ARAB REPUBLIC OF EGYPT

THE INSTITUTE OF NATIONAL PLANNING

Memo (1322)

(1322)

EXPERIMENTAL OPTIMIZATION BY SIMULATION TECHNIQUE

BY

MOTAZ KHORSHID

AND

Alain CHECROUN

March 1982

CAIRO
SALAH SALEM St.-NASR CITY

The purpose of this paper is twofold , (i) to propose an integrated framework for studying management systems by simulation and (ii) to evaluate the possibility of using Simulation as an experimental optimization technique.

A job shop simulation model , which can be used to test both materials handling and dispatching rules , was developed in order to demonstrate the applicability of the proposed procedures* .

---

* This paper is published in the proceedings of the international congress on applied System research and cybernetics - pergamon press - New york 1981.

1 - INTRODUCTION

The attractive features that computer simulation offers, have encouraged operations researchers and statisticians to improve its practice through the use of various statistical techniques to analyse the obtained results [11,12,13,16,17,18,19,23].

The commun conclusion of these studies indicate the need to consider two main problems when investigating simulated systems. First, the particular circumstances of simulation experiments that may lead to misinterpretation of results(1). Second, the difficulty to achieve the assumptions of the statistical theory, as independence, normality, and homogeneity of variances. So, either we hope that the selected techniques are not influenced by assumptions violation, or we manipulate simulation runs to match them.

The purpose of this research is two fold. First, to propose an integrated framework for studying complex systems by simulation. Second, to evaluate the possibility of using simulation to find the optimum solution. In order to demonstrate the applicability of the proposed procedures, we developed a model that can be used to test both materials handling and dispatching rules in a job shop production system.

As any statistical investigation, simulation user should consider the following steps when developing an experimental strategy:-

i) The choice of a sampling plan that specify how each test run is to be executed and how to determine simulation run length.

ii) The development of an experimental design that will yield the desired information.

iii) The selection of a data analysis technique in order to reach some conclusion about the simulated system.

---

(1) A list of these circumstances can be found in Fishman (11)P.262

In section 2, we propose an integrated mathematical base for studying management systems by simulation. The proposed optimization techniques, are presented in section 3. In section 4, we explain briefly the simulation model that will be the setting of the study. Finally, the design and analysis of two experiments are discussed in section 5.

2 - MATHEMATICAL FORMULATION

In many simulation models, the process of interest apears as a stochastic process[1].

$$\{ Y(t); -\infty \leq t \leq \infty \} \qquad (1)$$

Since our research concerne discrete event digital simulations, we assume that during a small interval of time, the process shows little, if any change so that observing $Y(t)$ at periodic intervals result in no loss of information. In a more detailed manner, the sequence,

$$\{ Y_t , t = 1,2\ldots, \infty \} \qquad (2)$$

corresponds to the process $\{Y(t)\}$ at all integer values of the index t.

The index t may be the time; for example Y_t may define the number of jobs in a production system at the instant t of simulation. It may simply denote order, for example Y_t may represent the waiting time for the t^{th} customer to receive service in a queueing model.

In order to study several processes of interest, generated by different environmental conditions or input specifications, we should aquire a quantitative characterization of each of them.

---

(1) we will consider only the stochastic simulation models as most management systems inevitably appear random to some degree in nature.

The mean of the process;

$$u = E\left[Y_t\right] \tag{3}$$

serves generally as the mathematical descriptor, and by definition, the variance[1];

$$\mathrm{Var}\,(Y_t) = E\left[(Y_t - u)^2\right] \tag{4}$$

and the autocovariance function,

$$R_s = E\left[(Y_t - u)(Y_{t+s}-u)\right];\ s = 1,2,\ldots, \tag{5}$$

Then regardless of the experimental objective and the type of simulated model, we should define a procedure for estimating "u" and for determining the accuracy of this estimator, i;e to select a sampling plan.

2.1 SAMPLING PLAN:

In simulation experiments sample size can be increased by prolonging simulation run, or by conducting separate runs. Consequently, there is two sampling plans, the autocorrelated observations and the replicated runs.

i) Autocorrelated observations:

Let $\{Y_t, t \in n\}$ a time series of length n observed during the simulation run, the mean u can be estimated by:

$$\bar{Y} = n^{-1} \sum_{t=1}^{n} Y_t \tag{6}$$

where $\bar{Y}$ is called "simular response".

In order to determine the accuracy of $\bar{Y}$, we need to estimate its variance 12 :

$$\mathrm{Var}\,(\bar{Y}) = n^{-1}\left[\mathrm{Var}\,(Y_t) + 2\sum_{s=1}^{n-1}(1-s/n)\,R_s\right] \tag{7}$$

---

(1) Assuming $\left[Y_t\right]$ a covariance stationary sequence [11,18]

Where $Var(Y_t)$ and R_s are defined by (4),(5) and can be estimated by the formulas [18] :-

$$\hat{Var}(Y_t) = (n-1)^{-1} \sum_{t=1}^{n} (Y_t - \bar{Y})^2 \qquad (8)$$

$$\hat{R}_s = (n-s)^{-1} \sum_{t=1}^{n-s} (Y_t - \bar{Y})(Y_{t+s} - \bar{Y}); \qquad (9)$$

$$s = 1,2,..,(n-1)$$

ii) Replicated runs :-

We can generate independent observations in simulation experiment by repeating the run using different random numbers. Let

k = number of runs
n' = number of observations per run
n = total number of observations (=n'k)
$Y_{t,j}$ = the observation generated at time t of run j.

The mean of each run is defined by:

$$Y_{n',j} = n'^{-1} \sum_{t=1}^{n'} Y_{t,j} \quad ; \quad j = 1,2,\ldots,k \qquad (10)$$

and then simular response is calculated by:

$$\bar{Y} = k^{-1} \sum_{j=1}^{k} Y_{n',j} \qquad (11)$$

Since the sequence $\{ Y_{n',j} ; j = 1,2,\ldots,k \}$ consist of k independent observations, $Var(\bar{Y})$ is given by:

$$Var(\bar{Y}) = Var(Y_{n',j})/k \qquad (12)$$

and can be estimated by:

$$\hat{Var}(\bar{Y}) = S^2/k \qquad (13)$$

where $S^2 = (k-1)^{-1} \sum_{j=1}^{k} (Y_{n',j} - \bar{Y})^2$

2.2 SIMULAR RESPONSE FUNCTION:

When experimenter select a sampling plan, he can proceed to the study of the response Y as a function of the environmental conditions or the experimental factors $(x_1, x_2, \ldots, x_p)$. Factors are categorized as qualitative and quantitative. Examples of qualitative factors are policy specifications, such that alternative dispatching rules in job shop, or discrete environmental conditions. Quantitative factors are examplified by input parameters that can usually be thought as continious variates.

Since the stochastic features are spowned in the simulation model by incorporating the random number seed as an integrated part of input specifications, the response $\bar{Y}$ becomes a random variable because it is a transformation not only of the experimental factors $(x_1, x_2, \ldots, x_p)$, but also of the randomly selected seed " r " [1] This relation is defined by:

$$\bar{Y} = \phi (x_1, x_2, \ldots, x_p; r) = \phi (\vec{x}, r) \qquad (14)$$

Although the random number seed " r " may be conceptually defined as a real number between o and 1, it could not be classified as quantitative factors, because $\bar{Y}$ will probably not be a continious function of it. Then it is unique among the other quantitative factors and expression (14) can be written:

$$\bar{Y} = \phi(x_1, x_2, \ldots, x_p) + \mathcal{E}(r) \qquad (15)$$

where $\mathcal{E}(r)$ is a random effect dependent upon the random number seed :

Further, if we assume that $\mathcal{E}(r)$ is independent of experimental factors and has zero expectation, the expected simular response can be defined as:

---

(1) a detailed discussion of this point can be found in Mihram [18]

$$E(\bar{Y}) = \phi(x_1, x_2, \ldots, x_p) = \phi(\vec{x}) \qquad (16)$$

It is the nature of the unknown function $\phi(\vec{x})$, termed simular response function, that we try to investigate by the simulation experiment.

In practical situations, any attempt to develop the exact form of $\phi(\bar{x})$ could not be justified from economical point of view. In addition, for many experimental purposes, it is unnecessary to consider the form of the true function. A flexible graduating function, will often be satisfactory to express the relation ship between $E(\bar{Y})$ and the p factors. Further more, many experimental strategies, divide the whole operability region of the factors space, into a number of smaller regions of immediate interest. Within these regions, the experimenter may feel it is reasonable to represent $\phi(\vec{x})$ by a known functional form, for example a polynomial, although he may know that such representation would be quite inadequate over the whole operability region.

As a result of the previous discussion, $\phi(\vec{x})$ may be approximated by:

$$E(\bar{Y}) \simeq f(x_1, x_2, \ldots, x_p ; \theta_1, \theta_2, \ldots, \theta_l) = f(\vec{x}, \vec{\theta}) \qquad (17)$$

Where f is a known functional form indexed by some unknown vector $\vec{\theta}$.

The way with which we investigate the function $f(\vec{x}, \vec{\theta})$, in order to yield information about simulated system, depends on the experimental objectives. Accordingly we distinguish two types of experiments, Exploratory and Optimization.

2.3 EXPLORATORY EXPERIMENTS :

If experimenter wishs to study the relative importance of the factors $\bar{x}$ as they affect the simular response $\bar{Y}$, he may use one of the following designs:

i) Screening designs

At the begining of investigation, specially with complicated simulation models, the experimenter may face the problem of so many factors. It may happen that not all the p factors are important but only a few, say p' factors. Therefore, he screen for them.

The statistical literature contains many designs, for example, fractional factorial designs [4,8,17], random disigns [24], group screening designs [21], and super saturated designs [3]. The investigator has to select the design which fit his particular experimental situation.

ii) Designs for estimating parameters

When experimenter has a prior knowledge about the simulated system, either due to theoritical background or from previous investigations. He may assume that a particular functional form $f(\vec{x},\vec{\theta})$ is a good approximation of the true response function $\phi(\vec{x})$, in such a way that bias due to inadequacy of $f(\vec{x},\vec{\theta})$ to represent $\phi(\vec{x})$ can be neglected. So his goal will be to select an experimental plan to estimate the unknown parameters $\vec{\theta}$ with high accuracy. Two basic approaches were proposed to develop an experimental design, either to use a simple factorial or fractional factorial design [4,8,17], or to select a design based on a variance criterion as D - optimal designs [15].

iii) Designs for exploring response surface

When knowledge about simulated system is limited, the object is to approximate, withen a given region of factors space, $\phi(\vec{x})$ by some graduating function $f(\vec{x},\vec{\theta})$ which most closely represent the true simular response function.

Accordingly the following design requirements have to be considered:

a) The design should allow sequentialization so that designs of higher order can be developed with minimum loss of information.

b) The design should consider not only sampling variation but also bias error.

c) The design should allow a check to be made on the representational accuracy of the postulated model.

2.4 OPTIMIZATION EXPERIMENTS:

The purpose of these experiments is to find the combination of factor levels at which simular response function $\Phi(\vec{x})$ is optimized.

A particular attention will be devoted, in the next section, to the explanation and the applicability of this experiment.

To summarize, any attempt to develop an experimental method for investigating management systems by simulation, should select a sampling plan which define an efficient procedure for estimating the variance of simular response $\bar{Y}$. The estimated variance measures the accuracy of results and then can be used to determine the appropriate run length. Having accomplished this task, an experimental strategy may be defined for investigating the interdependence between the simular response $\bar{Y}$ and the experimental factors.

3 - EXPERIMENTAL OPTIMIZATION TECHNIQUES

The choice of an experimental strategy that will yield an optimal solution depends on the type of factors in the simulation model. When all factors are quantitative, an optimum seeking routine can be used in order to find the combination of factor levels that optimize the response $\bar{Y}$. But the existance of some qualitative factors, as policy specification or operating rules, limit the search procedure to the choice between a number of experimental alternatives.

3.1 The search for an optimum combination of factor levels;

When all factors are quantitatives, the investigator will wish to find in the smallest number of simulation runs, the Point $(x^{\circ}_1, x^{\circ}_2, \ldots, x^{\circ}_p)$, withen the factors space, at which $\Phi(\vec{x})$ is a minimum or a maximum.

Since simular response function is not known in advance and is subjected to random variation, we think that the most reasonable strategy will be to fit a sequential program of investigation consisting of the following steps :-

i) Divide the whole region of interest into a number of small subregions, so that we can explore adequately a small subregion with a moderate number of simulation runs.

ii) Use the results obtained in one subregion to move to a second in which simular response $\bar{Y}$ is better.

iii) Repeate the previous steps until the attainment of a near stationary region where no improvement in the simular response can be achieved.

iv) In this limited region, conduct a more detailed experiment in order to determine the local nature of the function $\Phi(\vec{x})$.

In the following sections we discuss breifly the two main elements of this sequential program, seeking a near stationary region and exploring it.

3.1.1 Seeking a near stationary region

When the starting conditions of simulation are fairly remoted from the stationary point, an optimum seeking technique will be needed to move rapidly through the factors space to a near stationary region.

Brooks [6] compared four optimum seeking methods, steepest ascent, univariate, factorial, and random search. He concluded that, when sequential investigation is possible, steepest ascent seems to be superior to the others, except in case of large

number of factors, where random search is more efficient(1).

Recently, Smith [25] showed that random search should not necessarily be the search technique selected in practical situations even in case of so many factors and he recommended the use of the steepest ascent.

Since the steepest ascent method is explained in detail in Box and Wilson [5] and Davies [8], we just mention, here, some remarks that should be considered when applying the method to simulation experiment(2).

i) Since we use the error variance to test the adequacy of the fitted function and the significance of model parameters, an accurate estimate of the variance of $\bar{Y}$ is needed in order to avoid any wrong conclusion.

ii) As the steepest ascent method is affected by the size of the experimental error [5], we may try to reduce it, by selecting a minimum variance design (see section 2.3), by increasing simulation run length, and if possible, by using a variance reduction techniques.

iii) If possible, provision should be made to estimate some of higher order coefficients that were not included in the postulated model. The study of these coefficients will provide some indication of whether the assumption that these terms can be ignored is a reasonable one or not.

3.1.2 Exploring the near stationary region

The experimenter may arrive at a near stationary region either as the result of successive application of steepest ascent method,

---

(1) This is explained by the fact that, in random search algorithm, the number of experimental trials is not a function of the number of factors.

(2) The method will be explained using an example model in section 5

or because he has already found it at the begining of his investigation. In either cases, only immediate neighbourhood need be explored to determine the local nature of response function $\phi(\bar{x})$ and this may be done without excessively large number of experimental points.

Although many authers have ignored the exploration of near stationary region, and are only satisfied by finding the approximated optimum point, we think that it is an important step in case of simulation for the following reasons:-

First, it should be remembered that because of random error and possible lack of fit between fitted equation and the true response $\phi(\bar{x})$, it must not be implied immediatly that the true surface has a maximum (or minimum) at the selected point. So in practice further exploration and confirmatory runs should be performed arround the stationary point of the fitted surface.

Second, the discovery of factors dependence of a particular type may give us an idea about the cost of departure from the optimum point, if it was impossible to reach it in practice. For example, finding the direction of a stationary ridge means that we can know the different combinations of factor levels that optimize the response $\bar{Y}$. Then the choice between these alternatives can be decided according to the cost of each combination or according to an auxiliary response.

Two exploratory techniques are proposed in the statistical literature, Canonical analysis [5,8] and Ridge analysis [9]. The authors matched the two techniques in a single computer program in order to have more robust conclusions. This can be done by using canonical analysis to reveal the factor dependence withen the local stationary region, then using ridge analysis to evaluate the locus of the absolute maximum or minimum when augmenting the experimental region.

3.2 The choice between experimental alternatives

When simulation model contains qualitative factors, as managerial policies or operating rules, the search procedure will be reduced to the optimum choice between a number of experimental alternatives. More specifically, it is required to find the combination of factor levels corresponding to the best simular response $\bar{Y}$, such that the probability of correct selection (CS) is at least P^*, given that the difference Δ between the best and the next best simular response is at least Δ^*. This may be stated formally as

$$P_r\,(CS/\Delta \geq \Delta^*) \geq P^* \qquad (18)$$

The previous formulation of the problem permit the use of one of the multiple ranking procedures [2,17,22,23] . Most of these methods assume normality, independence and commun known or unknown variances.

In practical simulation models, the distribution of the response $\bar{Y}$ is not known, variances are not known and tend to differ, so either we manipulate simulation runs to meet these assumptions or we hope that the effect of their violation is negligible.

After consulting several multiple ranking procedures, the authers choosed three of them that seem to be attractive for simulation circumstances. The selected procedures are Bechhofer and Blumenthal [2], Paulson [22] , and Sasser et al [23] .

Bechhofer method is the only one extensivly tested for its sensitivity to assumptions violation, it is quite robust and relatively efficient [17] . Unfortunately, it cannot capitalize on favorable configurations of population means. Paulson's procedure give us the possibility to eliminate inferior populations, so it might be advantageous when comparing a large number of alternatives.