LECTURE 1. CONCEPTS OF SIMULATION AND MODELING

1.1 Generalities about systems, models and simulation

The word „*simulation*" derives from the Latin word „*simulatio*" that means the capacity to reproduce, represent or imitates something.

In mathematics, the term „*simulation*" has been used the first time by John von Neumann and S. Ulam in 1940-1944, when they conducted research of nuclear physics in the U.S. Together with N. Metropolis, Fermi and other mathematicians and physicists of the "Los Alamos" School, they also introduced a picturesque name in mathematics, namely the Monte Carlo Method. The name, of course, improper, comes from the fact that the first methods of generating/simulating random numbers were those offered by the roulette scores of the famous Monte Carlo casinos.

It is also said that simulation is more an art than a science.

1.1.1 Systems

The accelerated development of contemporary science and technology creates complexity, which is becoming more and more difficult to control, to master, to drive. In support of his efforts to master complexity, to know his components, to discover different laws governing it, man has created the notion of system.

The system is a set of interconnected elements (physical or logical components, laws, rules, etc.) that work together to achieve one or more purposes.

The element is a part of the system (a subassembly or a component) capable of performing a certain function within the system.

Examples:

- People live in social systems.
- Technological activity has produced complex physico-technical systems.
- A car is a system of components that work together to ensure the transport.
- The family is a system of cohabiting and raising children.

Systems are classified as *open* and *closed.*

An *open system* is characterized by:

- output corresponding to input into the system;
- the output is isolated from the input;
- the output has no influence on the input.

In an *open system*, the results of the past action do not influence and command the future action. The system does not notice and does not react to its own performance.

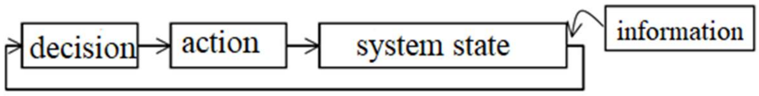
For *example*, a watch is an open system; he does not notice its own inaccuracies to correct it.

The closed system (feed-back system) is characterized by:

- outputs corresponding to inputs into the system;
- outputs depend on inputs;
- outputs influence the inputs.

A closed system is influenced by its own past behavior. On these systems, outputs can adjust inputs. A feed-back system system works like a closed loop that uses the results of the past action of the system to command the next action.

For *example*, a watch and its owner form a feed-back system. When the clock time is compared to the exact time that is taken as a landmark, the clock is fixed to eliminate the errors.



1.1.2 Models

Modeling is a method of studying processes and phenomena that is accomplished by substituting the real object of research. As a research method, it is quite old, the physical models by similarity, then those constructed by analogy, often replacing the actual object under investigation.

A model generally involves the representation of the system as a set of components interacting with one another.

The model can be:

- a duplicate of the system;
- a symbolic (e.g. mathematical) representation of the system;
- the system.

Models are representations of reality. If they were as difficult to handle as reality, their use would not have any advantage. Typically, models which are much simpler than reality can be built. Based on them, we can predict and explain with a high degree of accuracy, complex phenomena. The explanation is that although a large number of variables is required to describe a phenomenon, usually few of these are essential. It is important to find out those variables and the relationships between them.

Mathematical modeling occupies an important place in all modeling methods, especially through the facilities offered by computers with high memory capacity and high speed of work.

Mathematical models have emerged from the need to describe and formally study the behavior of a category of real systems in order to control and direct their future work. The elaboration of a mathematical structure along with a list of correspondences between the mathematical symbols and the objects of the concrete situation considered led to what we call a *mathematical model*.

Generally, *a model M of a system S* is another system *S'* that is equivalent from some points of view with S, but which is easier to study than *S*.

A system *S* is a structure of sets $S = \{T, X, U, V, Y, \varphi, \eta\}$, where

- ✓ *T* is the base *time* used for timing and ordering the events; this is a real number if the system is with continuous time, or integer, if the system is discrete;
- ✓ *X* represents the set of the *input* in the system;
- ✓ *U* is the set of the system input segments.

The *system input segment associated with the function u*: $T \rightarrow X$ is the graph of u on an interval $[t_0, t_1]$, e.g. $u([t_0, t_1]) = \{(t, u(t)) \mid t_0 \le t \le t_1\}$;

- ✓ *V* is *the set of system states*.

State is a concept of modeling the internal structure of the system, which contains its history and which affects its present and future and together with the form of inputs uniquely determines the outputs of the system.

✓ Y is *the set of system outputs*;
✓ φ is *the system response function* $\varphi : X \times V \rightarrow Y$;
✓ η is *the state transition function*.

Knowing the inputs and the corresponding output represents *the system´s behavior*.

A system model must meet the following three conditions:

• the model should reflect as accurately as possible the represented reality;
• the model should be a simplification of the represented reality;
• the model is by its essence an idealization of the represented reality.

1.1.3 Simulation

In the mathematical modeling process, the system components are associated with some variables/parameters, some known (controllable), called *input variables/input parameters*, other unknown (uncontrollable) called *output variables*/parameters. The links and interactions between the system components or the system links with the exterior are transposed into the mathematical model by *functional relationships* (equations and/or identities).

The aim of the model is to express the uncontrollable variables as functions of the controllable variables so that the performance criteria are met. Sometimes it is not possible to express all necessary connections, conditions and interdependencies as equations, which is why some of them are described by logical conditions or procedures that can only be manipulated by the computer.

The mathematical model completed with such procedures is a *simulation model*, which, starting from the values of the controllable variables (generated by special algorithms), will produce values of uncontrollable variables, offering variants from which one can choose the best one. Therefore, the simulation model produces experiments on the system it simulates, which allows the choice of those values of the variables and input parameters that lead to the desired performance.

The need to obtain information about a particular system before it is built has led to the apparition of *simulation*.

In designing systems, it is particularly important to obtain information about the system before it is actually realized; this is possible by applying the simulation technique.

Numerical simulation means the totality of mathematical and computational processes designed to study the real-time behavior of real systems with the help of numerical electronic computers, assuming that the evolution of these systems also involves random elements.

Numerical simulation is a technique that associates real system with an appropriate model called *simulation model*, which is the set of logical interactions of system components, as well as the mechanism of their change over time. The model is then used to produce, through the computer, the chronological sequence of states through which the system passes, considering its initial state.

Because in their evolution real systems are influenced by random causes whose effect must be emphasized in simulation models, one of the mathematical problems of numerical simulation is to generate/simulate with the computer some statistical selections on the different types of random variables and stochastic processes.

Another important issue related to building simulation models is that of accurate timing of simulated system state events using a variable called *the simulation clock* that is subject to a finite number of increments during the simulation.

Although it does not provide accurate solutions, simulation is an effective research technique for the physical phenomena that cannot be perceived by humans and for those perceived but impossible to analyze, as well. The need for simulation lies in the fact that often real systems cannot be studied directly, either because of the qualitative or quantitative difficulties of the phenomena, or because of the complexity (large variables of entry and exit, the large number of possible states, the complexity of the functions φ and η, etc.).

The study of the decision variants using models has the following advantages:

- of an economic nature;

- shortens the time for obtaining the solutions;

- allows the analysis of a large number of variants by changing the initial conditions, having the advantage of returning to the response version required by the user.

Using a real system for experimentation can disrupt the activities of a field in which the system is studied, some less inspired variants may have unpredictable implications.

In the case of systems that do not exist yet, a system-building plan can be obtained based on certain optimization criteria for inputs and/or outputs of the represented system.

For example, if we use the simulation to design a dam, its dimensions and strength can be determined by the computer simulation of a model that provides the requirement of the average electrical power and the random factors such as the volume of rainfall over the established time intervals. Real experiments are not practical because the dam once built cannot be modified.

Simulation also involves some disadvantages:

- building simulation models requires special training; it is said that simulation is rather an art than a science, which is learned in time and by experience;

- simulation results are approximate and not accurate, and sometimes difficult to interpret;

- most system outputs are random variables (based on random inputs) whose distributions must be known or determined.

1.1.4 Types of simulation models

In many scientific areas, three types of simulation models are used:

• imitative models
• analogical models
• symbolic models

Imitative models have the following features:

➢ translate reality to another scale, larger or smaller, in order to observe the behavior of that reality;
➢ imitate the reality, which means that an imitative model resembles the phenomenon it represents, but it differs in size;
➢ is an image of reality.

Imitative models are specific, concrete (physical) and difficult to manipulate for experimental purposes.

Example: building projects, geographic maps, automobile mock-up etc.

Imitative models of the sun and planets are diminished, while atomic models (Bohr's model, for example) are increased.

Analogical models are specific to a process or phenomenon whose behavior is unknown. To be studied, a realistic model of a phenomenon or process that has analogies is used.

Analogical models have the following features:

- use certain properties to represent other properties;
- are less specific, less concrete, but easier to handle than imitative models.

For *example*, one can use:

- models of hydraulic systems for studying electrical or transport systems;
- historical analogies for the prognosis of the development of society in a particular country;
- level curves on a topographic map to represent the height of the relief.

Symbolic models have the following features:

- use letters, numbers or other symbols to represent the characteristics of a reality;
- correlations between the characteristics of reality have led to writing of appropriate mathematical relations and thus to the creation of an abstract (mathematical) model.

A simulation model is a particular type of mathematical model of a system.

Simulation models can also be classified as follows:

• static or dynamic

• deterministic or stochastic

• discrete or continuous.

Static models are those that meet the following conditions:

- do not explicitly take into account the variable time;
- reflects invariant and timeless situations and states;
- solutions can also be obtained analytically.

Dynamical models are those that:

- take into account the variation and the interaction in time of the variables considered;
- incorporates the time as the fundamental characteristic, being a state variable;
- can be solved using the simulation technique.

Deterministic models are those in which:

- all variables are not random;
- operative characteristics are equations of a certain form;
- solutions of these models are obtained in an analytical way.

Stochastic models are those that:

- contain one or more random input variables and so one of the operative characteristics is given by a density function;
- random inputs lead to random outputs;
- events do not happen with certainty, but with a certain probability.

In the case of deterministic models, events occur or do not occur, and in the case of their apparition there is a certainty based on clear rules. In the case of random models, events occur or do not occur, but the model's inference rules do not provide the certainty of their occurrence. These models are typically solved using the simulation technique, the analytical methods being inefficient.

Discrete models are those models in which changes of the variable' states are made at discrete moments of time.

Continuous models are those in which changes in variable states occur continuously.

Examples

1. A plant can produce three types of electronic devices in one month. The number of hours needed to produce one device is: 4 – for the first type device, 2 –for the second type and 3 for the third type. The benefit from the production and sale of these devices is as follows: 50 lei for the first type, 100 lei for the second type and 75 lei for the third type. The sales are limited to 250 devices of the first type, 120 of the second type and 80 of the third type. Knowing that the factory is working 160 hours a month, determine the monthly work schedule, containing the number of devices of each type that will be produced, so that the benefit obtained is maximum.

Denoting by x_1, x_2, x_3 the number of devices of each type, the model can be written as:

$$\begin{cases} 4x_1 + 2x_2 + 3x_3 \le 160 \\ x_1 \le 250 \\ x_2 \le 120 \\ x_3 \le 80 \\ x_1, x_2, x_3 \ge 0 \\ \max(50x_1 + 100x_2 + 75x_3) \end{cases}$$

This is a static deterministic model.

2. *The newspapers salesman problem*

A newspaper vendor has to order daily a number of newspapers so that the income from their sale is maximum. Every day he orders a certain number of newspapers, selling some or all of them. Every newspaper sold brings a certain gain. Unsold newspapers can return, but it causes a certain loss. The number of newspapers sold is variable from day to day, the probability of selling a certain number of newspapers in a day may be estimated on the basis of sales from the previous period.

Let us consider the following variables:

N = the number of newspapers ordered each day;
c = earnings from the sale of a newspaper;
l = loss due to a unsolicited newspaper;
r = demand/number of newspapers sold in one day;
$p(r)$ = the probability to sell r newspapers in one day;
P = gain in one day.

If one day the number of newspapers sold (demand) is greater than or equal to the number of newspapers ordered ($r \ge n$), the earnings will be

$$P(r \geq n) = nc.$$

If one day the number of newspapers sold (demand) is less than the number of newspapers ordered ($r < n$), the earnings will be

$$P(r < n) = rc - (n-r)l.$$

The average daily earning will be

$$\overline{P} = \sum_{r=0}^{n} p(r)[rc - (n-r)l] + \sum_{r=n+1}^{\infty} p(r)nc$$

This model is stochastic because *N* is a controllable variable, *r* is an uncontrollable variable and *c*, *l* are constants. By solving this model, one should find *n* for which $\bar{P}$ is maximum.

1.2. Description of simulation models

In the mathematical modeling process, the model is representative for the physical system if the causality condition is met, which leads to the classification of the elements in:

- *input elements (cause)* that form the input vector $x = (x_1, x_2, \ldots, x_m)$,
- *output elements (effect)* that form the output vector $y = (y_1, y_2, \ldots y_n)$,

both vectors being generally random.

In the absence of any information on the structure of the system, this is mathematically described by the functional dependence between the output vector and the input vector: $y = f(x)$.

Input variables can be:

- *deterministic variables*, which are obtained according to well-defined rules
- *stochastic variables* that are computer generated by performing generation algorithms, the generation depending on the input parameters that characterize these variables.

The simulation step is by definition a stage in which all input variables take constant values during the execution of the program.

Remarks:

- output variables depend on input variables; the dependence is determined by the logical structure of the considered simulation model;
- a value of an output variable is the result of executing a step of the computation program associated with the model;
- if at least one of the input variables is stochastic, then at least one of the output variables is stochastic.

Of great importance in building the simulation model is the process of "moving" the system over time; for this, it is necessary to introduce a special variable called the *simulation clock* to measure the real-time leakage of the system simulation in order to maintain the correct order of events in time.

The simulation models also contain:

- *functional relationships*: identities and/or equations;
- *operative characteristics*, that are used to express by mathematical relations the interactions of the variables and the system behavior.

An operative characteristic is usually a hypothesis (statistical or not) or a mathematical equation that links the input variables with the states system or with the output variables. If these variables are stochastic, the operative characteristics take the form of some probability density functions, and the statistical parameters of the operative characteristics will be found among the input parameters of the model. These parameters act as input variables in the simulation model and must be estimated in advance from statistical observations of the process or system to be simulated.

After building the simulation model, the simulation itself, as an experiment, consists in varying the values of the input variables and parameters of the system and deducting, based on the model, as a result of the calculations, their effects on the output variables.

We distinguish two types of simulation:

- *discrete*, if the model variables can take only certain discrete values;
- *continuous*, if the model variables can have any value at certain real intervals.

In the case of discrete time simulation, the simulation clock advances from one event to another and not continuously. In the case of continuous time simulation, the variables describing the state of the system change their values continuously with respect to the time.

1.2.1 Stages of building a simulation model

Building simulation models is an ample process that generally involves the following steps:

1. *Set-up the problem*, stage in which the objectives of the simulation are set:

- questions to be answered, which need to be clear;
- the assumptions to be tested, which must be accompanied by acceptance criteria;
- the effects to be estimated.

2. *Collecting, analysis, interpretation and primary data processing*

At this stage, the following are established:

- the observation data necessary for studying the system;
- the ways of collecting observation data.

This step is essential because the collection of erroneous data has major consequences in obtaining the final results, which is why a preliminary analysis and their interpretation are necessary for detecting possible inconsistencies with reality. A primary data processing is performed, then data are converted and forwarded to be organized into files for the computer use. Observation data is required to estimate the parameters of the operative characteristics of the model to be built, initialization of the input variables of the model, and its validation.

3. *Formulation of the simulation model*

To build a mathematical simulation model of a system, its components are associated with some variables and parameters, some of which are:

- known (controllable) called *input variables or parameters*;
- unknown (uncontrollable) called *output variables or parameters*.

Interactions between the system components or system connections to the outside are found in the mathematical model in the form of *functional relationships*. Among the model relationships there are one or more functions that bind different variables and measure the system performance.

Because in their evolution the real systems are influenced by random factors whose effect is emphasized in the simulation model, some of the input variables of the model are random variables with known distribution functions. Hence the need for the simulation model to contain routines that generate these input variables.

The simulation model must contain:

- variables that describe the status of the system components (state variables);
- an agenda to memorize the events that occur in the system;
- routines for producing (generating) different types of events.

Building a simulation model differs from one problem to another, which is why some generally valid rules cannot be established. However, there may be some rules that need to be taken into account in building the simulation model.

One of these refers to the number of variables that the model uses; a too large number would create difficulties in establishing functional relationships, would make the model less flexible and the calculation time would be higher. There is no need to reach the other extreme of the exaggerated simplification of the model by using a small number of variables, because in this case it could lose some of the essential aspects of the problem.

Of a great importance in building simulation models is obtaining a low computing time, which allows simulation of different system variants with reasonable cost.

Another requirement to be taken into account when constructing simulation models is the modalities by which the correctness of the model and the variants to be simulated by means of the electronic computer can be verified.

4. *Estimation of the input parameters of the model*

The input parameters of the mathematical simulation model are estimated by statistical methods, using the data collected (in the first stage) about the real system. Operational features may take the form of equations or systems of equations that depend on certain parameters that can be estimated using specific techniques of regression analysis.

5. *Evaluation of the model performance and parameter testing*

This step aims at verifying the model before its programming:

- verifying that the input parameters of the model were well estimated using statistical tests;
- checking that the model contains all the essential variables and parameters as well as the functional relationships necessary to represent the essential interdependencies of the real system.

If the operative characteristics take the form of statistical hypotheses concerning the distributions of the input variable, then the matching tests (chi-squared, Kolmogorov-Smirnov test) are applied to verify these hypotheses. If these checks reveal that a question or hypothesis is not correctly formulated, it means that either the variables and parameters were not well-chosen or the input parameters were not well estimated. If there are other inconsistencies in the model, then all previous steps will be resumed in order to correct them.

6. *Description of the simulation algorithm and writing the calculation program*

Based on the results of the previous steps, the calculation algorithm is constructed which represents the logical sequence of the events to be reproduced with the electronic computer. To make it easier to program, the algorithm is represented by a logical schema; program writing follows either by using a high-level programming language or by a special simulation language.

The choice of the programming language depends on several factors, including: the simulation time required, the form in which the simulation results must be printed out, the programmer's experience, etc.

Simulation languages (specialized) make much easier a system description and its behaviour application over time. They can greatly facilitate modeling, which makes them superior from this point of view to general programming languages.

7. *Model validation*

The model validation, that is, establishing its suitability for reality, is usually a complex and difficult task. The value of a model in relation to its contribution to the study of the concrete modeled situation is determined by its degree of adequacy, i.e. by the way the predictions match the observations.

Validation methods for mathematical simulation models are not unique.

Model validation can be done by:

- testing the model in a particular case where the solution is known or can easily be deduced by analytical means;
- comparing the simulation results with the data obtained by observing similar systems or by comparison with the past evolution of the real system that was simulated.

Model variants that turn out to be inappropriate are modified to get solutions that are consistent with the reality.

8. *Planning the simulation experiments*

At this stage, we assign the values of the input variables and parameters that cover the real situations of the system, for selecting the variant that meets the user's requirements.

9. *Analysis of simulated data*

The results of the simulation tell us what the system's "reaction" is to changing the values of the input variables, and more, we will look for the answers to the questions asked at the beginning. This is possible by collecting simulated data, processing it by calculating the statistics for the significance tests and then interpreting only the results.

1.2.2. The simulation clock

Through the simulation model, the electronic computer produces successively different events that represent the changes that take place over time within the system. In order to be able to maintain the correct order of these events and to specify, after each step of the simulation, the timeframe in which the system evolution was simulated at that step, it is necessary to introduce into the simulation model a special variable called clock. At every step of the simulation a "clock" increase must be generated that adds to the clock's magnitude recorded at the previous step.

There are two types of clock: with a *fixed* (constant) *increment* and with a *variable increment*.

The simulation based on the constant clock method consists in generating a constant increase of the clock every time and then analyzing the state of the various elements of the system generating all the possible events occurring in the time interval of length c. Then a new increase will be generated, that will be added to the clock and the mentioned analysis is repeated.

The logical scheme of the simulation model must then fully describe the evolution of the system over a time interval of length c; simulation of the system over a large interval of time will be obtained by repeating a number of times the algorithm with respect to the length time interval *c.*

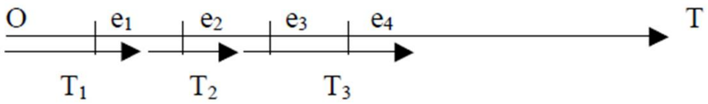
So for constant clock simulation models, the size of the clock T is of the form:

$$T = c \cdot j \ (j = 0, 1, 2, \ldots)$$

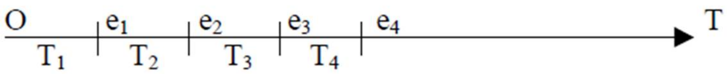
where j is an integer representing the number of iterations of the simulation algorithm.

In the case of *the variable clock*, the (variable) value of the clock increase is equal to the length of the time interval between the occurrences of two consecutive events. In other words, the magnitude of the clock's increase is equal to the time interval from the current state to the moment of occurrence of the nearest future event.

In the case of the constant clock model, we have:



In the case of the variable clock model, we have:



The variable clock method rigorously assumes the order of all occurrences of successive events, so that each new occurrence corresponds to an increase in the clock.

In the case of constant clock models, for example, if the clock value at a moment is $T = ck$, then in this phase the simulation algorithm will generate the events to occur within the time interval $[c(k\text{-}1), ck)$ and then advance the clock to $T = c(k + 1)$.

A simulation model based on the constant clock method considers the group of events produced in the interval $[c(k\text{-}1), ck)$ as if it had occurred at the moment ck. Consequently, the method based on the constant clock method, unlike that based on a variable clock, causes the group of events occurring over a time interval of c to be synchronized at the end of that interval. For this reason it is advisable to choose the constant as small as possible. This will increase the computing time.

On the other hand, increasing the constant c due to the synchronization of events that occur at a big interval of time will increase the approximation of the model and will reduce the calculation time.

The constant clock method is preferable to that of the variable clock, especially in terms of the ease with which the simulation algorithm can be constructed.