

# Optimization and Simulation Methodology Applied to Improve a Computational Simulator

Adriana Gaudiani

National University of General Sarmiento, Buenos Aires,  
Argentina  
agaudiani@campus.ungs.edu.ar

Emilio Luque

Dolores Rexachs  
emilio.luque@uab.es  
dolores.rexachs@uab.es

Computer Architecture and Operating Systems  
Department, Universitat Autònoma de Barcelona,  
Bellaterra, Barcelona, Spain

## ABSTRACT

In our research, we elaborate a search and optimization methodology to achieve a higher prediction quality of a computational model of a physical system. A problem that simulators of physical systems must address is the uncertainty of their parameters and their impact on the output results, causing prediction errors. We addressed this problem by performing a two-phase optimization using a simulation methodology. This proposal was able to achieve this goal using the least computational resources. The average percentage improvement obtained in the simulation error compared to those of the untuned simulator reached 60%.

## KEYWORDS

optimization and searching, simulator calibration, adjusted parameters, parametric simulation, Monte Carlo-KMeans heuristic

### ACM Reference Format:

Adriana Gaudiani, Emilio Luque, and Dolores Rexachs. 2024. Optimization and Simulation Methodology Applied to Improve a Computational Simulator. In *Proceedings of June 26–28, 2024 (11th ACM Celebration of Women in Computing: womENCourage 2024)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 1 INTRODUCTION

Computational models that simulate natural phenomena are designed to behave as closely as possible to the real system. The simulated data differ from the real data due to multiple factors. One problem that these numerical models must handle is the parameter's uncertainty [3]. The lack of information or unavailability of some input data causes uncertainty in the values of the model parameters and their impact on output results, causing significant prediction errors [5] [6].

We proposed a simulation methodology implemented in two stages for tuning a computational model with the lowest possible computing cost. We applied computer science methods and parallel programming techniques to reduce the computational cost and

energy expenditure when they are combined. Our proposal was applied to a river channel model. In particular, this simulation model simulates the Parana River channel in Argentina [7].

An exhaustive search always guarantees finding a solution, if a solution exists. The computational cost is high because it is proportional to the search space dimension [4]. This option is often impractical. Therefore, an alternative optimisation technique is required to find an approximate good solution, which efficiently uses computing time and resources. Our work proposed an optimization and simulation algorithm. This is a two-stage methodology to get this good solution. We develop a tuning function that measures the performance of the system. We established a quality index for each simulation scenario resulting from some combination of the adjustment parameters. This index gives us a measure of the error (RMSE) between the simulated data and the real data. This optimization method searched for a local minimum based on these indices.

In the first stage, we used a computational process based on an iterative Monte Carlo-type algorithm combined with a K-Means clustering method that allowed us to identify the promising regions for optimization based on the neighbourhood structure of the problem aiming to reduce the feasible region [1]. In the second stage, we used an exhaustive reduced search on the promising areas obtained in the first stage. This gave us a configuration of fitted parameters that best fit the observed data. Our methodology found a local optimum that allowed us to achieve successful results and the best agreement between observed and simulated data. We explain exhaustively in [2].

We want to be clear that the simulator is considered a black box and when a simulation is started, it is fed with a complete simulation scenario containing the input data and parameter values. On the other hand, computing the optimal set of points is a combinatorial problem. It grows exponentially with the number of parameters.

## 2 COMPUTATIONAL MODEL TUNING

We verified our tuning methodology using a river simulation model, specifically a simulator of the Parana River channel in Argentina [7]. The domain of the river channel is divided into 76 sections. We have daily water height data (observed data) at the 15 monitoring stations along the Paraná River channel. The output data is the simulated height of the water for each simulation day and at each monitoring station. The model parameters must be set in each of the 76 sections to create each of the simulation scenarios. When we select 3 adjustment parameters for the 76 sections and select a range of 3, 4 and 4 different values respectively, we obtain 110,592

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*11th ACM Celebration of Women in Computing: womENCourage 2024, Madrid, Spain*,  
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM  
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

configurations, then we are faced with a combinatorial explosion of scenarios. Under these conditions, an exhaustive search becomes an intractable computational problem. The experimental results obtained using this tuning method are presented in the following section. Our two-phase methodology is an optimization scheme formulated to get a "good" configuration that optimizes this problem. We defined a function which calculates the quality index for each scenario. The objective function is formulated to search for the best parameter configuration that minimizes simulation error.

We developed a fit function to calculate the quality index,  $Iq$  of each scenario and each  $\vec{x}$  is a combination of  $n$  parameter values in the domain  $\Omega \in \mathcal{R}^n$ .

$$\begin{aligned} & \text{minimize} && Iq : f(\vec{x}) \\ & \text{subject to} && \rightarrow x \in \Omega \end{aligned} \quad (1)$$

In the second stage of our research, we combined the quality index  $Iq$ , of each simulation scenario with the number of cities (monitoring stations) along the riverbed that improved with each tuning execution scenario.

### 3 EXPERIMENTAL RESULTS

We selected the following experimentation to display our successful results. We configured the process for a river simulation period of 360 days. The search space was defined with 3 adjustment parameters. The different configurations are set by combining the range values of the parameters. Two parameters can take 3 values and one 4. The methodology described above provided the best parameter configuration (a scenario) for each location under the conditions established for the experimental case. The optimization methodology found the configurations that improved the prediction for 3 cities, Goya, San Martín and Rosario. In the last stage of the methodology, by selecting the best group of configurations and carrying out the exhaustive search in the reduced space, the improvement was 30-60%. The improvements achieved in these cities are detailed in Table 1. The details of this experience are explained in depth in [2]. We compared the root mean square error obtained when applying our tuning methodology to the simulator with the simulator without tuning.

1

**Table 1: RSME and Gain achieved with the optimal configuration**

Station	Scenario		Gain
	tuned Sim.	untuned	
Goya	0.472	0.782	40%
San Martín	0.274	0.686	60%
Rosario	0.327	0.485	33%

### 4 CONCLUSIONS AND FUTURE WORK

Our work provides a methodology that does not require knowledge of the simulated system and enables its application to other simulations of physical systems. It is only necessary to know their adjustment parameters and the range of values of each one.

We could have launched a parametric simulation by running the simulator with each parameter value combination in an exhaustive

way. This would be an easy strategy to carry out because we can look for the computing resources necessary to speed up the execution of this process. We can turn to cloud resources if we can pay for them. Our methodology uses computer science and efficient computing to reduce the search space for tuning parameters. This is a solution that avoids 110,592 unnecessary simulator runs, only one of which matters, saving significant energy costs. Finally, we only need 5392 simulator runs. The total gain is measured in the number of simulator runs and exceeds 20X as shown in Figure 1.

Exhaustive search solution (Simulations)	Resulting simulations by Monte Carlo (Reduced)	Resulting simulations by K-Means (Reduced)	Search over the reduced search space (Reduced)	Sum of simulated scenarios
110.592	5200	8	192	5200+192

**Figure 1: Tuning methodology improvement**

These results motivated us to advance our research by reducing the search space for parameters, for which we took advantage of the knowledge of the system. The values of the parameters maintain similarities in adjacent areas of the river, which is why they should not present great variation between them. This information is included in new research on the topic.

### ACKNOWLEDGEMENTS

This research has been supported by the Agencia Estatal de Investigación (AEI), Spain and the Fondo Europeo de Desarrollo Regional (FEDER) UE, under contract PID2020-112496GB-I00 and partially funded by the Fundacion Escuelas Universitarias Gimbernat (EUG).

### REFERENCES

- [1] K.D. Berends, J.J. Warmink, and S.J.M.H. Hulscher. 2018. Efficient uncertainty quantification for impact analysis of human interventions in rivers. *Environmental Modelling & Software* 107 (2018), 50 – 58. <https://doi.org/10.1016/j.envsoft.2018.05.021>
- [2] Adriana Gaudiani, Alvaro Wong, Emilio Luque, and Dolores Rexachs. 2023. A computational methodology applied to optimize the performance of a river model under uncertainty conditions. *The Journal of Supercomputing* 79 (2023), 4737–4759. <https://doi.org/10.1007/s11227-022-04816-6>
- [3] S. K. Jain and V. P. Singh. 2019. *Hydrological Cycles, Models, and Applications to Forecasting*. Springer Berlin Heidelberg, Berlin, Heidelberg, Chapter Springer Berlin Heidelberg, 311–339. [https://doi.org/10.1007/978-3-642-39925-1\\_20](https://doi.org/10.1007/978-3-642-39925-1_20)
- [4] Farkhondeh Khorashadi Zadeh, Jiri Nossent, Befekadu Tadesse Woldegiorgis, Willy Bauwens, and Ann van Griensven. 2019. Impact of measurement error and limited data frequency on parameter estimation and uncertainty quantification. *Environmental Modelling & Software* 118 (2019), 35–47. <https://doi.org/10.1016/j.envsoft.2019.03.022>
- [5] Hilary K. McMillan, Ida K. Westerberg, and Tobias Krueger. 2018. Hydrological data uncertainty and its implications. *WIREs Water* 5, 6 (2018), e1319. <https://doi.org/10.1002/wat2.1319> arXiv:<https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wat2.1319>
- [6] G. Merkurjeva, Y. Merkurjev, B. V. Sokolov, S. Potryasaev, V. A. Zelentsov, and A. Lektuers. 2015. Advanced river flood monitoring, modelling and forecasting. *Journal of Computational Science* 10 (2015), 77 – 85. <https://doi.org/10.1016/j.jocs.2014.10.004>
- [7] Mariano Re, Martin Sabarots Gerbec, and Lucas Storto. 2015. Estadística de Niveles en el Delta del río Paraná mediante Modelación Hidrodinámica. In *Séptimo Simposio Regional sobre Hidráulica de Ríos - Uruguay*. [https://www.researchgate.net/publication/284728090\\_Estadística\\_de\\_Niveles\\_en\\_el\\_Delta\\_del\\_río\\_Parana\\_mediante\\_Modelación\\_Hidrodinámica](https://www.researchgate.net/publication/284728090_Estadística_de_Niveles_en_el_Delta_del_río_Parana_mediante_Modelación_Hidrodinámica)