Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale, Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Tune Decomposition Schemes for Large-Scale Mixed-Integer Programs by Bayesian Optimization

Guido Sand[a*], Sophie Hildebrandt[a,b], Sina Nunes[a], Chung-On Yip[a] and Meik Franke[b]

[a] Pforzheim University of Applied Science, School of Engineering, Pforzheim, Germany
[b] University of Twente, Faculty of Science and Technology, Enschede, The Netherlands
* Corresponding Author: guido.sand@hs-pforzheim.de.

ABSTRACT

Heuristic decomposition schemes like moving horizon schemes are a common approach to approximately solve large-scale mixed-integer programs. The authors propose Bayesian optimization as a methodological approach to systematically tune parameters of decomposition schemes for mixed-integer programs. This paper discusses detailed results of three studies of the Bayesian optimization-based approach using hoist scheduling as a case study: Firstly, two objectives of the tuning problem are examined considering sequences of incumbent solutions found by the Bayesian optimization. Secondly, the Bayesian optimization is applied to a set of test instances of the hoist scheduling problem using four types of acquisition functions; they are compared with respect to the convergence of the tuning problem solutions. Thirdly, the scaling behaviour of the Bayesian optimization is studied with respect to the dimension of the space of tuning parameters. The results of the three studies show that the solutions found by the Bayesian optimization converge quickly in smaller and larger tuning parameter spaces using different types of acquisition functions.

Keywords: Derivative Free Optimization, Machine Learning, Mixed-Integer Programming

MOTIVATION

Decomposition Schemes for MIPs

Heuristic decomposition schemes are a common approach to approximately solve large-scale mixed-integer programs (MIPs) as a sequence of partial MIPs. A typical example are moving horizon schemes applied to scheduling problems. Decomposition schemes usually exhibit parameters which can be used to tune their performance. Examples for parameters of moving horizon schemes are the horizon length and the step size of its movement.

Existing Tuning Methods

30 publications on decompositions schemes for mixed-integer programs (both linear and non-linear) were analysed with respect to systematic tuning methods. The publication dates range from 2002 to 2025; 26 of the publications are journal papers, 3 are conference papers and 1 is a book section. The mixed-integer programs cover batch scheduling, maintenance scheduling, project scheduling, supply network planning, production routing, bin packing, locomotive assignment, healthcare planning, flowsheet optimization, lot sizing and other optimization problems. The decomposition schemes cover temporal, area, activity, scenario and geographic decomposition as well as rolling horizon and other schemes.

Tuning parameters were clearly identified in 25 publications; in the other 5 publications the existence of tuning parameters could not be decided. Only in 12 publications the tuning parameters were tuned, in 9 of them by grid search [e.g. 1,2,3], in 2 by heuristics and in 1 they were chosen randomly. No publication with an optimization-driven tuning method was found.

Bayesian Optimization Based Tuning

In a previous paper [4], Bayesian optimization was proposed as a methodological approach to tune decomposition schemes for mixed-integer programs. This approach is reasonable since the tuning problem is a black-box optimization problem with an expensive to evaluate objective function: Each evaluation of the objective function of the Bayesian optimization requires the (approximate) solution of the MIP applying the parametrized decomposition scheme. The mentioned paper demonstrated by a case study that the proposed approach is feasible and effective in principle. The case study was a mixed-integer hoist scheduling model along with a moving horizon decomposition scheme (see below).

Scope of this Paper

This paper discusses results of three studies of the Bayesian optimization-based approach using the same case study: Firstly, two objective functions of the tuning problem are examined considering sequences of incumbent solutions found by the Bayesian optimization. (An incumbent solution is the best solution found up to a certain iteration.) Secondly, the Bayesian optimization is applied to a set of test instances of the hoist scheduling problem using four types of acquisition functions; they are compared with respect to the convergence of the tuning problem solutions. Thirdly, the scaling behaviour of the Bayesian optimization is studied with respect to the dimension of the space of tuning-parameters.

CASE STUDY

Hoist Scheduling Model

Hoist scheduling is the task of assigning several steps of several jobs to several stations over time. Figure 1 shows an exemplary layout of a batch plant with a hoist (left) and exemplary recipes defining the sequence and duration of steps for two jobs. In addition to the usual job shop scheduling constraints, the hoists must be considered. Hoists transport product carriers between the stations and constrain the solution space due to their finite transportation times. A typical objective of hoist scheduling is makespan minimization. Hoist scheduling problems can be formulated as MIPs.

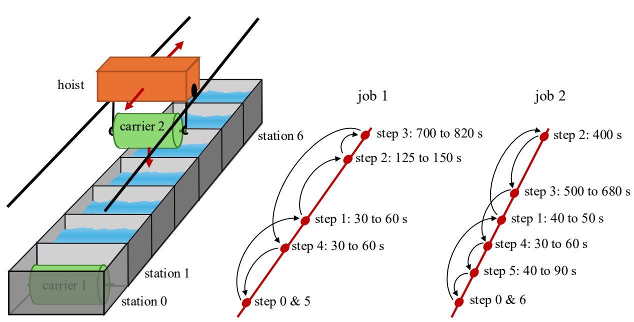

Figure 1. Hoist scheduling. Left: Exemplary layout of a batch plant with hoist. Right: Exemplary recipies.

In this paper, the MIP-model by Aguirre at al. [1] is used as a case study. This model considers one hoist, stations with a capacity of one carrier and non-circulating carriers. In the context of this work, it serves as an example for an NP-complete problem, that calls for decomposition schemes for large-scale instances.

Heuristic Decomposition Scheme

The heuristic decomposition scheme used in this paper is an extension of the moving horizon scheme described by Aguirre et al. [1]. It comprises two sequential phases to schedule jobs from a given list of jobs: a construction phase and an improvement phase.

In the construction phase, jobs 1 to NSJ (number of selected jobs) are scheduled by solving a partial MIP, then jobs 1 to NFJ (number of fixed jobs) are fixed and jobs NFJ+1 to NSJ+NFJ+1 are (re)scheduled, and so on. The number of partial MIPs to be solved is given by the number of jobs divided by NFJ. The construction phase generates a first feasible schedule, which may be improved in the improvement phase.

In the improvement phase, jobs 1 to NRJ (number of released jobs) from the schedule are rescheduled by solving a partial MIP. If the schedule is not improved, jobs 2 to NRJ+1 are rescheduled, and so on. If the schedule is improved, the improvement phase re-starts from the first job in the schedule. The improvement phase terminates if no further improvement is observed. In contrast to the construction phase, the number of partial MIPs to be solved during the improvement phase cannot be predicted.

The tuning parameters NSJ, NFJ and NRJ are integer in nature and have a lower bound of 1. The upper bounds of NSJ and NRJ are the total number of jobs; the upper bound of NFJ is NSJ: NFJ≤NSJ. In addition to the integer parameters, the accepted optimality gap (AOG) of the partial MIPs (termination criterion of the MIP solver) is considered as a continuous tuning parameter within the bounds of 0 % to 100%. Aguirre et al. considered only NSJ and NRJ as tuning parameters; this paper extends the tuning parameter space by NFJ and AOG.

Architecture of the Tuning Method

Figure 2 illustrates the algorithmic components for the tuning of decomposition schemes for large-scale MIPs and the information exchanged between them.

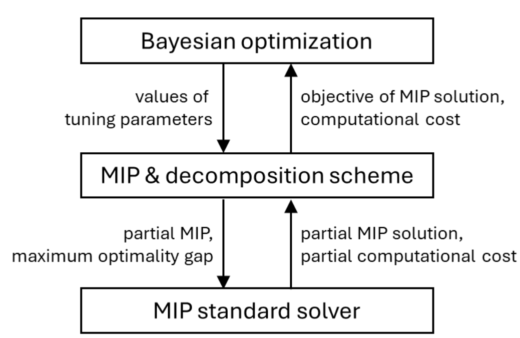

Figure 2. Algorithmic architecture for MIP decomposition scheme tuning.

The Bayesian optimization generates the next values of tuning parameters for the decomposition scheme;

the next values are based on previous values of the objective of the MIP solution or the computational cost to apply the decomposition scheme. For the studies in this paper, the implementation SMAC3[1] is used.

Based on the values of the tuning parameters, the decomposition scheme generates a sequence of partial MIPs from the MIP. The MIP and the decomposition scheme are implemented using Pyomo[2].

The MIP standard solver solves the partial MIPs with at least the AOG provided by the Bayesian optimization. It returns the partial MIP solution along with its partial computational cost. For the studies in this paper, CPLEX[3] is used as the MIP standard solver.

TUNING PROBLEM OBJECTIVES

In the first study two objective functions of the tuning problem are examined considering sequences of incumbent solutions found by the Bayesian optimization.

Motivation for Different Objectives

Applying heuristic decomposition schemes to large-scale MIPs has two natural and competing objectives: The computational cost shall be minimized while the objective of the MIP shall still be optimized. For the case study at hand, the computational cost is measured by the CPU-time needed to apply the decomposition scheme to the MIP. The MIP objective is to minimize the makespan of the hoist schedule, measured in time units TU.

Setup of Computational Studies

Heuristic decomposition schemes are useful for large-scale MIPs with computational cost that prohibit their solution. However, the studies in this paper are limited to small-scale MIP-instances that can (exactly) be solved at reasonable computational cost to ensure that all partial MIPs can be solved to proven optimality.

Four instances of the hoist scheduling model comprising six jobs each are considered; the instances differ in the number of steps and stations and in the speed of the hoist.

The number of jobs is kept constant to keep the size of the tuning parameter space constant. In this first study the tuning parameters NSJ und NRJ vary between 1 und 6, while NFJ is set to 1 and AOG is set to 0. This leads to a tuning parameter space with only 36 feasible (integer) solutions, such that a grid search with complete enumeration of all solutions is possible. Only one objective of the tuning problem (computational cost or makespan) is optimized at a time (without limiting the other).

Computational Results

A grid search in the NSJ-NRJ-space of the tuning problem leads to similar shapes of each of the objective functions for the four instances. Therefore, only the results for instance 1 are discussed in the sequel.

Figure 3 shows the results of the grid search for the computational cost objective combined with the incumbent solutions found by the Bayesian optimization (red line). Note, that only the integer-valued points are feasible while the values of the objective function are interpolated between them to improve the clarity of the figure.

The computational cost objective exhibits multiple minima located at the corners of the tuning parameter space: (1;1), (6;1), (1;6) and (6,6). The Bayesian optimization is initialized at the worst solution (4;4); after finding two local optima it terminates at the global one (1;1). This solution is intuitively reasonable for the computational cost objective, since the partial MIPs in the construction phase and the improvement phase are of minimal size and cheap to solve.

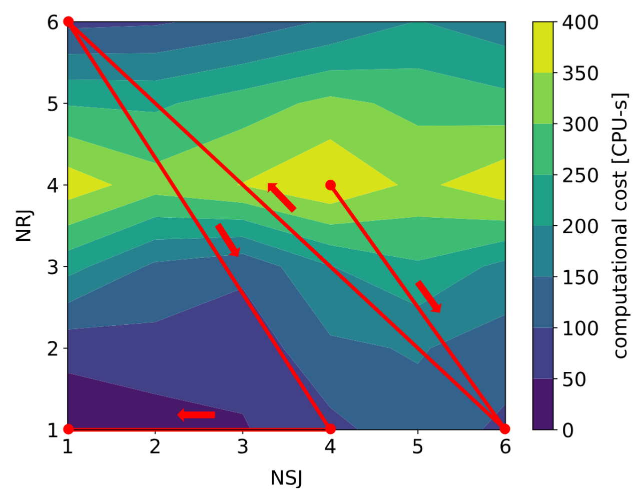

Figure 3. Computational cost objective of instance 1 and incumbent solutions found by the Bayesian optimization.

Analogue to Figure 3, Figure 4 shows the results for the makespan objective. The maximum of the makespan objective (1;1) is at the same point as the minimum of the computational cost objective. This exemplifies the competing character of the two objectives and demonstrates the need to balance them appropriately.

Analogue to above, the Bayesian optimization is initialized at the worst solution (1,1). It terminates after the first incumbent solution (6;1). The early termination is obviously favoured by the weakness of the objective function: Since 27 out of 36 solutions are optimal, an optimum is easy to find.

Note, that NSJ=6 or NRJ=6 means that the MIP is neither decomposed in the construction phase nor in the

[1] https://automl.github.io/SMAC3/main/index.html#

[2] https://www.pyomo.org/

[3] https://www.ibm.com/products/ilog-cplex-optimization-studio

improvement phase. The optimality of solutions (6;1), (6,2), ... (6,6), (5,6) ... (1;6) is intuitively reasonable: Decomposing the MIP bears the risk of losing its optimum, such that by not decomposing it the optimum is secured.

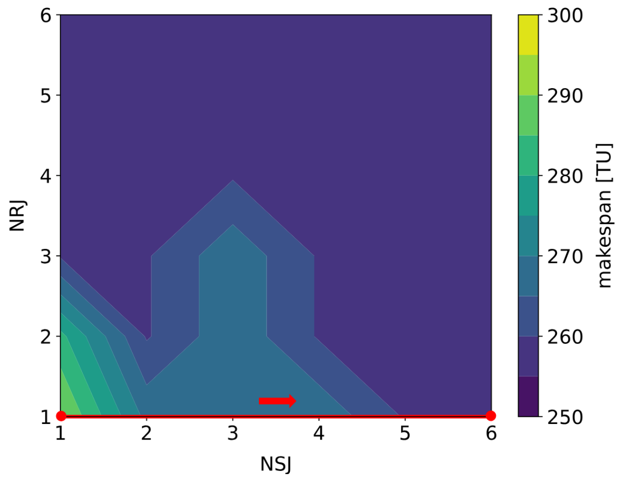

Figure 4. Makespan objective of instance 1 and incumbent solutions found by the Bayesian optimization.

ACQUISITION FUNCTIONS

In the second study the Bayesian optimization is applied to a set of test instances of the hoist scheduling problem using four types of acquisition functions; they are compared with respect to the convergence of the tuning problem solutions.

Impact of the Acquisition Function

Acquisition functions steer the selection of the next point in the solution space of the Bayesian optimization. Different types of acquisition functions balance the exploitation and the exploration of the search space in different ways. Therefore, the type of acquisition function is assumed to be a design parameter of the Bayesian optimization with a major impact on its performance.

The following four common acquisition functions are applied to the four small-scale instances and the two objective functions from the first study: "expected improvement" (also applied in the first study), "probability of improvement"," upper confidence bound" and "Thompson sampling".

Setup of Computational Studies

For each acquisition function and each objective function, the objective values of the incumbent solutions of the four instances are plotted over the iterations of the Bayesian optimization. To make them comparable, the objective functions are normalized such that the starting point equals 100% and the optimum equals 0%. Note that the optimum in the tuning parameter space is known from the first study. For each iteration, the minimum, the maximum and the mean value of the normalized objective are displayed.

Computational Results

The results for the acquisitions functions "expected improvement"; "upper confidence bound" and "Thompson Sampling" are similar, while the convergence using the "probability of improvement" function is slower. For the makespan objective the Bayesian optimization converges after 2-3 iterations to an optimum, while for the computational cost objective the number of iterations is higher. (This observation is consistent with study 1.) Therefore, only the results for the computational cost objective and the acquisition functions "probability of improvement" and "expected improvement" are discussed (see Figures 5 und 6).

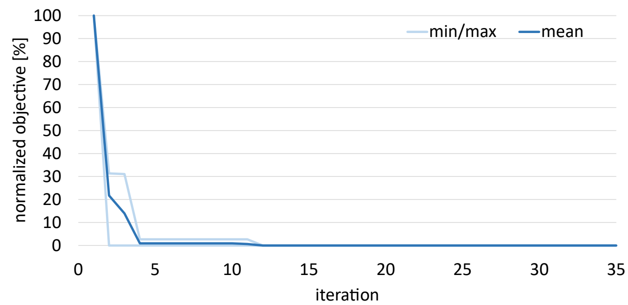

Figure 5. Convergence of computational cost objective with acquisition function "expected improvement".

The acquisition function "expected improvement" considers both, the likelihood of improvement of the next solution and the magnitude of its potential improvement [5]. It balances exploration and exploitation of the tuning parameter space and is therefore well suited for objective functions with multiple local optima.

As Figure 5 shows, the instances converge after 2 to 12 iterations to the global optimum (with a normalized objective of 0%). After 4 iterations the maximum normalized objective over all instances is around 3%. The instances terminate after 14 to 35 iterations. The relevant termination criterion is the maximum number of 16 retries i.e. after 16 iterations no new solution in the tuning parameter space is found.

In contrast to the acquisition function "expected improvement", the acquisition function "probability of improvement" selects the next solution such that the objective value is most likely to improve [6]. It typically selects the next solution close to the last solutions, such that it behaves more exploitative than explorative.

Applying the acquisition function "probability of improvement" to the case study leads to slower convergence to inferior solutions compared to the other acquisition functions. As Figure 6 shows, the instances terminate after 11 or 13 iterations before the final solution with normalized objectives between 1% and 13% is reached.

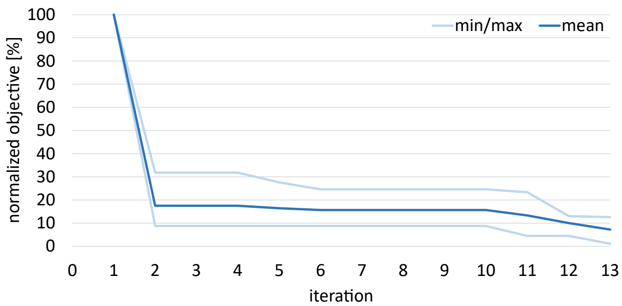

Figure 6. Convergence of computational cost objective with acquisition function "probability of improvement".

TUNING PARAMETER SPACE DIMENSION

The third study analyses the scaling behaviour of the Bayesian optimization with respect to the dimension of the space of tuning parameters.

Extension of the Tuning Parameter Space

The proposed Bayesian optimization-based approach is particularly useful for tuning parameters spaces of a size that prohibit full grid search. Therefore, the two-dimensional tuning parameter space considered in studies 1 and 2 is extended by two additional dimensions: the integer-valued NFJ and the continuous AOG. (Both were introduced in section "Heuristic Decomposition Scheme" but bound to fixed values in studies 1 and 2.) The same settings as in study 1 were applied.

Comparative Analysis of the Convergence

To comparatively analyse the convergence speed, the same type of plots with the same normalization is used as in Figures 5 and 6. Figures 7 and 8 show the convergence of both objective functions in the larger tuning parameter space.

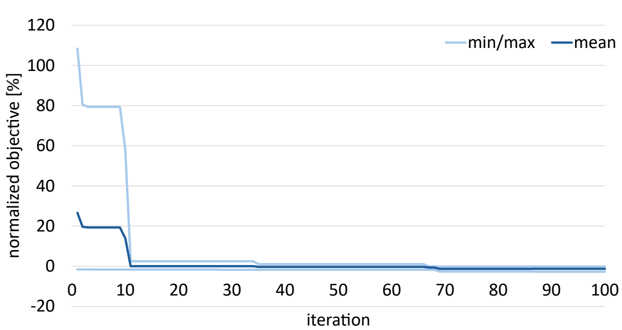

Figure 7. Convergence of computational cost objective with four-dimensional tuning parameter space.

As the full grid-search is prohibited, the worst solutions in the tuning parameter space are not know such that they cannot be used to initialize the Bayesian optimization (as in Figures 5 and 6). Instead, arbitrary initialization points were chosen for Figures 7 and 8.

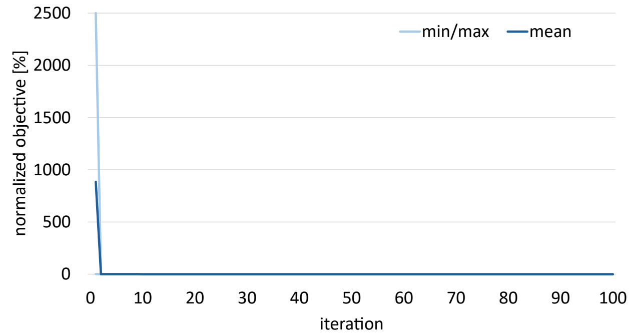

Figure 8. Convergence of makespan objective with four-dimensional tuning parameter space.

The worst normalized objective value in Figure 7 is 108% at iteration 1. This shows that the larger search space of this study 3 contains worse solutions than the smaller search space of studies 1 and 2. On the other hand, the Bayesian optimization terminates after 100 iterations at a solution with a normalized objective vale of -3%. This shows that the additional dimensions of the tuning space (NFJ and AOG) provide additional optimization potential that was exploited by the Bayesian optimization.

The convergence speeds in Figures 5 (smaller tuning parameter space) and 7 (large tuning parameter space) are comparable. Figure 7 shows that the maximum normalized objective over all instances is around 3% after 11 iterations (compared to 4 iterations in Figure 5).

Figure 8 shows similar results for the makespan objective than Figure 7 for the computational cost objective: The worst normalized objective in Figure 8 is 2500% (compared to 100% in Figure 6), and the maximum normalized objective over all instances is around 1% already after 2 iterations.

Comparative Analysis of the Optimized Tuning Parameters

The optimized tuning parameters for the two objectives are shown in Tables 1 and 2.

Table 1: Optimized tuning parameters for computational cost objective.

Instance	NSJ	NRJ	NFJ	AOG
1	1	1	1	78%
2	4	6	3	86%
3	1	1	1	100%
4	6	6	1	91%

Table 2: Optimized tuning parameters for makespan objective.

Instance	NSJ	NRJ	NFJ	AOG
1	6	2	2	0%
2	3	3	1	2%
3	3	3	1	2%
4	5	5	5	9%

For the computational cost objective (see Table 1), most of the values of the integer-valued tuning parameters NSJ, NRJ and NFJ are at their boundaries. This result is reasonable since solving many small or one medium-sized MIP is computationally cheap (compare Figure 3).

The AOG is close to or at 100%, which is also reasonable since the partial MIPs terminate the earlier the larger the AOG is.

For the makespan objective (see Table 2), most of the values of the integer-valued tuning parameters NSJ, NRJ and NFJ are away from their boundaries (contrasting Table 1). This observation can reasonably be explained by the weak optimum of this objective (compare Figure 4).

The AOG is close to or at 0% (also contrasting Table 1). This result is also reasonable since smaller optimality gaps push the MIP solver to better solutions of the partial MIPs.

CONCLUSIONS AND FUTURE WORK

In this work, a Bayesian optimization-based approach to tune decomposition schemes for MIPs was studied in detail using hoist scheduling as a case study. The results support the following two main conclusions:

Firstly, decomposition schemes comprise tuning parameters which are sensitive to the two natural objectives of decomposing, namely minimizing the computational cost while still optimizing the MIP objective; these two objectives are competing.

Secondly, a reasonable method to tune the parameters of a decomposition scheme is Bayesian optimization due to its following properties:

- The performance of the Bayesian optimization is insensitive against changes of its acquisition function. If exploration and exploitation are not unbalanced, good solutions are found after a few iterations.
- The performance of the Bayesian optimization is insensitive against changes of the parameters of the MIP. The convergence behaviour of different instances of the MIP is similar.
- Multiple local optima and the global optimum in the space of tuning parameters can be found by the Bayesian optimization.
- The Bayesian optimization is efficient in smaller and larger mixed-integer tuning parameter spaces. Its performance in two- and four-dimensional spaces is comparable.

In its present state, Bayesian optimization-based approach exhibits limitations. Future work deals with the approaches to relax them: Multi-objective optimization will be studied; large-scale MIP which cannot be solved to proven optimality will be examined; and methods to deal with multiple MIP-instances will be developed.

REFERENCES

1. Aguirre A, Méndez C, Sànchez Á, Mier M. Applying MILP/Heuristic Algorithms to Automated Job-Shop Scheduling Problems in Aircraft-Part Manufacturing. *IJIE* 5(10):26-41 (2014) https://incubadora.periodicos.ufsc.br/index.php/IJIE/article/view/3043
2. Kelly JD, Mann JL. Flowsheet decomposition heuristic for scheduling: a relax-and-fix method. *Comp.Chem.Engg.* 28:2193–200 (2004). https://doi.org/10.1016/j.compchemeng.2004.03.009.
3. Paquay C, Limbourg S, Schyns M, Oliveira JF. MIP-based constructive heuristics for the three-dimensional Bin Packing Problem with transportation constraints. *Int.J.Prod.Res.* 56:1581–92 (2018). https://doi.org/10.1080/00207543.2017.1355577.
4. Hildebrandt S, Sand G. Hyperparameter optimization of matheuristics for hoist scheduling. In: ESCAPE34/PSE24. Ed: Manenti F, Reklaitis GV. Elsevier (2024) https://doi.org/10.1016/B978-0-443-28824-1.50488-9
5. Mockus J, Tiesis V, Zilinskas A. The application of Bayesian methods for seeking the extremum. *Towards Global Optimisation* 2:117–129 (1978)
6. Kushner, H J. A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise. *J. Fluids Eng.* 86(1):97–106 (1964) https://doi.org/10.1115/1.3653121