afdelingen (IR, PMR, MS en CMR-spektrometrie) vormen de 4 knooppunten van dit netwerk. In vele opzichten komen de eigenschappen van het onderzochte netwerk overeen met deze van een 'open' netwerk. Er is namelijk geen correlatie gevonden tussen het aantal aanwezige monsters in de verschillende knooppunten van het netwerk. Het aanbod naar de knooppunten is onafhankelijk van de toestand in het knooppunt zelf. Daarnaast zijn de uitgang en ingang van het netwerk onafhankelijk van elkaar

en niet geautocorreleerd. De histogrammen van het aantal binnenkomende en vertrekkende monsters per dag zijn totaal verschillend gebleken. De histogrammen van de doorlooptijden in de knooppunten worden het best benaderd door een k = 2 Erlang kansdichtheidsverdeling. De histogrammen van het aantal binnenkomende monsters door een Poisson verdeling. De parameters van de verdelingen van aanbod en analysetijden werden bepaald.
Hoofdstuk III geeft de resultaten weer van een literatuuronderzoek naar enkele theoretische modellen uit de wachttijdentheorie. Niettegenstaande deze modellen tè eenvoudig zijn om een complex systeem als een laboratorium exact te kunnen beschrijven, kan toch een redelijke schatting verkregen worden van de gevoeligheid van het laboratorium voor een aantal faktoren. De belangrijkste faktoren zijn ondermeer: de bezettingsgraad van het personeel en instrumentarium. De doorlooptijd wordt oneindig groot bij een bezettingsgraad naderend tot 1. Van even grote invloed is de grootte van de 'overhead' van het personeel. Indien de som van 'overhead' en bezettingsgraad tot 1 nadert, wordt de doorlooptijd eveneens oneindig groot, zelfs bij lage bezettingsgraden. Door het aanleggen van prioriteitsregels tussen verschillende groepen monsters, kunnen deze monsters sterk verschillende doorlooptijden verkrijgen.
Hoofdstuk IV beschrijft een digitaal simulatiemodel van een laboratorium voor structuuranalyse. Met dit model kunnen de in hoofdstuk III voorspelde effecten geverifieerd en gekwantificeerd worden. Daarenboven wordt het mogelijk om situaties door te rekenen waarvoor geen vereenvoudigde theoretische modellen beschikbaar zijn zoals b.v. het invoeren van toestandsafhankelijke beslissingen. Mogelijke strategieën voor monsterprioriteit, monsterrouting, personeelsallocatie en afbreken van analyses zijn beschreven. Statistische methoden om het model te valideren zijn op hun bruikbaarheid getoetst.
In hoofdstuk V wordt de simulatie van de actuele situatie bij Philips-Duphar besproken. Het monsteraanbod naar het laboratorium en de verschillende afdelingen kon beschreven worden door één Poisson verdeelde monsterstroom per opdrachtgever. Het model beschreef niet alleen in voldoende mate de situatie die gebruikt werd om het model op te stellen, maar voorspelde tevens het gedrag van het laboratorium een jaar vooruit. Dit betekent dat zowel de gevonden correlaties tussen de verschillende variabelen, als de gemiddelde doorlooptijden en het aantal wachtenden in de verschillende afdelingen in model en reële situatie, niet significant verschillend zijn. Een autoregressief model van de eerste orde beschrijft het gedrag van het aantal wachtende monsters over de tijd in het simulatiemodel. De snelheid en grootte van de fluctuaties

van het aantal wachtende monsters in model en realiteit zijn dus niet significant verschillend.
Gebleken is dat de variatiecoëfficient van de doorlooptijd van de monsters minimaal wordt indien in de afdelingen absolute prioriteit wordt verleend aan de monsters die intern zijn doorgestuurd.
Experimenten met het model, ondermeer via 'experimental design' technieken opgezet (hfst. VI), tonen aan dat de doorlooptijd van de monsters gevoeliger is voor de gemiddelde analysetijd. Dit betekent dat bij standaardisatie van analysemethoden eerder gezocht moet worden naar een verlaging van de gemiddelde waarde dan van de spreiding van de analysetijd. De omschakeling van een gedecentraliseerde organisatie (waar het personeel slechts ervaring heeft met 1 analysemethode) naar een gecentraliseerde organisatie heeft slechts zin indien de apparatuur geen knelpunt vormt en het personeel voldoende ervaring heeft in de methoden buiten hun specialiteit. Prioriteit op basis van een geschatte interpretatietijd, waarbij monsters met een hoge geschatte waarde (> 3x gemiddelde waarde) absolute voorrang moeten geven aan alle andere monsters, kan de gemiddelde doorlooptijd met 20% doen afnemen. Het is verder mogelijk een 24 uurs service in te voeren voor een bepaalde groep monsters (vereisen nauwelijks enige interpretatietijd). Bij gelijkblijvende laboratorium organisatie is de doorlooptijd sterk afhankelijk van verhoging van het aanbod (50% toename van de doorlooptijd bij 15% toename van het aanbod). Vooral de PMR afdeling vormt hierbij het knelpunt. Een afweging van de waarschijnlijkheid dat de verschillende afdelingen het analyseprobleem kunnen oplossen en hun bezettingsgraad reduceert de doorlooptijd ongeveer met 15%. Het batchgewijs aanbieden van de monsters heeft over het algemeen een verhogende invloed op de doorlooptijd (faktor 1,5 tot 2), indien geen besparing op de overhead verkregen wordt door een batchgewijze analyse. Het beperken van de overige werkzaamheden (b.v. eigen onderzoek) van het personeel tot deze periodes waarop slechts weinig monsters in het laboratorium aanwezig zijn (N < 20) heeft slechts een beperkte invloed op de doorlooptijd (< 10%). Uitbreiding van het personeel zonder uitbreiding van het instrumentarium zal slechts geringe invloed hebben. De gehanteerde prioriteitsregels tussen monsters van verschillende opdrachtgevers zullen vooral de monsters met een relatief laag aanbod sterk beïnvloeden. Geregelde controle van de doorlooptijden van de verschillende groepen monsters is dus noodzakelijk.

CURRICULUM VITAE

B.G.M. Vandeginste

12 november 1943 : geboren te Kortrijk;

juni 1962 : eindexamen Oude Humanioria (Latijnse Wiskunde) aan het St. Amandscollege te Kortrijk;

juli 1965 : diploma van technisch ingenieur, behaald aan het Hoger Technisch Instituut te Oostende;

van september 1965 tot

april 1973 : in dienst bij de afdeling Chemische Technologie van de Technische Hogeschool te Delft;

september 1965 : begin studie aan de Technische Hogeschool te Delft;

25 januari 1973 : diploma scheikundig ingenieur (met lof);

april 1973 : wetenschappelijk medewerker aan de Katholieke Universiteit Nijmegen, Faculteit der Wiskunde en Natuurwetenschappen, Afdeling Analytische Chemie, tevens begin van dit onderzoek.

STELLINGEN

I

De door Eckschlager voorgestelde werkwijze voor de berekening van de kosten per bit geleverde informatie en voor de optimalisatie van de geleverde informatie per tijdseenheid, houdt ten onrechte geen rekening met de organisatie van het laboratorium.
K. Eckschlager, Anal. Chem., 49 (1977) 1265.

II

De bewering van Liteanu en Panovici dat analytische systemen met een significant van nul verschillende waarde voor de autocorrelatiefunctie bij $\tau=1$, onstabiel zijn, is onjuist.
C. Liteanu, I.I. Panovici, Talanta 24 (1977) 196.
C. Liteanu, E. Hopîrtean, Z. Anal. Chem. 288 (1977) 59.

III

De kwalitatieve resultaten van Allen en McMecking omtrent de detectiegrens van twee overlappende Gausse banden m.b.v. de tweede afgeleide waren reeds eerder kwantitatief afgeleid.
G.C. Allen, R.F. McMecking, Anal. Chim. Acta., CTO 103 (1978) 73.
B.G.M. Vandeginste, L. de Galan, Anal. Chem., 47 (1975) 2124.

IV

Tattershall vermeldt niet aan welke eisen i.r. spectra moeten voldoen om, door digitaal aftrekken van spectra, afzonderlijke componenten in een mengsel te kunnen identificeren.
B.W. Tattershall, Anal. Chem., 49 (1977) 772.

V

Het feit dat identieke waarnemingen van de NaCl interferentie op het Cu signaal verkregen door middel van vlamloze atomaire absorptie, tot tegengestelde interpretaties kunnen leiden (occlusie en verdamping), duidt erop dat de resultaten van de studie van het interferentie mechanisme nog steeds een hypothetisch karakter dragen.

D.J. Churella, T.R. Copeland, Anal. Chem., 50 (1978) 309.
E.J. Crobih, D.P. Matousek, Anal. Chem., 50 (1978) 2.

VI

Daar Fujiwara et al. bij hun studie van de ruis in analytische vlammen, alleen het amplitude van de ruis hebben onderzocht, is hun evaluatie van spektrale ruis zeer onvolledig.
K. Fujiwara, A.H. Ullman, J.D. Bradshaw, B.D. Polland, J.D. Winefordner, Spectrochim. acta. 34B (1979) 137.

VII

Het wereldmodel van Jørgensen voorspelt een verviervoudiging van de wereldbevolking en een verzevenvoudiging van de totale wereldproductie (bruto nationaal product BNP), voor het jaar 2030. Daar deze voorspelling slechts gebaseerd is op een extrapolatie van het BNP en de gevonden correlatie tussen bevolkingsgroei en BNP, is zijn conclusie dat een te ver doorgevoerde ontwikkeling van de geïndustrialiseerde wereld hiervan de oorzaak is, ongegrond.
S.E. Jørgensen, Ecol. Modelling 1 (1975) 199.

VIII

De ontwikkeling van de analytische scheikunde in Nederland zou ermee gebaat zijn indien men goede nota nam van de uitspraak van het Department of Energy (U.S.A.), dat een eerste prioriteit dient gegeven te worden o.a. aan de studie van de analytische chemie als systeem.
H.C. Laitinen, Anal. Chem., 51 (1979) 785.

IX

Het gelijkstellen van de interne spanningen tussen Vlamingen en Walen met deze tussen katholieken en protestanten in Noord-Ierland in een veel gebruikte Nederlandse schoolatlas, komt een juiste beeldvorming in Nederland van de Belgische politieke situatie niet ten goede.
De grote bosatlas, Wolters-Noordhoff, Groningen (1976), pag. 93.

B.G.M. Vandeginste

Nijmegen, 11/1/1980