Proceedings of the 1983
Winter Simulation Conference
S. Roberts, J. Banks, B. Schmeiser (eds.)

SIMULATION METHODOLOGY: STATISTICAL ASPECTS

G. Arthur Mihram, Ph. D.
Post Office Box No. 1188
Princeton, N.J. 08542

The present paper, of a tutorial nature, relates the literature of our simulation methodology since the publication of SIMULATION: STATISTICAL FOUNDATIONS AND METHODOLOGY [Academic Press, 1972 (1970)]. Explained will be the expansion of its Principium of Seeding not only to the Fundamental and the General Principia of Seeding but also to a Third Principium of Seeding. Of course, all three principia relate to the procedure(s) by which we simulationists can ensure that we meet the conditions (e.g., independence, experimental error, blocking) of extant statistical methodology, the conditions by which we may properly design and analyse experiments with a stochastic, computerised, and algorithmic (simulation) model. The paper (tutorial) will extend the current literature of statistical and simulation methodology by calling attention to a new class of pseudo-random number generators.

1. INTRODUCTION

A number of issues relating to the application of statistical methodology have arisen since the publication of SIMULATION: STATISTICAL FOUNDATIONS AND METHODOLOGY (1972). That book introduced the Principium of Seeding as the principle by which we simulationists can assure that the simular responses (model outputs) which our models generate can be properly statistically analysed.

Though the Principium of Seeding remains valid, it became clear to the present author that many practitioners were not sure of its motivation(s). The Principium of Seeding was worded as two principles (Mihram, 1974; 1976a; 1976b): the Fundamental Principium of Seeding and the General Principium of Seeding.

The Fundamental Principium of Seeding ensures randomness, statistical independence, and a proper measure of experimental error, whereas the General Principium of Seeding, when applied, not only assures that these three conditions for subsequent statistical analysis will be met but also reveals that experimentation with random number seeds per se has a very limited application in simulation methodology. The General Principium of Seeding, nonetheless, reveals the condition under which the statistical technique known as random blocking (also called 'variance components') can be applied in simulation methodology (i.e., in simular experimentation).

A Third Principium of Seeding was subsequently introduced (Mihram, 1979) when it became clear that our special-purpose simulation programming languages were failing to enforce adequate discipline on simulationists conducting simular experimentations with a stochastic model. The need for meeting the statistician's requirements for randomness, statistical independence, and experimental error can readily be forgotten by analysts whenever the simulation programming language used to author the model does not itself require that the seed specifications be conducted in accordance with the Principia of Seeding. The Third Principium of Seeding was introduced in order to correct a tendency among simulationists to believe that one pseudo-random generator is "better" than another (of the same class) of differing parameters. Since the multiplicative and mixed congruential pseudo-random generators are merely card-shufflers, no one shuffling (parameter specification) can possibly be better than any other proper shuffling.

The present paper will then discuss the three Principia of Seeding, noting that adherence to them has also been revealed to be a matter of ethics, of professionalism (Mihram, 1981) among us simulationists. The relationship between pseudo-random number generators and encryption techniques will also be explored (Mihram, 1982).

2. TERMINOLOGY

2.1 Outline of the Paper

The present section will introduce and define a number of terms for the ensuing sections. The very next section (3) will review the six-stage Scientific Method (cf. Mihram, 1972: Figure 5.2), introducing therein a more appropriate term (viz., Perscrutation) for that activity which some simulationists have incorrectly (Fishman and Kiviat, 1968) termed "verification" and others not sufficiently definitively (Mihram, 1976c) called "scrutinisation".

The second subsequent section (4) will review the three Principia of Seeding in the context of the final three stages of a simulation model's development: i.e., in the three stages in which experimentation with an already programmed simulation is conducted: Perscrutation [III], Confirmation [IV], and Inference [V].

The third subsequent section (5) will then review the recent literature of three statistical techniques in simulation methodology: (A) antithetic variates; (B) blocking by means of seeds themselves; and, (C) regenerative processes, noting that there is only very limited applicability of any one of the techniques in simulation methodology.

The fourth subsequent section (6) will note, in the context of the Principia of Seeding, that our pseudo-random number generators not only are "card-shufflers" (Mihram, 1979a) but also are exemplars of encryption techniques (Mihram, 1982).

The paper then closes (Section 7) by noting the moral (i.e., the ethical) issues raised by Admiral Bobby Inman (1982) in the process for publishing material dealing with encryption techniques. That adherence to the Principia of Seeding is itself an ethical issue (Mihram, 1982) is presented in the context of the very motivation for conducting simulation methodology in concordance (Mihram, 1983) with the established, six-stage Scientific Method: viz., national survival.

2.2 The Definitions

That computer programming is algorithmic, and therefore not mathematics, has been earlier established (Mihram, 1977, e.g.) and then alluded to (Manthey and Moret, 1983). Some scientists (e.g., Darwin) express themselves in a natural language, using first-person grammar; others (e.g., Newton) present their explanation of some phenomenon in the language of mathematics, using a third-person grammar. A computer programming language, however, is <u>not</u> mathematical; rather it is of second-person grammar, is algorithmic.

The fact that <u>some</u> mathematical proofs are constructive (e.g., Euclid's Algorithm) does not really contradict the facts that scientists employ mathematics as a third-person description of nature and that computer programming is of second-person grammar. A mathematical proof, if expressed as an algorithm, is not science; it is merely art.

<u>Science</u> is that human activity devoted to the search for the very explanation for some naturally occurring phenomenon. As such, a <u>scientist</u> is one who seeks to provide the very explanation of (i.e., the truth about) some naturally occurring phenomenon. The scientist is required therefore to be an artist, for he must construct a model in order to provide his explanation (Mihram, 1977).

When the present author uses the term, "simulation," he is referring to a completed (i.e., a programmed) algorithmic model in a machine-readable format. A simulation is therefore a model, not an exercising of that model.

To be more explicit, a <u>computerised simulation</u> is the model within the memory of the computer, whereas a <u>simulating</u> is an encounter (Mihram, 1972) with the computerised simulation; i.e., a simulating is the exercising of the model to mime once, over a single time period, the system being modelled.

I believe that it was John McLeod who introduced the term, "<u>simuland</u>": that real system being modelled. This is a helpful term indeed, particularly useful in discussing the Confirmation stage of a model's development, at which stage one requires a two-sample statistical test: one sample from the simulatings (encounters), a separate sample of observations from the simuland.

There does exist an adjectival form of simulation: viz., <u>simular</u>: of, or having to do with, the simulation—as opposed to having to do with the simuland. The <u>simular response</u> becomes then the vector-valued, <u>time-oriented</u>, realisation of a simulating:

$$R(t) = [R_1(t), R_2(t), \ldots, R_J(t)], \quad t = 0,1,2,\ldots,T, \qquad [1]$$

wherein $\underline{T}$ represents the <u>simular duration</u> and $\underline{J}$ the number of time-dependent attributes for which the simulation model's executive routine is capable of maintaining a trace, or record.

It follows that the term, "<u>simular experimentation</u>," is defined as a meaningfully structured set of $\underline{n}$ simulatings with a simulation.

One should note that the present paper does not preclude the application of any of the statistical techniques to be discussed to a computerised, stochastic, and yet truly mathematical model, such as one written as the "input data" for a CSMP or a DYNAMO processor (See Mihram, 1976c, e.g.). Indeed, the Principia of Seeding apply equally well to encounters with (experimentation with) these computerised mathematical models.

3. THE SCIENTIFIC METHOD

In either event, simular experimentation is conducted for one of two purposes: (A) to establish a model's credibility; or (B) to make inferences regarding a simuland. The first matter has led (Mihram, Innis, et al., 1974) to the understanding of the six-stage Scientific Method:

0. EXTANT KNOWLEDGE—the 'Null Stage' of a model's development, representing the total cumulation of all the content of all of Mankind's museums and libraries in existence;

I. SYSTEMIC ANALYSIS—the first stage of a scientist's model's development, representing both his confrontation with a yet-to-be explained naturally occurring phenomenon and his mental reflexions on those observations, (i.e., Stage I is '<u>Observations + Reflexions</u>'), leading to:

II. POSTULATED HYPOTHESIS—the second stage of a scientist's model's development, representing the artwork (sculpturing, writing, programming) necessary to provide the explanation derived by the scientist as a result of his reflexions on his observations (Stage I, above). This Stage begins with a decision as to which modelling format (e.g., algorithms) the scientist is to employ and concludes with the completion of that artwork (e.g., his simulation programme);

III. MODEL PERSCRUTATION—the third stage of a scientist's model's development, the first of the two stages dealing with the credibility of the explanation provided at the end of Stage II (above), and deals with the logical and grammatical rectitude (precision) of the model itself. Any error such as a logical or typographical error in a sequence of mathematical equations submitted by a scientist to an editor) is corrected by the scientist's returning to Stage II [See Figure 1.], there to alter the model (artwork) appropriately and to then subject the corrected model itself to further perscrutation (Stage III) before proceeding to:

IV. CONFIRMATION—the fourth stage of a model's development, the second of the pair of stages dealing with the credibility of the explanation (model) provided at the end of Stage II. This stage deals with the compatibility of any impeccably logical deduction (derived from accepting the thoroughly perscrutated model as correct) with the real world. Any error so discovered is likely the result of an error in Stage I (Systemic Analysis), to which the scientist must return so as to reformulate his explanation (Stage II: Postulated Hypothesis), then subject it to its own Perscrutation (III) before returning to this Confirmation (IV) stage for further credibility testing; if perchance an incompatibility is discovered in this Confirmation Stage, yet is actually the result of a previously undetected logical or grammatical error (such as an undiscovered 'bug' in the simulation model), then the correction is made by a direct return to Stage II, rather than a return to Stage I [Figure 1 makes clear these alternative cybernetic paths (Mihram, 1975) for correction.], before returning again to Stages III and IV;

V. INFERENCE—the final stage of a scientific model's development, concerned solely with making impeccably logically derived deductions from the thoroughly scrutinised and confirmed model. In effect, a new model (book, computer tape, e.g.) has been placed on the library shelves, thereby augmenting Human Knowledge (Stage 0), and permitting further human progress. (e.g., Mihram, 1976c; Mihram, 1983; and their many referents.)

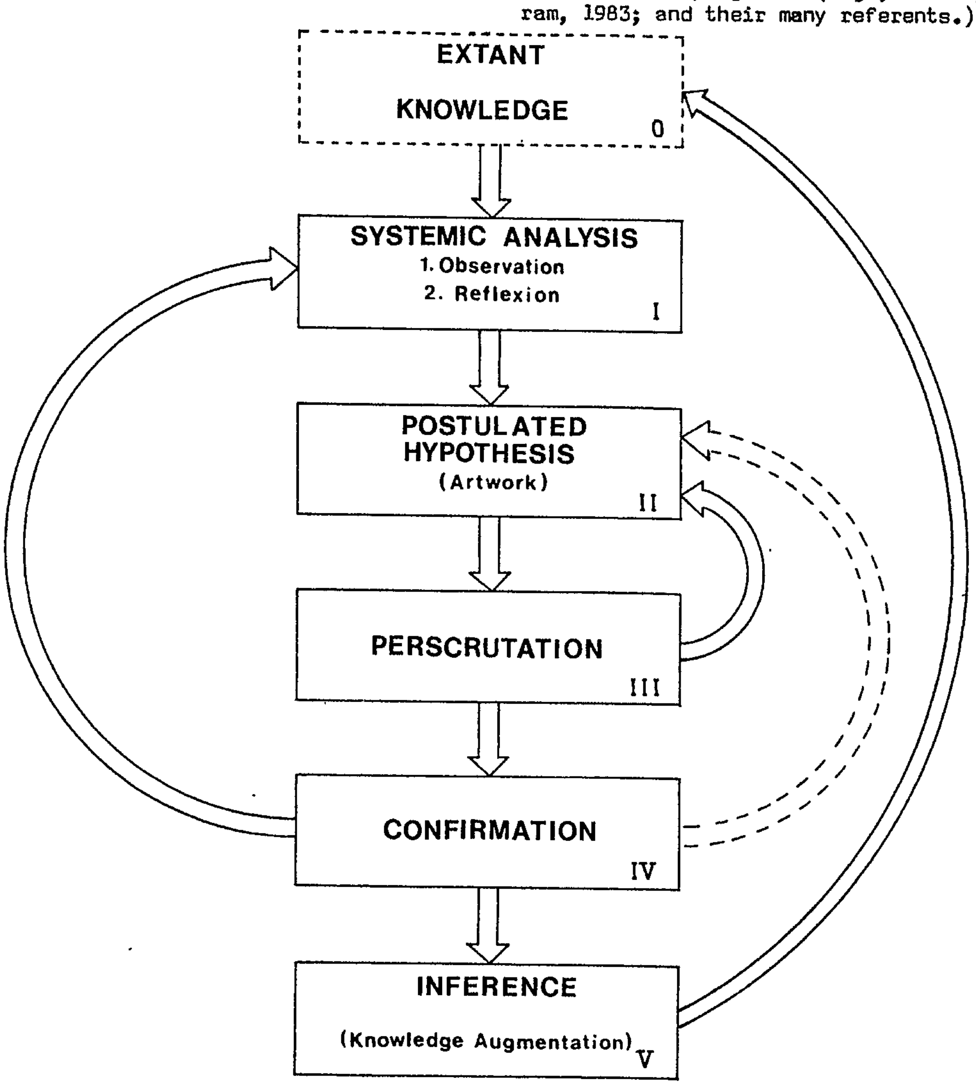

Figure 1: The Scientific Method

The present discussion (outline) of the Scientific Method is terse in presentation, but has been discussed at length in the literature of our simulation methodology (e.g., Mihram, 1976c; and its referents) as well as that of our companion disciplines(e.g., Mihram, 1983; its referents; and Mihram and Mihram, 1982, particularly its Table II). For the present, however, one should note that the six-stage model-building process (the Scientific Method: Figure 1) is itself confirmed (e.g., Mihram and Mihram, 1982, particularly its own Figure 1 depicting how Popper's Three Worlds are unified as one) by an examination of Nature's own two earlier model-building methodologies (genetic, then neural) for ensuring survival. The Scientific Method becomes therefore the model-building process by which national survival is ensured.

4. THE THREE PRINCIPIA OF SEEDING

Our goal in simulation methodology is therefore the desire to see that simulation models are constructed in accordance with the Scientific Method. Whenever we opt to construct a stochastic simulation of a system, however, we must adhere also to the principles which will assure that the resulting simular experimentation is conducted in accordance with the conditions set by the techniques of statistical methodology. In particular, experimentation with a stochastic simulation implies that we meet the statistician's requirements for:

Randomness
Statistical independence
Experimental error and/or
The distinction between 'blocks' and 'variance components'.

4.1 Fundamental Principium of Seeding

As introduced in Equation [1], the simular response, emanating from a simulating, is a time-ordered, multivariate (vector) record of the J attributes being maintained by the particular simulation programming language's (SPL's) executive routine. For the stochastic simulation, any one of these simular responses is a transformation of the total set of the model's input conditions, though this transformation would seldom be capable of expression in terms of a mathematical formula:

$$R_j(t) = R_j(\vec{S};\ \vec{x},T), \qquad [2]$$

for any $t = 0,1,2,\ldots,T$,
for any $j = 1,2,\ldots,J$,

where (as in Eq. [1]) T is the simular duration specified by the analyst and J is the total number of attributes being maintained by the SPL's executive routine (in accordance with the model's author's instructions), yet where

$$\vec{x} = (x_1,x_2,\ldots,x_p)$$

is the set of p environmental specifications required by the computerised simulation, and where

$$\vec{S} = (S_1,\ S_2,\ \ldots,\ S_K) \qquad [3]$$

is the set of K seeds required to initiate the model's stochastic behaviour in any simulating.

Quite generally, then, one could compute/determine:

$$Y(T) = f[\vec{R}(t),\ t = 0,1,2,\ldots,T] \equiv Y(\vec{S};\ \vec{x},T) \qquad [4]$$

as the simular response, hereinafter taken for the expediency of discussion to be univariate, where f is any function which the analyst might meaningfully specify.

The vector, $\vec{S}$, of seeds can contain any number of any of the five phyla of seed types (Mihram, 1979b), but in any event the vector [3] can be juxtaposed, as on a single input data card or field, to become:

$$U = (S_1 | S_2 | \ldots | S_K), \qquad [5]$$

which, if viewed with a decimal, or bit, point placed afront, becomes a variable restricted to the range, (0,1), of the unit interval.

The simular response [4] therefore becomes

$$Y(T) = Y(U;\ \vec{x},T), \qquad [6]$$

with the (p+1)-vector $(\vec{x},T)$ constituting the nonstochastic vector of input conditions required by the computerised simulation.

In this form [6], the simular response becomes a statistician's transformation of a uniformly distributed random variable, U. Indeed, for any fixed (p+1)-vector,

$$(\vec{x},T) = (\vec{x}^*,T^*),$$

one could conceptually initiate a barely finite sequence of m encounters with the computerised stochastic model, one simulating (encounter) for each successive seed value: $1,2,\ldots,m$, where m is the largest value assignable to the (juxtaposed) seed [5]; actually,

$$m = \prod_{k=1}^{K} m_k, \qquad [7]$$

where m_k is the number of values assignable to the kth of the model's K seeds.

As the transformation, even via the implicit function defined by the algorithmic simulation itself, of a uniformly distributed random variable, the simular response [4] becomes itself a random variable, one having its own probability distribution function, a p.d.f. which conceptually could be formed from the histogram of all m simular responses:

$$Y(1;\ \vec{x}^*,T^*),\ Y(2;\ \vec{x}^*,T^*),\ \ldots,\ Y(m;\ \vec{x}^*,T^*),$$

in which case the model's environmental conditions $(\vec{x}^*,T^*)$ would merely become parameters for the pdf:

$$g_Y(y;\ \vec{x}^*,T^*). \qquad [8]$$

One of the fundamental requirements of the statistician is that one draw a random sample from the p.d.f. [8] at hand. The realisation (Mihram, 1972) of this requirement led to the (Fundamental) Principium of Seeding:

> The juxtaposed seed, U, for a stochastic, computerised, simulation model must be randomly and non-repetitively selected from among the set of m admissible seed values.

Adherence to this principle ensures not only the randomness required by the statistician but also the statistical independence required by most of the statistical tests employed by the simulationist during his perscrutation tests (Stage III, above). In the scrutinisation of a dynamic, stochastic, simulation model, we simulationists formulate an hypothesis, one which would have to be true of the stochastic simular responses for some particular input set $(\vec{x}^*,T^*)$ IF INDEED the model had been programmed correctly. For example, we might expect that the mean value of g_Y (See Eq. [8]) would be a known value, that g_Y be a particular density function, or that the time series be of a particular type (e.g., a normal process), in which case we would express the null hypothesis and proceed with the corresponding one-sample perscrutation test (Mihram, 1979c: Table II).

A statistician would view the simular response (Eqns. [4] and [6]) through the following notation:

$$Y(T) = r(\vec{x},T) + e(\vec{S};\ \vec{x},T), \qquad [9]$$

wherein $r(\vec{x},T) = E[Y(T)]$

is the mean (expected) value of the simular pdf [8] and

$$E[e(\vec{S};\ \vec{x},T)] = 0,$$

with e representing the <u>simular experimental error</u>.

Similarly, in the subsequent stage (IV: Confirmation) of a stochastic model's development, one requires a random sample of n_1 simular responses in order to compare this sample with a random sample of n_2 observations from a real-world system being modelled by some particular specification $(\vec{x}^*,T^*)$. A listing of these two-sample statistical tests for Model Confirmation has appeared in the literature (Mihram, 1979c: Table III).

4.2 The General Principium of Seeding

In the ultimate stage (V: Inference) of a stochastic simulation model's development, the already thoroughly scrutinised and arduously confirmed model will still require statistical techniques for analysis. Adherence to the Fundamental Principium of Seeding still meets in this final stage of the model's development the statistician's requirements for randomness, for statistical independence, and for experimental error. But, the meeting of this last requirement is even more important in the simular experimentation conducted with the programmed, scrutinised, and confirmed model.

The analyst at this point seeks to compare (statistically) the simular responses emanating from two differing environmental conditions (See Eq. [6]):

$$Y_1(T) = Y(U;\ \vec{x}_1^*,T) \text{ and}$$
$$Y_2(T) = Y(U;\ \vec{x}_2^*,T).$$

Again, statistical hypotheses would be tested: e.g., the equivalence of the corresponding cumulative distribution functions (See Eq. [8]):

$$G_{Y_1}(y_1;\ \vec{x}_1^*,T) \text{ and}$$
$$G_{Y_2}(y_2;\ \vec{x}_2^*,T);$$

or, the hypothesis that their two mean values are equivalent, thereby testing whether one systemic configuration $(\vec{x}_1^*)$ is or would be better (larger) than the other $(\vec{x}_2^*)$.

More generally, one can apply the statistician's experimental designs by adhering to the Fundamental Principium of Seeding (See Mihram, 1979c: Table IV): multiple rankings, analysis of variance, response surface methodology, time series analysis). However, a clear misunderstanding of the requirement for the Principia of Seeding has appeared (Schruben and Margolin, 1978) in the literature of statistical methodology. That misunderstanding has been corrected (Mihram, 1979d and its referents), though some statisticians (and simulationists) may still believe that they can conduct experiments with the seeds of large-scale stochastic simulations, perhaps due to the limited distribution of Mihram (1976b).

The General Principium of Seeding (e.g., Mihram, 1976b) states that:

> Some non-empty subset of the <u>K</u> seeds required by a dynamic, stochastic, and computerised simulation must be randomly and non-repetitively selected from among their collective set of admissible values.

Whenever a subset of a model's <u>p</u> environmental specifications, $\vec{x}$, remains fixed in defining a set of <u>n</u> model encounters, a <u>block</u> of simular responses has been defined. Whenever a (proper) subset of a model's <u>K</u> seeds, $\vec{S}$, remains fixed in defining a set of <u>n</u> model encounters, a <u>random block</u> of simular responses has been defined.

It is quite possible to hold constant <u>all</u> <u>p</u> of a model's environmental specifications, $\vec{x}$, throughout the definition of <u>n</u> model encounters, so that the adherence to the Fundamental Principium of Seeding would merely assure that a random sample of responses would result—as discussed in the material above leading to Equation [8]. Furthermore, even if each of the p-vectors, $\vec{x}$, is varied from encounter to encounter among the <u>n</u> simulatings, adherence to the Fundamental Principium of Seeding assures that experimental error (See Eq. [9]) will be present in the data used in any ensuing statistical analysis. The methodology of statistical experimental designs is directly applicable in simulation methodology for simulationists who adhere to the Fundamental Principium of Seeding.

However, the entire K-vector of seeds cannot be fixed throughout the definition of the <u>n</u> simulatings in a simular experimental design. To so fix $U = U_0$ would not permit any measure of experimental error.

As a specific example, suppose that the <u>only</u> stochasticity in a model of a grocery store is the sequence of a day's customer arrival times, and that these random arrival times are provided as a tape of actually recorded observations from a day in the past. The simulationist expects that there exists a file of these daily tapes, so that he may randomly select a tape from this file for each model's encounter. If he should hold constant the model's environmental conditions $(\vec{x}^*,T^*)$ for a set of <u>n</u> model encounters, he would find <u>no</u> variation in the <u>n</u> simular responses: <u>unless</u> from encounter to encounter he alters randomly the seed value (i.e., he selects randomly a tape of the pre-recorded random customer arrival times).

Yet, a <u>proper</u> subset of a model's K seeds <u>can</u> be held constant throughout a simular experimental design, PROVIDED THAT the analyst is aware that he has thereby defined and must account for a <u>variance component</u>: what the statistician calls <u>"random blocking"</u> (Mihram, 1976b). In the preceding example, if some other aspect of the model were being handled stochastically (e.g., customer's time at the checkout counter), then to initiate each successive encounter with the model by the same single day's record of customer arrival times would define a (random) block of simular responses; of course, in this instance, the randomness of the customer's times at the check-out counter would need to be spawned by adherence to the (General) Principium of Seeding.

4.3 The Third Principium of Seeding

Thus, the Fundamental and the General Principia of Seeding really form the same requirement; the General Principium, however, reveals to the statistician (and the simulationist) how some very limited experimentation with random number seeds can be conducted. This type of experimentation, nonetheless, is valid only in the context of the statistical feature of "random blocks," providing a 'variance component' rather than a 'block effect' in the subsequent statistical analysis of variance.

In the aforementioned example of a stochastic simulation of a grocery store, the repetitive use of the very same pre-recorded sequence of customer arrival time in each of the n encounters in a simular experimental design means that one is actually making inferences about the grocery store's performance on one particular day (on one particular Monday, e.g.), so that one is not able to make inferences regarding the general performance of the store until one has removed the effect of this variance component (the random block effect).

However, the Third Principium of Seeding (Mihram, 1979a) was introduced in reaction to a noted deficiency in our extant simulation programming languages. These languages (SPL's) typically employ a multiplicative (or a mixed) pseudo-number generator (See, e.g., Mihram, 1972: p. 55), yet the one (two) parameter(s) of the generator is (are) imbedded within the SPL itself and are not made accessible for change by the analyst. Hence:

> At the outset of a simular experiment of n successive encounters with a stochastic simulation model which employs one or more pseudo-random number generators, each generator's parameters should be selected randomly (and non-repetitively) from among its set of admissible values.

Adherence to this Third Principium of Seeding ensures that the analyst will be using a "well-shuffled deck" of the number cards provided by each pseudo-random number generator (Mihram, 1981) used in the n simulatings. With respect to the four requirements of statistical techniques, adherence to this Third Principium of Seeding ensures randomness. There should be no reason to favour any randomly shuffled deck of cards over any other randomly shuffled deck.

Hence, the three Principia of Seeding are principles, the adherence to which provides the qualities required for the application of statistical techniques (including time series analyses) in simulation methodology. They ensure the randomness and statistical independence of the simular responses required in the one-sample statistical tests used in Stage III (Perscrutation) of a model's development; they ensure the randomness, statistical independence, and the experimental error required in the two-sample statistical tests used in Stage IV (Confirmation) of its development, and, they ensure that not only these three qualities but also those called "experimental units," "blocking," and "variance components" are present as desired in the experimental designs employed with the already perscrutated and assiduously confirmed model in Stage V (Inference). One may examine Tables I, II, and III of Mihram (1976b) for a delineation of the statistical tests thereby applicable in the three respective stages of a model's scientific development.

5. THREE STATISTICAL TECHNIQUES

Adherence to the three Principia of Seeding makes available therefore an extensive 'catalog' of statistical techniques for application in the experimental stages (III, IV, V) of a model's development. Three other topics in the statistical literature do seem to require some clarification, particularly with respect to their presentation in our literature of simulation methodology: (A) antithetic variates; (B) blocking by means of seeds themselves; and, (C) regenerative processes.

5.1 Antithetic Variates

Two types of antithetic variates, the simple and the generalised, were suggested in straightforward Monte Carlo estimates of integrals (Hammersley and Handscomb, 1964), the purpose being to save computational time by employing functionally related 'seeds', U and V, which would induce a negative correlation between the integrands, h(U) and h(V), in

$$\theta \equiv \int_0^1 h(x)dx,$$

when estimated by

$$\theta^* \equiv n^{-1}\sum_{i=1}^{n} h(U_i) \text{ and}$$
$$\hat{\theta} \equiv n^{-1}\sum_{i=1}^{n} h(V_i), \qquad [11]$$

wherein $V_i = v(U_i)$ is a function continuous and of slope 1 over the unit interval except at (at most) a finite number of points, where the U_i, $i = 1,2,\ldots,n$ are randomly and independently selected uniformly distributed random variables.

The case for $V \equiv v(U) = (1-U)$ provides the simple antithetic variates, whereas the functional forms acceptable to v for the application of the generalised antithetic variates have been discussed earlier (e.g., Mihram, 1974) and there related directly to the mixed congruential techniques' transformations themselves. Indeed, the inapplicability of any of the four variance-reducing techniques of Monte Carlo analysis has been demonstrated, the direct result of the fact that the simular response function (Eq. [10]) is actually

$$r(\vec{x},T) = m^{-1}\sum_{i=0}^{m-1} Y(i;\vec{x},T), \text{ or}$$
$$= \int_0^1 Y(\alpha;\vec{x},T)d\alpha, \qquad [12]$$

so that neither the integrand nor its mathematical expression or properties are likely to be known.

A more recent paper (Ramsay and Wright, 1979) has confirmed the result that it is foolhardy indeed to apply antithetic variates blindly in simular experimentation. Variance augmentation rather than the desired reduction shall likely result.

5.2 Random Blocking and the Split-plot

The statistical technique of random blocks (variance components) has been discussed above and its applicability in simular experimentation presented in the context of the principled seeding of computerised stochastic simulations.

Another topic in the statistical literature is the split-plot (experimental) design (Mihram, 1978). The model, once computerised, becomes for the statistician the experimental unit; a treatment is then applied by specifying the model's p environmental conditions, $\vec{x}$. Randomisation of the treatment to the experimental unit occurs by the specification of the seed, $U = \vec{S}$, to the computerised model. Blocks of experimental units are defined by holding constant q ($< p$) of the environmental conditions constant throughout the n encounters of a simular experimental design, whereas (as discussed above) random blocks (variance components) arise whenever any proper subset of the computerised model's K seeds remains fixed at the initiation of each of the n encounters. The split-plot design occurs in simulation methodology only whenever some facet (sub-routine) of a model could be replaced in toto by a more detailed sub-routine, one miming some systemic property with much greater detail; by defining the properties (say x_{p+1}, x_{p+2}) of this more detailed simulation, one would be introducing a split-plot effect.

5.3 Regenerative Processes

A third statistical technique deemed by many (e.g., Crane and Iglehart, 1975) to be appropriate to simulation methodology is the application of regenerative processes. Unfortunately, as will be shown below (See also Mihram, 1976a), the technique is quite inappropriate for large-scale stochastic simulation models because the technique's requirement for statistical independence can seldom (and could hardly be expected to) be met.

The simular response is indeed a multiple time series, or multivariate stochastic process,

$$\left\{\vec{R}(t)\right\}_{t=0}^{\infty} = \left\{\vec{R}(\vec{S};\ \vec{x},t)\right\}_{t=0}^{\infty}, \qquad [13]$$

as reflexion on Eqns. [1] and [2] reveals. Indeed, since any stochastic process is an ensemble of time-dependent functions (of realisations), it follows that, for any fixed (p+1)-vector of environmental conditions, a single realisation from a simular stochastic process results from each different specification of the (juxtaposed) seed, $U = \vec{S}$ [See Eq. [5].).

Regenerative processes become useful whenever a stochastic process, say $\{X(t)\}$, has realisations which return from time to time to a previously held (i.e., some specific) position or state. With the assumptions that each realisation (with probability one) will return again and again to this state and that these successively observed, state-to-state, segments of the realisation are statistically independent, then one would hope that he could obtain n realisations without needing to specify n random number seeds in accordance with the Principia of Seeding; yet:

$$\begin{aligned}&[R(U;\ \vec{x},t),\quad t=0,1,2,\ldots,t_1]\\&[R(N_1(U);\ \vec{x},t),\quad t=t_1,t_1+1,\ldots,t_1+t_2] \qquad [14]\\&[R(N_2(N_1);\ \vec{x},t),\quad t=t_1+t_2,\ t_1+t_2+1,\ \ldots,\ t_1+t_2+t_3],\quad \text{etc.,}\end{aligned}$$

wherein t_i represents the time (random) between the (i-1)st and the ith regeneration, and N_i is the value (location) of the pseudo-random number generator at the time of the ith regeneration.

One notes immediately that the successive segments of the realisation (See Eq. [14]) are very, very likely to be statistically dependent, for, after all, the responses in each segment are transformations of the same random variable, U. Furthermore, since the typical simulation model's response (See Eq. [13]) is multivariate, one could hardly expect that the values of the interesting variate (say, R_i) would be statistically independent of the concomitantly randomly varying R_j, all $j \neq i$.

One need be VERY, VERY cautious indeed in applying the techniques appropriate to regenerative processes in statistical methodology.

6. PSEUDO-RANDOM NUMBERS AND ENCRYPTION

Since any of the multiplicative (or mixed) congruential pseudo-number generators is itself a "shuffler" of numbered cards, it should not be too surprising to learn that mathematicians had already sighted them as ideal "shufflers" of, say, the letters of an alphabet and therefore had seen them as devices for encrypting symbols. All one need do is assign to every symbol (i.e., to every key on a keyboard) a number $\{0,1,2,3,\ldots,m\}$, then secretly shuffle these numbers so as to transmit the code (symbol) now in shuffled position j instead of the symbol representing j. (Alternatively, one could transmit for the symbol which had been assigned value j the symbol assigned to the value resulting immediately from the (mixed) pseudo-random number generator when seeded by j: There exist both an ordinal and a cardinal technique for encrypting symbols.).

More details of this procedure have been delineated earlier (Mihram, 1982), with reference to the "prime factor" technique for encryption (Richards, 1982a) as well. Indeed, the prime factor technique provides another method for shuffling a deck of numbered cards (Mihram, 1983b), so that theoretically-minded (i.e., number-theoretic) simulationists now have an opportunity to compare other "card-shuffling" techniques for their application in simulation methodology.

7. CONCLUSIONS

The fact that our pseudo-random number generators can serve as encryption techniques has raised a second ethical issue (For the first, see Mihram, 1982; for the second, see Inman, 1982.) in their use. One's value-ranking (e.g., nation vis-a-vis family, nation vis-a-vis religious friends) is a matter of his integrity, though it has been shown that certain value-rankings are more valuable to national survival than others (e.g., Mihram, 1983a; Mihram, 1982). Indeed, it may have been brought to the attention of Richards (1982b) that there are many other American academics who did not like the apparent (though perhaps unintended) slap-in-the-face which his paper (Richards, 1982a) had given to Admiral Inman's request (Inman, 1982) that papers dealing with encryption be first submitted for review to, say, the National Security Agency.

The other ethical matter raised by our pseudorandom numbers is directly the motivation for the Principia of Seeding: viz., to ensure that randomness was, without any possibility of bias, instilled into the n encounters constituting a simular experimental design. Simply put, if our machine-readable stochastic simulations are to merit shelving in our national archives, then their truth(s) will only be enhanced by our adopting a professional attitude (Mihram, 1982) toward their honestly-conducted seeding.

We would do well as professionals to ignore the contemporary "concerns" regarding academic freedom (e.g., Kolata, 1982) and be prepared to note that one's academic freedom (i.e., his freedom to conduct and report his research without topical constraint [See TECHNOMETRICS 19: 225-I, 1977.]) is not constrained because certain topics of the day are to be reviewed first by a legislated national agency (CIA, FBI, NSA, e.g.) before being despatched for publication. Surely, return receipts for registered/certified mail carrying manuscripts to such agencies can be used to establish one's priority of publication or to reveal to academic administrative superiors whether an excessive delay is being encountered by a young, probationary academic who has patriotically submitted a manuscript to a national security agency. In this regard, the 1822 precedent could be noted: the suppression of publication of a Frenchman's manuscript for reasons of national security (Smith, 1982).

A look at the OXFORD ENGLISH DICTIONARY might also prove useful to us professional simulationists concerned about national survival/security. For the entry, "liberty," one would infer the formula:

LIBERTY ≡ FREEDOM + RESPONSIBILITY,

leaving one to ask whether certain chosen people who call for 'academic freedom' do so so loudly as to make rather inaudible the more proper call: the call for 'academic liberty'. Perhaps, given the established survival value of adherence to the Scientific Method (Mihram, 1975b), we should avoid usage of the term, "scientific freedom," and employ instead "scientific liberty," implying thereby our concern for our national responsibility.

The U.S. Supreme Court ruled on 29 June 1981, by a vote of 7 to 2, that one's highest priority must be the nation. The second verse of Katherine Lee Bates's "America the Beautiful" underscores clearly [WASHINGTON POST, 28.I.1981: p. E-12] the Court's fundamental point (Burger, 1981) in deciding the Agee caper:

> "It is 'obvious and unarguable' that no government interest is more compelling than the security of the nation."

REFERENCES

Burger Chief Justice WE (1981), Excerpts from Court Opinion on Revocation of Passport on Security Grounds, New York Times, 30.VI.1981: p. B-8.

Crane MA and Iglehart DL (1975, Simulating Stable Stochastic Processes, Operations Research, Vol. 23: pp. 33-45.

Fishman GS and Kiviat PJ (1968), Statistics of Discrete-event Simulation, Simulation, Vol. 10, pp. 185-195.

Hammersley JM and Handscomb DC (1964), Monte Carlo Methods, Methuen, London.

Inman Admiral Bobby R (1982), National Security and Technical Information, Central Intelligence Agency, Washington, DC 20505 (07.I.82).

Kolata G (1982), CIA Director Warns Scientists, Science, Vol. 215, p. 383.

Manthey MJ and Moret BME (1983), Computational Metaphor and Quantum Physics, Commun. Assoc. Comput. Machin., Vol. 26, p. 137.

Mihram GA (1972), Simulation: Statistical Foundations and Methodology, Academic Press, New York.

Mihram GA (1974a), Blocking in Simular Experimental Designs, J. Statist. Comput. Simul., Vol. 3, pp. 29-32.

Mihram GA (1974b), Variance Reduction and Random Number Generation, Modeling and Simulation, Vol. 5, pp. 845-848.

Mihram GA (1975a), Telecybernetics: An Essential Feature of Simular Credibility. In: Proc., 1975 Summer Computer Simulation Conference, Soc. Comput. Simul., La Jolla, Calif., pp. 1235-1242.

Mihram GA (1975b), An Epistle to Dr. Benjamin Franklin. Exposition-University Press, New York.

Mihram GA (1976a), Seeds: Antithetic Variates, Blocking, and Regenerative Processes. In: Proceedings, Hawaii International Conference on the System Sciences, Western Periodicals, Honolulu, pp. 148-150.

Mihram GA (1976b), Statistical Techniques in Simulation. In Selected Papers on Simulation, R. Chattergy (ed.), Western Publishers, Los Angeles, pp. 14-21.

Mihram GA (1976c), Simulation Methodology, Theory and Decision, Vol. 7, pp. 67-94.

Mihram GA (1977), Making Sure Computer Scientists Are Scientists, Commun. Assoc. Comput. Machin., Vol. 20, pp. 120ff and 452.

Mihram GA (1978), Prospect for Split-plot Designs in Simular Experimentation, 1978 Abstracts, Amer. Statist. Ass'n, Washington, DC, pp. 228-229.

Mihram GA (1979a), Pseudo-random Number Generators Are Really Card-shufflers. In: NCC '79 Personal Computing Proceedings, A.F.I.P.S., Arlington, VA pp. 318-326.

Mihram GA (1979b), Monte Carlo and Sampling Experiments: A Tutorial, Modeling and Simulation, Vol. 10, pp. 2185-2191.

Mihram GA (1979c), Simular Experimentation: Subjectivists and Objective Probabilists. In: Proc., Xth Pittsburgh Conf. on Model./Simul., W. G. Vogt and M. H. Mickle (eds.), pp. 2193-2202.

Mihram GA (1979d), Corrective Note Regarding Margolin's Assignment of Pseudo-random Numbers, Bulletin, Inst. Math. Statist., Vol. 8, pp. 90-91.

Mihram GA (1981), Are the Principia of Seeding Merely a Philosophical, or Rather, an Ethical Matter?, Modeling and Simulation, Vol. 12, pp. 765-770.

Mihram GA (1982), Moving Simulation Methodology Forward with Sociobiology, Modeling and Simulation, Vol. 13, pp. 1053-1057.

Mihram GA (1983a), Terminating Popper's UNENDED QUEST. In: Relation Between Major World Problems and Systems Thinking, G. E. Lasker (ed.), Soc. Gen. Syst. Research, Louisville, Ky., pp. 283-290 (1982).

Mihram GA (1983b), Representing Pluralism in Societal Modelling, forthcoming.

Mihram GA, Innis G. et al. (1974), Bibliography on Simular Credibility, Simulation, Nov. 1974, p. vi.

Mihram GA and Mihram D (1982), Model Realism, Simular Credibility, and Scientific Truth: Synonymous or Not? In: General Survey of Systems Methodology, L. Troncale (ed.), Soc. Gen. Syst. Research, Louisville, Ky., pp. 233-240 (1981).

Ramsay TE and Wright RD (1979), Variance Reduction Techniques: Some Limitations and Efficiency Limits, Modeling and Simulation, Vol. 10, pp. 2083-2090.

Richards I (1982a), Invisible Prime Factor, American Scientist, Vol. 70, pp. 176-179, March-April.

Richards I (1982b), Unbreakable Codes, American Scientist, Vol. 70, pp. 351-352, July-August.

Schruben LW and Margolin BH (1978), Pseudo-random Number Assignments in Statistically Designed Simulation and Distribution Sampling Experiments, J. Amer. Statist. Ass'n, Vol. 73, pp. 504-520.

Smith CS (1982), Damascus Steel, Science, Vol. 216, p. 242.