Chapter 2
SimExplorer: Programming Experimental Designs on Models and Managing Quality of Modelling Process

Florent Chuffart
Cemagref- Laboratoire d'Ingénierie des Systèmes Complexes, France

Nicolas Dumoulin
Cemagref- Laboratoire d'Ingénierie des Systèmes Complexes, France

Thierry Faure
Cemagref- Laboratoire d'Ingénierie des Systèmes Complexes, France

Guillaume Deffuant
Cemagref- Laboratoire d'Ingénierie des Systèmes Complexes, France

ABSTRACT

This article describes Simexplorer, a computer framework for managing simulation experiments and, to some extent, the scientific quality of the modelling process. An information system, included in the framework, insures the traceability of the experiments and their reproducibility and thus contributes to the modelling process quality management. Moreover, this information system provides facilities for sharing and exchanging components of experiment scenarios. The authors illustrate the use of the framework on a simple example of modelling process.

INTRODUCTION

The theory of design of experiments (DoEs) has been initially developed for real, non-simulated experiments in agriculture in the 1920s (Chen (2003) refers to 1926 publication by Fisher), and in engineering, psychology, and so forth, since the 1950s. In real experiments it is impractical to investigate many factors; ten factors seems a maximum, because otherwise the number of experiments to perform becomes too high. Moreover, it is hard to experiment with factors that have more than a few values.

DOI: 10.4018/978-1-4666-0333-2.ch002

These restrictions do not apply when experimenting on models (simulations). Indeed, computer codes may have hundreds of inputs and parameters each with many values. The experiment is a simulation, which generally has a very low cost, compared with real experiments. Consequently, a multitude of scenarios may be simulated. Therefore a change of mindset of simulation experimenters took place, to adapt theory and practice to this new situation (Kleijnen et al., 2005).

In this context, it is particularly important to easily access the advances of the research on simulation experiments. The goals of such a numerical experiment can be (Oden, 2006; Bayarri et al., 2007)

- Calibration and validation,
- Sensitivity Analysis—either global or local—or "What If" analysis,
- Optimization,
- Risk Analysis.

Each of these types of experiments requires particular methodological approaches, and specific experimental designs. Using the adequate method is also a strong requirement to insure a good scientific quality of modelling.

The practise of experiments on models is already widespread, and will most probably keep increasing in the future. Indeed, we notice the following evolutions of modelling processes:

- The increasing computing power makes it possible to develop more and more complex models, coupling sub-models of different types representing industrial, natural or social processes, or representing explicitly all the individuals of populations (individual or agent based models).
- Such complex models generally lack an analytical theoretical framework that fits them. Therefore, it is necessary to explore experimentally their behaviours and confront them with observation data or expertise.
- In this context, the modelling process involves loops where simulation experiments on the model suggest new hypotheses and possible changes, leading again to experiments.
- These experiments are performed through a specific application, developed for the particular model and generally poorly reusable in other modelling processes. Moreover, keeping the trace of all the modelling steps becomes rapidly difficult without a specific tool. Yet, this trace is often necessary to justify the modelling choices, and to insure the scientific quality of the process.
- There is a growing requirement of Reproducible research (RR) (Buckheit & Donoho, 1995), stating that all the results of a paper can be reproduced. This means that the whole simulation process is reproducible.

Last, the increasing complexity of the models implies frequently a team development, which enhances the necessity to take into account a multi-user context. This multi-user perspective is also important for sharing methods and libraries which are connected to different types of experiments.

SimExplorer aims at addressing the needs appearing in this context, in particular, to provide:

- A programming environment devoted to the development of simulation experiments, that can be used with any executable model,
- An easy access to well established libraries for experiment designs and corresponding data treatment,
- An information system, in a client server configuration, providing the possibility for sharing or exchanging components of experiments, and for tracing them.