USING SIMULATION AND RELIABILITY CONCEPTS TO SET STARTING TIMES IN MULTI-TASK PROJECTS WITH RANDOM DURATION AND A COMMON DEADLINE

Angel A. Juan[(a)], Laura Calvet[(a)], Carles Serrat[(b)], Sara Hatami[(a)]

[(a)]IN3-Computer Science Dept., Open University of Catalonia, Barcelona, Spain
[(b)]Dept. of Mathematics, Technical University of Catalonia, Barcelona, Spain

[(a)]{ajuanp, lcalvetl, shatami}@uoc.edu, [(b)]carles.serrat@upc.edu

ABSTRACT

In production and transportation logistics, it is frequent to have projects in which different tasks must be processed in parallel and completed before a common deadline. Usually, it is not desirable to assign scarce resources to a new project earlier than strictly necessary, since they might be currently occupied in other projects. When these tasks have random completion times, some natural questions arise: (*i*) how much can we delay the starting time of each task in the new project so that it can be completed by the deadline with a given probability? and (*ii*) how to compute these 'optimal' starting times when not all tasks need to be necessarily finished to consider the project as completed? This paper proposes a hybrid approach, combining reliability concepts with simulation and optimization techniques, to support decision makers in finding the optimal starting times for each task under the described uncertainty scenario.

Keywords: project scheduling, reliability analysis, simulation, optimization

1. INTRODUCTION

In a global economy, production and transportation logistics are becoming increasingly complex due to the existence of joint projects where different departments (or even companies) cooperate to manufacture a final product or provide a given service. A project can be frequently decomposed as a set of independent and parallel tasks (e.g., production lines), some of which need to be finished in order to consider the project as completed on a given deadline (Figure 1).

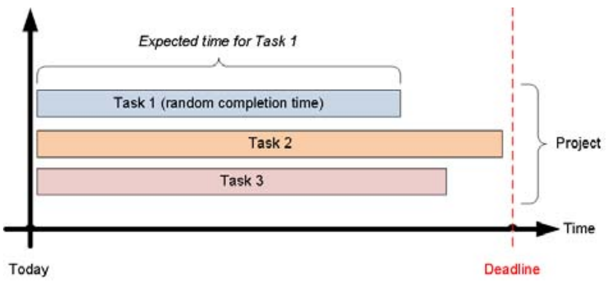

Figure 1: A Simple Project with Three Parallel Tasks

Notice that finishing all the tasks in a project might not be a necessary condition in order to consider the project as completed. In effect, depending on the final product or service specifications there might be a certain degree of redundancy in these tasks, which allows the project to be completed –at least partially– if some combinations of these tasks are satisfactorily concluded by the deadline. Thus, for instance Figure 2 shows a simple logical representation of a project ending condition: the associated project will not be terminated (i.e., it will "survive" the deadline) as far as any of the following combinations occur: (*i*) Tasks 1 and 2 are still not accomplished by the deadline; or (*ii*) Tasks 1 and 3 are still not finished by then. Thus, while in classical reliability analysis the goal is to increase the probability that a system survives a given target time, in our study the goal will be related to reducing the probability that a project survives a user-defined deadline.

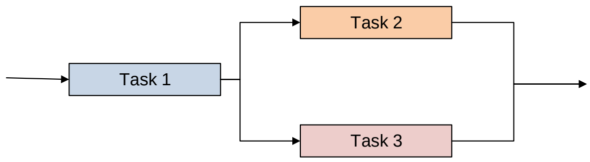

Figure 2: Logical Representation of Ending Condition

From a managerial perspective, it is usually not efficient or desirable to assign scarce production resources to a new project earlier than strictly necessary, especially if these resources are currently occupied in other projects. Thus, a manager might be interested in finding the starting time for each task in the new project that maximizes the total delay time while ensuring that the project will be completed by the deadline (Figure 3). Also, the manager might be interested in analyzing how these starting times might vary as different combinations of concluded tasks account for the project termination. In this paper, we will assume that each task is a process which duration can be modeled as a random variable. Since the problem becomes stochastic, the goal will be to maximize total delay time subject to achieving project competition by the deadline with a user-specified probability.
To support the manager during this complex decision-making process, we propose a hybrid approach in which simulation is combined with reliability analysis and optimization. One of the main benefits of our approach is that it does not make any assumption on the probability distributions employed to model the random

duration times associated with each task, i.e.: being based on simulation, there is no need to assume exponential or Normal times –instead, any empirical or theoretical distribution based on historical data can be employed to model the behavior of these times.

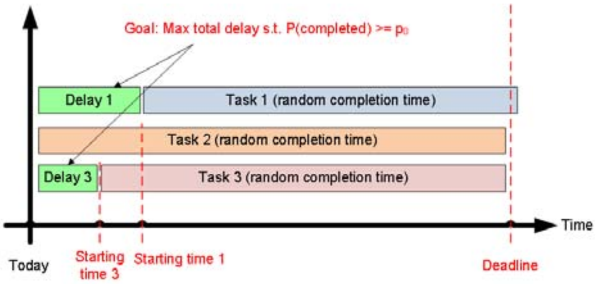

Figure 3: Determining Starting Times

The rest of the paper is organized as follows: Section 2 provides a review of related work on the use of simulation methods in reliability analysis. Section 3 describes the main ideas behind our hybrid simulation-optimization approach. An illustrative numerical example is discussed in Section 4, and the obtained results analyzed in Section 5. Finally, we conclude this paper by summarizing its main findings and highlighting future research lines in Section 6.

2. SIMULATION IN RELIABILITY ANALYSIS

Reliability or survival analysis of time-dependent systems is a research area with applications in engineering (Skrzypczak et al. 2017, Zheng and Chang 2017) as well as in experimental and social sciences (Topaloglu et al. 2016, Levin and Kalal 2003). Many works have discussed the benefits of maintenance policies in systems reliability. Some of these works highlight the fact that system management concepts, such as aging, repair obsolesce and renovation, are not easily captured by analytical models (Borgonovo et al. 2000).

As has been pointed out by many authors, when dealing with real-life complex systems only simulation techniques –such as Monte Carlo simulation (MCS) or discrete-event simulation (DES)–, can be useful to obtain credible predictions for system reliability and availability parameters (Billinton and Wang 1999). In fact, simulation has been revealed as a powerful tool in solving many engineering problems (Grasas et al. 2016, Sobie et al. 2018). This is due to the fact that simulation methods can model real-systems behavior with great detail (Dubi 2000). In addition, simulation methods can provide supplementary information about system internal behavior or about critical components from a reliability point of view. Applications of simulation techniques in the reliability field allow to model details such as component dependencies, dysfunctional behavior of components, etc. (Labeau and Zio 2000, Labeau and Zio 2002). Examples of works which deal with applications of simulation techniques in reliability are the ones of Barata et al. (2001) –who employs MCS in the modeling of components' degradation processes taking place in nuclear power plants–, or Barata et al. (2002), who use MCS in the modeling of repairable multi-component deteriorating systems (e.g., offshore structures and aerospace components). Juan and Vila (2002) proposed a MCS approach for determining the reliability of complex systems, while Faulin et al. (2007, 2008) proposed both MCS and DES algorithms for dealing with reliability and availability issues in complex telecommunication networks. Similarly, Cabrera et al. (2014) proposed a hybrid simulation-optimization approach for optimizing the availability of distributed computer systems. For additional examples of works linking simulation and reliability, the reader is addressed to Faulin et al. (2010).

3. OUR SIM-OPT SOLVING APPROACH

Following our discussion in the Introduction section, and under an uncertainty scenario in which the completion time associated with each task is a random variable, our goal is to maximize the total delay time – sum of tasks' delay times– subject to satisfying a user-specified probability of completing the project on or before a given deadline. In our approach, we assume that the specific probability distributions modeling each task's completion time is known. This could be achieved, for instance, by fitting historical data on the duration of each task by the most appropriate probability distribution –in case no theoretical distribution offers a reasonable fitting, then an empirical distribution could be used instead. Then, for each task, it is possible to plot its survival or reliability function based on the fitted probability distribution. Notice that the introduction of a delay in the starting time of task i shifts the associated survival function to the right, thus increasing the probability that task i cannot be terminated before the common due date (Figure 4).

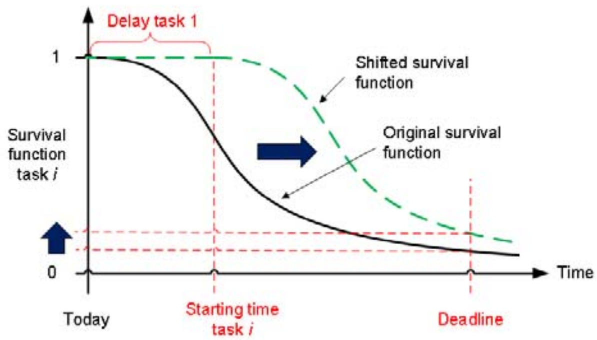

Figure 4: Effect of Delays on Task's Survival Functions

Thus, the question for the decision maker is how much can we extend these delays (i.e., increase the reliability level of each task at the deadline) without violating the probability constraint on the project completion. In order to provide an informed answer to this question, the following simulation-optimization method is proposed:

1. For each task i in the project ($i = 1, 2,\ldots, m$), use MCS to generate n random observations

ISBN 978-88-97999-85-0; Affenzeller, Bruzzone, Jiménez, Longo and Piera Eds.

(with n large enough) of its completion time, i.e., t_{i1}, t_{i2}, …, t_{in}.

2. For each task i, consider the decision variable d_i, which represents the delay time associated with task i. Then, consider the n observations of the 'shifted' completion time for task i given by $d_i + t_{i1}, d_i + t_{i2},\dots, d_i + t_{in}$.
3. Use the shifted completion times from the previous step and the logical termination condition of the project to obtain n observations for the project completion time – which, at this stage, will be a function of the decision variables d_{i1}, d_{i2}, …, d_{im}. Typically, the logical termination condition of the project is represented by a series of minimal paths that need to be terminated (all of them) before the project can be considered as completed. Each of these paths is composed of a series of tasks, and it terminates as soon as one of these tasks is finished (Faulin et al., 2008).
4. At this point, either an exact optimization solver (if m is reasonably low) or a metaheuristic algorithm is used to obtain (near-)optimal values for the decision variables. Notice that the probabilistic constraint makes the model non-linear, and thus linear programming solvers cannot be used.
5. Once (near-) optimal values for d_{i1}, d_{i2}, …, d_{im} have been obtained, it is convenient to perform several reliability-related studies, including: obtain statistics (average, variance, and quartiles) on the project's completion time, estimate the project survival function by measuring the probability that it might be finished at different target times, analyze how this survival function varies as different delay levels are considered, etc.

4. A NUMERICAL EXAMPLE

For the shake of illustrating the previous ideas with a numerical example, the following project was generated, solved and analyzed.

- Our project consists of 9 independent tasks to be carried in parallel. Each of these tasks has a random completion time which follows a probability distribution as depicted in Table 1, where the second column 'Alpha' refers to the shape of the Weibull distribution and the third column 'Beta' to the scale.
- For each task, a total of 10,000 random observations have been generated using the Minitab statistical software (any other statistical package could have been used instead). Then, an Excel model was constructed with these random observations and according to the procedure described in the previous section (Figure 5a).

Table 1: Probability Weibull Distributions for Each Task

Task	Alpha (shape)	Beta (scale)
1	2.8	1.8
2	2.7	1.7
3	2.6	1.6
4	2.5	1.6
5	2.4	1.4
6	2.2	1.2
7	2.3	1.3
8	2.1	1.1
9	2.0	1.0

	A	B	C	D	E	F	G	H	I	J	K
5			Tasks								
6		Weibull	1	2	3	4	5	6	7	8	9
7		alpha (shape)	2,8	2,7	2,6	2,5	2,4	2,2	2,3	2,1	2,0
8		beta (scale)	1,8	1,7	1,6	1,6	1,4	1,2	1,3	1,1	1,0
9		Observation	Task completing time -- We(alpha, beta)								
10		1	1,1367	0,7810	1,6554	0,9660	1,2363	1,1417	0,4388	1,3037	1,3375
11		2	1,2635	1,6395	1,1317	0,5584	1,9697	1,1294	1,4368	0,7213	0,6071

Figure 5: Simulated Random Completion Time for Each Task

- The delay time for each task is considered as decision variable. For each observation, the shifted completion time associated with delay time is calculated (Figure 6).

	L	M	N	O	P	Q	R	S	T
5	Tasks								
6	1	2	3	4	5	6	7	8	9
7	Delay time (decision variables) with 0 <= delay <= deadlinie								
8	0,5143	0,5182	0,5214	0,5153	0,5214	0,5142	0,5145	0,5176	0,5183
9	Delay time + Task completing time								
10	1,6510	1,2993	2,1768	1,4813	1,7577	1,6559	0,9534	1,8213	1,8558
11	1,7778	2,1578	1,6531	1,0737	2,4911	1,6436	1,9514	1,2389	1,1254

Figure 6: Shifted Completion Time for Each Task

- The project will be considered completed as soon as all the 7 paths (combination of tasks) represented in Table 2 are finished. For a path to be finished, it is enough that any of the tasks in it has been terminated. Therefore, the completion time for a given path is equal to the minimum shifted completion time of the task belonging to the given path. Thus, the project could be completed even if not all tasks have been finished by the deadline. For each observation, the maximum shifted completion time of the paths is considered as project finishing time (Figure 7, U9:AB11).

Table 2: Project-completion Paths

Paths of tasks		
1 – 4 – 7	2 – 4 – 7	3 – 6 – 9
1 – 4 – 8 – 9	2 – 4 – 8 – 9	N/A
1 – 5 – 6 – 9	2 – 5 – 6 – 9	N/A

- The common deadline was set to 2.0 time units (e.g., months, years, etc., in coherence with the units of the Weibull parameters), while the user-specified terminating probability was set to 0.90. The project termination situation is checked if it is completed before the determined deadline or no and consequently the probability of the project termination is calculated (Figure 7, AC3:AD7).

	U	V	W	X	Y	Z	AA	AB	AC	AD
1										
2										
3								E[time]	Sum delays:	4,66
4								1,5609	Deadline:	2,00
5	Paths (all need to be completed for the project to finish)							StDev	User Prob.:	0,90
6	1	2	3	4	5	6	7	0,3364	Observations:	10.000
7	Tasks to be finished for the path to be completed							Project finishing time	P(finished):	0,90
8	1v4v7	1v4v8v9	1v5v6v9	2v4v7	2v4v8v9	2v5v6v9	3v6v9			
9	Path completion time								Completed?	
10	0,9534	1,4813	1,6510	0,9534	1,2993	1,2993	1,6559	1,6559	1	
11	1,0737	1,0737	1,1254	1,0737	1,0737	1,1254	1,1254	1,1254	1	

Figure 7: Near-optimal Solution for the Example

In the proposed non-linear model, the objective is to minimize the summation of the delays. As constraints, *a*) the probability of the project termination should be more or equal than the user-specified probability, *b*) delay variable decisions should be less than deadline and finally *c*) these variables should be positive. Components of the constructed model (objective and constraints) in GRG Nonlinear solver are shown in Figure 8. The model was 'optimized' using the GRG Nonlinear solver integrated in the Excel Solver plug-in.

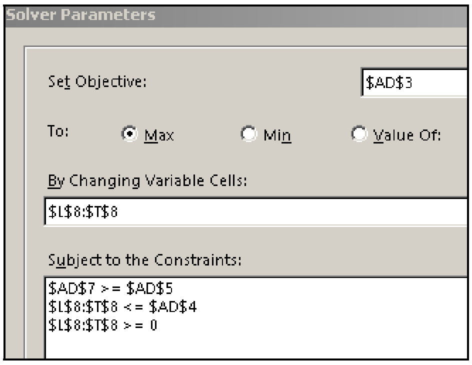

Figure 8: Non-linear solver parameters (objective and constraints)

The optimization process took about 2.5 minutes in a standard laptop computer (Intel Core i5 CPU @ 2.3 GHz, 8 GB RAM). Figure 6 (L8:T8) shows a partial screenshot of the Excel model, including the near-optimal delay values per task provided by the non-linear solver, which are the following ones: d_1 = 0.5143, d_2 = 0.5182, d_3 = 0.5214, d_4 = 0.5153, d_5 = 0.5214, d_6 = 0.5142, d_7 = 0.5145, d_8 = 0.5176, d_9 = 0.5183. These values were obtained by the non-linear solver after setting 0.5 as the initial value for all the tasks' delays (we have noticed that, probably due to convergence issues of the solver, the final solution provided might vary somewhat depending on these initial values).

5. ANALYSIS OF RESULTS

Figure 9 shows two survival functions for the project. These survival functions have been generated using the Kaplan-Meier estimator (Meeker and Escobar 1998) for non-censured data. In this case, the data refer to the observations provided by our method for the project completion times. One set of observations was obtained without assuming any delay in the tasks (i.e., all tasks were initiated from the very beginning), while the other set was obtained using the optimal delay values provided by the Excel solver.

Notice how the delayed survival function is shifted-to-the-right with respect the non-delayed one. Note also that, as it was anticipated, the delayed survival function intersects the deadline at 2.0 exactly when the survival probability reaches the value 0.1.

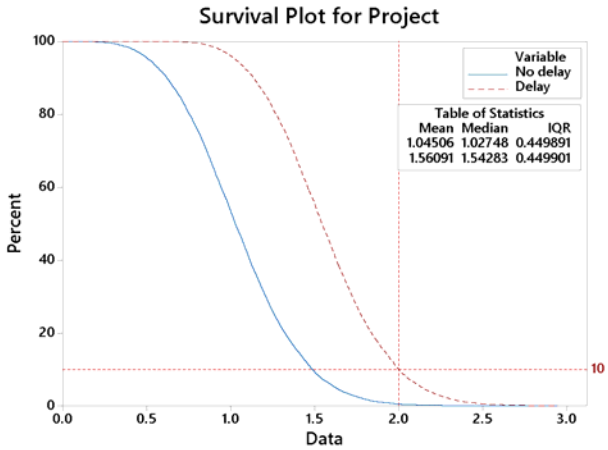

Figure 9: Survival Function for Project

Finally, Figure 10 shows a trade-off plot between total delay and probability of completing the project on or before the deadline. As expected, the higher the completion probability the lower the total delay that can be used –i.e., the sooner the resources must be assigned to the project.

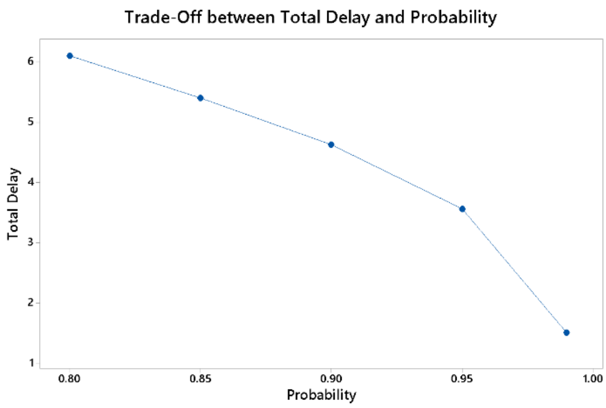

Figure 10: Trade-Off between Delay and Probability

6. CONCLUSIONS

Setting starting times for multi-task projects under uncertainty conditions can be a challenging problem for managers. This paper proposes a simulation-optimization method able to provide informed decisions on how to set these starting times, so that a given probability of project completion is achieved. Our method is combines simulation and optimization the best of our knowledge, it is the first time that such a relevant problem is discussed in the scientific literature and a solving approach is proposed.

Our results contribute to quantify how the survival function of the project evolves as more delay is included in the process. They also provide numerical values for quantifying the trade-off between the user specified probability of project termination and the maximum delay that can be considered without violating the probabilistic constraint on project completion.

Several research lines emerge from this preliminary work. First, a more robust non-linear solver should be used for larger cases. In fact, metaheuristic approaches could be a good alternative to exact methods for solving large-scale instances in short computing times. Also, other project topologies can be explored, including interdependencies among tasks. Finally, the tasks themselves could be complex processes, e.g., flow-shops, that also require some pre-optimization process.

ACKNOWLEDGMENTS

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness and FEDER (TRA2013-48180-C3-P, TIN2015-66951-C2-2-R, MTM2015-64465-C2-1-R), and the Erasmus+ program (2016-1-ES01-KA108-023465).

REFERENCES

Barata J., Guedes-Soares C., Marseguerra M., and Zio E., 2001. Modelling components degradation processes of Monte-Carlo simulation. In: Zio E., Demichela M. and Piccinini N., editors. Safety and Reliability: towards a safer world. Proceedings of the European conference on safety and reliability ESREL 2001, 16-20 September, Torino, Italy, pp. 879–886.

Barata J., Guedes-Soares C., Marseguerra M. and Zio E., 2002. Simulation modelling of repairable multi-component deteriorating systems for ‘‘on-condition’’ maintenance optimisation. Reliability Engineering and System Safety, 77, 255–267.

Billinton R. and Wang P., 1999. Teaching distribution systems reliability evaluation using Monte Carlo simulation. IEEE Transaction on Power Systems, 14, 397–403.Blischke W.R. and Prabhakar-Murthy D.N., 2000. Reliability. Modeling, prediction, and optimization. New York: Wiley.

Borgonovo E., Marseguerra M. and Zio, E., 2000. A Monte Carlo methodological approach to plant availability modelling with maintenance, aging and obsolescence. Reliability Engineering and System Safety, 67, 61–73.

Cabrera G., Juan A., Lazaro D., Marques J. and Proskurnia I., 2014. A simulation-optimization approach to deploy Internet services in large-scale systems with user-provided resources. Simulation: Transactions of the Society for Modeling and Simulation International, 90 (6), 644–659.

Dubi A., 2000. Monte Carlo applications in systems engineering. New York: Wiley.

Faulin J., Juan, A., Martorell, S. and Ramirez-Marquez, J., 2010. Simulation Methods for Reliability and Availability of Complex Systems. Springer.

Faulin J., Juan A., Serrat C. and Bargueno V., 2007. Using Simulation to determine Reliability and Availability of Telecommunication Networks. European Journal of Industrial Engineering, 1 (2), 131-151.

Faulin J., Juan A., Serrat C. and Bargueno V., 2008. Predicting availability functions in time-dependent complex systems with SAEDES simulation algorithms. Reliability Engineering and System Safety, 93 (11): 1761–1771.

Grasas, A., Juan, A. and Ramalhinho, H., 2016. SimILS: A Simulation-based extension of the Iterated Local Search metaheuristic for Stochastic Combinatorial Optimization. Journal of Simulation, 10(1): 69-77.

Juan A. and Vila A., 2002. SREMS: System Reliability Using Monte Carlo Simulation with VBA and Excel. Quality Engineering, 15 (2) 333–340.

Labeau P.E. and Zio E., 2000. The mathematics of component-based and system-based Monte Carlo sampling in realistic systems engineering. Proceedings of the MRR2000, 2, 667–670.

Labeau P.E. and Zio E., 2002. Procedures of Monte-Carlo transport simulation for applications in system engineering. Reliability Engineering System Safety, 77, 217–228.

Levin M. and Kalal T., 2003. Improving product reliability: strategies and implementation. New York: Wiley.

Meeker W. and Escobar L., 1998. Statistical methods for reliability data. New York: Wiley.

Skrzypczak I., Słowik M. and Buda-Ożóg L., 2017. The Application of Reliability Analysis in Engineering Practice – Reinforced Concrete Foundation. Procedia Engineering, 193,144-151.

Sobie C., Freitas C. and Nicolai M., 2018. Simulation-driven machine learning: Bearing fault classification. Mechanical Systems and Signal Processing, 99, 403-419.

Topaloglu M., Caldibi E. and Oge G., 2016. The scale for the individual and social impact of students' social network use: The validity and reliability studies. Computers in Human Behavior, 61, 350-356.

Zheng L. and Chang Y., 2017. Risk assessment model of bottlenecks for urban expressways using survival analysis approach. Transportation Research Procedia, 25, 1544-1555.

AUTHORS BIOGRAPHY

Angel A. Juan is Associate Professor at the IN3-Open University of Catalonia. He holds a PhD in Industrial Engineering and a MSc in Mathematics. He completed a pre-doctoral internship at Harvard University and a postdoctoral internship at the MIT Center for

Transportation and Logistics. His research interests include applications of randomized algorithms and simheuristics in logistics, production, and Internet computing. He has published over 50 JCR-indexed articles in these fields.
Laura Calvet is Assistant Professor at the University of Barcelona and PhD student at the IN3-Open University of Catalonia. She holds a MSc in Statistics and Operations Research, a BSc in Economics and a BSc in Applied Statistics. Her work is related to the combination of statistical learning and simheuristics for solving complex combinatorial optimization problems under uncertainty.
Carles Serrat is Associate Professor of Applied Statistics in the Department of Mathematics at the Technical University of Catalonia (Barcelona, Spain). He holds a PhD and BS in Mathematics. Dr. Serrat develops his research in the context of the Institut d'Estadística i Matemàtica Aplicada a l'Edificació (IEMAE).
Sara Hatami is a postdoctoral researcher at the IN3-Open University of Catalonia. She holds a PhD and a MSc in Industrial Engineering. Her research focuses on the development of metaheuristics and simheuristics for industrial optimization problems, using traditional and modern Operations Research techniques and modeling complex real-life problems.