Replication of Computational Results Report for "*Green Simulation with Database Monte Carlo*"

ALESSANDRO PELLEGRINI, Sapienza, University of Rome, Italy

This article presents the reproducibility results associated with the article "*Green Simulation with Database Monte Carlo*" by Mingbin Feng and Jeremy Staum. The authors have uploaded their artifact to Zenodo, which ensures a long-term retention of the artifact. The artifact, which is based on a set of R scripts, allows to easily regenerate data for the figures and the tables, it completes successfully, and allows to reproduce all the experimental results in the article.

The article can thus receive the *Artifacts Available*, the *Artifacts Evaluated—Functional*, and the *Results Reproduced* badges.

CCS Concepts: • **General and reference** → Empirical studies; *Experimentation*; **Validation**; **Verification**.

Additional Key Words and Phrases: Replication of Computational Results, ACM TOMACS

1 INTRODUCTION

The paper *Green Simulation with Database Monte Carlo* by Mingbin Feng and Jeremy Staum [1] tackles simulation scenarios in which multiple simulation runs are executed, using the same model. The proposed approach stores the results of previous experiments in a database, in order to reuse them in future runs, to keep the execution time of the simulation low and potentially under a fixed computational budget. In the article, the authors show that using idle time to create databases of simulation output allows to reduce the variance in the experiments.

2 REPLICATION OF COMPUTATIONAL RESULTS

This section documents the reproducibility process, and compare the reproduced results with the data taken from the original article. The authors have given permission to use their original results in this report.

2.1 Software download and installation

The authors have provided a link to a permanent repository on Zenodo—the artifact is located at http://doi.org/10.5281/zenodo.3905334—making it permanently available (*Artifacts Available badge*). The artifact is composed of a set of R scripts, which can be sourced in the R interpreter from the command line.

Author's address: Alessandro Pellegrini, pellegrini@diag.uniroma1.it, Sapienza, University of Rome, Rome, Italy.

The amount of dependencies required to run the artifacts is extremely reduced, and typical of an R environment. In particular, to run the provided scripts, the following R packages must be installed: `randtoolbox`, `pracma`, `ggplot2`, and `scales`.

2.2   Quality of the artifact

The artifact is relevant to the associated article, as it allows to re-generate all the tables and figures, and it is easy to use for reproducibility purposes thanks to the availability of a couple of scripts (`File2_Example1_Uniform.R` and `File2_Example2_Reinsurance.R`) which, when sourced in the R interpreter, run all the experiments, generate the figures, and dump on the shell the values of the tables. No component required to re-run the experiments and generate the final plots is missing (*Artifacts Evaluated—Functional*).

2.3   Replicating the experiments

Replication of the experiments has been carried out on a machine equipped with a quad-core Intel® Core™ i7-7600U CPU, equipped with 16 GB or RAM, running Arch Linux with kernel version 5.7.5, and using R 4.0.2.

The artifact is easy to run, and all data required to support the reproducibility are provided (*Artifacts Evaluated—Functional*). To re-run the experiments, around one hour is required—the execution is single threaded, but the two scripts can be launched in parallel. The experiments have been reproduced using both the same random seed used in the original article, and using different random seeds.

This section provides a short description of the reproducibility results. There is a total of 3 Figures and 2 Tables in the original paper to be reproduced—Figures 4 and 5 in the original article are only visual representations supporting the text. The figures provided in this report are related to one execution using different random seeds—of course, using the same seed as in the original experimentation yields perfectly similar results.

(a) Original.   (b) Reproduced.
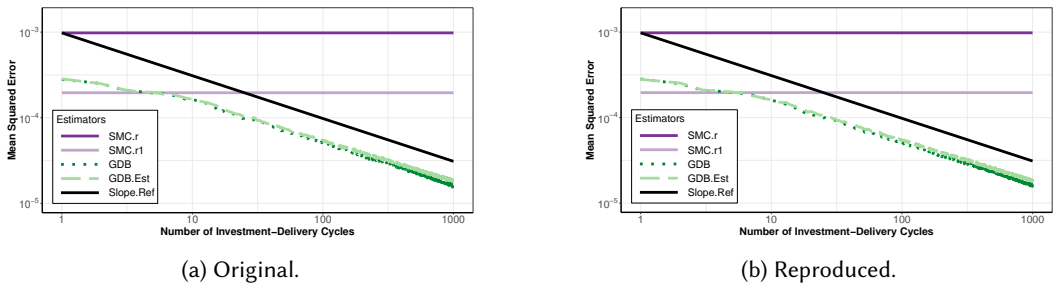
Fig. 1.   Log-log plot of estimated MSEs for different estimators in tail probability estimation.

The results in Figure 1 are comparable. They report the estimated Mean Squared Error (MSE) for the estimators presented in the article, with data gathered from 1,000 simulation iterations. Quantitative values associated with the estimators are also reported in Table 1. By the results, we can observe that the largest deviation from the original results is in the order of 10^{-5}, which is perfectly acceptable, considering the different random seeds used and the somewhat limited number of iterations.

Figure 2 shows the results associated with the estimation of c_1 (one of the properties of the underlying random field of the simulation model) and α (another property, the convergence factor) over 10,000 sample paths of repeated experiments. The comparison between the original and the

Table 1. MSEs in the n-th repeated experiment for different estimators, expressed in 10^{-4}.

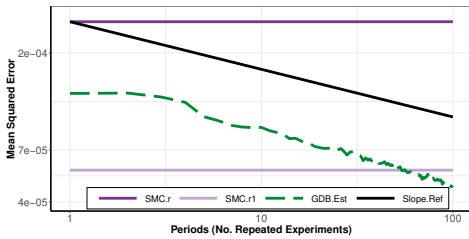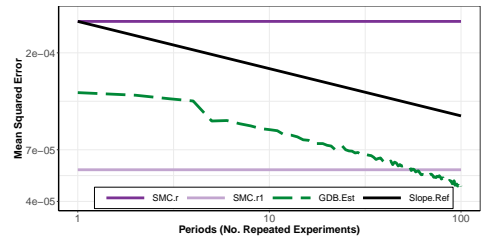| Original.						Reproduced.				
	$n = 1$	$n = 10$	$n = 100$	$n = 1000$			$n = 1$	$n = 10$	$n = 100$	$n = 1000$
SMC.r	9.80	9.80	9.80	9.80		SMC.r	9.80	9.80	9.80	9.80
SMC.r1	1.96	1.96	1.96	1.96		SMC.r1	1.96	1.96	1.96	1.96
GDB	2.81	1.64	0.52	0.16		GDB	2.83	1.60	0.51	0.16
GDB.Est	2.85	1.64	0.54	0.18		GDB.Est	2.80	1.62	0.55	0.19
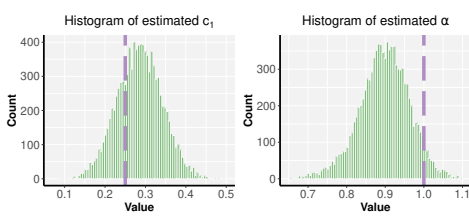
(a) Original.
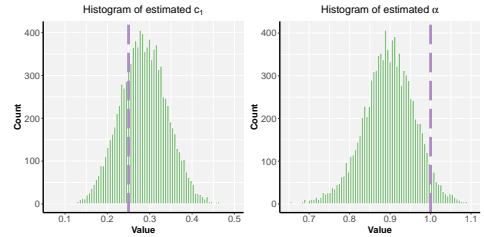
(b) Reproduced.

Fig. 2. Histograms of estimated parameters in the GDB procedure. True values derived from underlying random field are shown as dashed purple line in each histogram.

reproduced results is also fine. We observe a slight change in the shape of the distribution, but the deviation in the maximum is around 7%. Given the stochastic nature of the experiments, this is perfectly acceptable.

(a) Original.

(b) Reproduced.

Fig. 3. Log-log plot of estimated MSEs for different estimators in CAT bond pricing.

Table 2. MSEs in the n-th repeated experiment for different estimators, expressed in 10^{-4}.

Original.					Reproduced.			
	$n = 1$	$n = 10$	$n = 100$			$n = 1$	$n = 10$	$n = 100$
SMC.r	2.80	2.80	2.80		SMC.r	2.80	2.80	2.80
SMC.r1	0.56	0.56	0.56		SMC.r1	0.56	0.56	0.56
GDB.Est	1.29	0.89	0.47		GDB.Est	1.30	0.87	0.47

Also, the results in Figure 3 are perfectly comparable. They report the estimated MSE for the SMC and the GDB.Est estimators, and the black line's slope is computed as the average of the estimated α's. Quantitative values associated with the estimators are also reported in Table 2. As it can be observed, the trend of the curves is very similar, and the quantitative results show a largest deviation from the original results in the order of 10^{-5}.

Overall, the plots and the tables which have been generated are consistent with the original publication (*Results Reproduced*).

REFERENCES

[1] Mingbin Feng, and Jeremy Staum. 2020. Green Simulation with Database Monte Carlo. *ACM Transactions on Modeling and Computer Simulation* 31, 1 (Dec. 2020)