



Design Of Experiments for Complex (and complicated!) Models

14-15/10/2015

IMF-CSIC
c/Egipciaques, 15 - Barcelona



@MariaPereda_
#DOECM
#SIMULPAST

The SimulPast - CONSOLIDER project is organising, the 14th and 15th of October 2015, a two-day tutorial on the analysis of complex simulation models. María Pereda, a postdoctoral researcher of Group 10 (University of Burgos) and working with Case Study 3, will teach you how to apply statistics and machine learning to analyse your models.

Motivation

The Ockham's razor principle is generally accepted when producing a model but this does not imply that your models are going to be always *stupidly simple*. It may happen that your model has a large number of parameters. Even though these would not be in the hundreds, this situation "many parameters" situation is more common than we would expect. For the purpose of this course, we will consider that a model has a large number of parameters when they are more than five. It may also happen that it is difficult to establish a direct correspondence between parameters' values and the reality under study. These situations make the analysis of the influence of the parameters in the model output really complicated.

Using statistical sampling and machine learning techniques can help us to attempt a first approach to the understanding of the implications of the parameters on our model, and to guide the conception of more profound and design-detailed experiments. Moreover, we need to ensure that we are covering the whole parameters' space in our exploration, which is not always the case with traditional approaches, such as Monte Carlo sampling.

In this course you will learn how to design experiments for a model with a large number of parameters using advanced statistical sampling and to analyse simulations results with machine learning techniques. The course is divided in a theoretical and a practical part. For the practical part, we will use NetLogo and R programming languages. Although it is not necessary to have programming background to understand the course (the code is going to be provided), some background would be of help for better participation in the practical session.



www.simulpast.es

Practical information and Schedule

14/10/2015 Day 1	First day is theoretical basis.
9:30 – 11:00	Session 1: 1. Introduction. a. Objectives of model analysis b. Analysis of complex models 2. Design of experiments. a. Definition b. Why? c. Sensitivity Analysis. Limitations of traditional approach
11:00 – 11:30	Coffee break.
11:30 – 13:00	Session 2: 3. Sampling of the parameters' space. a. Monte Carlo sampling b. Latin Hypercube sampling
13:00 – 14:00	Lunch break.
14:00 – 15:30	Session 3: 4. Brief intro to Machine Learning techniques. a. CART: Classification And Regression Trees b. Random Forests c. Application to the study of a model
15:30 – 16:00	Coffee break.
16:00 – 17:30	Session 4: (continued) Brief intro to Machine Learning techniques.
15/10/2015 Day 1	Second day is hands-on. We will learn how to develop these approaches using R and NetLogo with a toy-guided example. The code will be provided.
9:30 – 11:00	Session 6: Hands-on.
11:00 – 11:30	Coffee break.
11:30 – 13:00	Session 7: Hands-on.
13:00 – 14:00	Lunch break.
14:00 – 15:30	Session 8: Informal session with María. In these informal sessions (8 and 9), we will discuss any doubt you may have about the course, the techniques, other related models, or applications of machine learning you may be curious about. Feel free to ask whatever question you have, and I will do my best to answer you all. Let's have some fun talking!
15:30 – 16:00	Coffee break.
16:00 – 17:30	Session 9: (continued) Informal session with María.

This is a two days course that will be held at IMF-CSIC, Egipcíaques 15, Barcelona, 08001.

Places are limited to **15** participants and these will be allocated on a first-come first-served basis. Members of the SimulPast project have precedence, but the event is open to other people as well. Anyone interested in participating should send an email (with #DOECM as subject) to simulpast@gmail.com by **Friday 9th** October 2015.

For the practical part (second day), participants are encouraged to bring their own laptop with R, R Studio and NetLogo installed. We will also provide the necessary installation executables for those who will need them.