Development of a hybrid metaheuristic-based dual ascent method for reliability optimization of IT systems

Viktor Morozov[1,*,†] and Vladyslav Deineha[1,†]

[1] *Taras Shevchenko National University of Kyiv, 24, Bohdan Gavrilishin Str., Kyiv, 04116, Ukraine*

Abstract

The article considers the problem of optimal reservation in serial-parallel systems, where it is necessary to maximise reliability under a limited budget. A model with a nonlinear objective function and resource constraints is constructed, reflecting the practical conditions of critical IT system design. To solve it, a hybrid approach is proposed, combining the Dual Ascent method with the metaheuristic Particle Swarm Optimisation method, which solves the internal optimisation problem. A numerical experiment showed the advantages of the hybrid method compared to the classical Dual Ascent, where internal optimisation is performed by gradient descent. The resulting hybrid method demonstrated higher system reliability, superiority in finding the optimal solution, and more complete use of the budget. In addition, the results proved to be stable for the initial generations, confirming the practical reliability of the approach. Thus, the combination of Dual Ascent and PSO proves its effectiveness in optimal redundancy tasks, ensuring a balance between strict enforcement of constraints and high quality of searching for optimal configurations.

Keywords

reliability optimization, optimal redundancy, series-parallel systems, dual ascent, particle swarm optimization, hybrid method

1. Introduction

Optimisation tasks play an important role in industry, engineering and applied mathematics. Optimisation is understood as the process of finding the best solution among a set of acceptable options, given certain constraints. Depending on the problem, this may involve minimising costs, time or risks, or maximising the reliability, efficiency or quality of an IT system's performance. Optimisation methods are widely used in a variety of fields, from economics and logistics to energy and information technology [1] [2].

One important area is reliability theory, which quantifies a system's ability to perform its functions despite component failures. Reliability theory is applied in various fields, such as energy [3]. In practice, it is impossible to create completely failure-resistant components, so the probability of failure is always different from zero. To ensure the required level of reliability of critical systems - energy facilities, transport infrastructure, data centres - a redundancy approach is used.

Reserving involves introducing additional elements (reserves) that take over functions in the event of failure of the main part of the system. There are different system structures: series, where the failure of one element leads to the failure of the entire system; parallel, which remain operational as long as at least one of the elements is working; and combined (e.g. series-parallel), which combine the properties of both previous structures [4]. A common practical approach is series-parallel redundancy, as it allows for a significant increase in reliability without excessive cost increases.

The question naturally arises: how many and what kind of reserves need to be established to achieve the desired level of reliability with limited resources? Obviously, an unlimited increase in

[1] *SNE 2025: Workshop on Software and Knowledge Engineering, November 19-20, 2025, Almaty, Kazakhstan*
* Corresponding author.
† These authors contributed equally.
viktor.morozov@knu.ua (V. Morozov); thedynkan@knu.ua (V. Deineha)
0000-0001-7946-0832 (V. Morozov); 0009-0008-3123-2302 (V. Deineha)

the number of reserves is economically and technically impractical. This is where the task of optimal redundancy comes in – finding a distribution of reserves that ensures maximum system reliability within fixed constraints (e.g., budget, weight, or power consumption). Optimal redundancy is an important task when designing critical systems. The stability of power systems, the continuity of data centres, and the safety of transport and medical complexes depend on its solution [5].

This paper considers an approach based on a combination of different optimisation methods. Gradient methods are a classic tool for finding extremums, based on moving in the direction of the target function gradient, and work well for smooth problems [6]. Dual Ascent allows constraints to be effectively taken into account by introducing Lagrange multipliers and gradually updating them. Particle Swarm Optimisation, in turn, is a metaheuristic swarm method that searches for the optimum using the collective behaviour of 'particles,' making it suitable for complex and nonlinear problems. The proposed hybrid approach, which combines Dual Ascent with PSO to solve the internal optimisation problem, provides a balance between strict enforcement of constraints and the ability to find high-quality solutions in complex search spaces.

The scientific novelty of the proposed approach lies in the integration of two heterogeneous optimization methods into a unified hybrid structure. Unlike prior reliability optimization studies, which use only metaheuristic algorithms or only gradient methods, the proposed method integrates Dual Ascent's constraint-handling precision with PSO's stochastic global search. This combination allows for effective work with constraints, ensuring faster convergence and solving the dual problem in complex system optimal reservation tasks.

The proposed method is applicable to a wide range of reliability optimization problems, regardless of the specific system configuration. As an example, it can be applied to large-scale series-parallel systems representing real IT infrastructures such as data centers, and distributed computing networks. Since each subsystem in a series-parallel structure corresponds to a node or stage with several parallel elements, the method maintains computational efficiency even with increasing system size. The optimization process naturally accounts for resource and budget limitations, which allows maintaining both reliability and cost-effectiveness in large-scale environments. This makes the approach suitable for practical reliability optimization, in particular, in large information infrastructures.

2. Analysis of recent research and publications

Review article [7] on optimal redundancy includes types of redundancy, series-parallel schemes, component parameter uncertainties, system reliability modelling, and methods for solving a specific problem. This shows that there are many approaches to the topic of optimal redundancy and indicates the relevance of the topic. It also discusses metaheuristic methods as options for solving redundancy problems and also points out the peculiarities of the methods, noting that developing new approaches is valuable. Paper [8] provides an overview of approaches to solving various reliability optimisation problems, the application of these methods to different types of design problems, as well as heuristics, metaheuristic algorithms, and exact methods. This is relevant to our work and highlights the complexity of the optimal redundancy problem and the need for fast computational methods. Article [9] presents a systematic review of the application of evolutionary algorithms to reliability optimisation and redundancy problems. The authors evaluate the effectiveness of different approaches to solving the problem of reserve allocation in complex system structures. This work is relevant to our research because it confirms the feasibility of using evolutionary methods in optimal reserve allocation problems and emphasises the need to search for new combined approaches.

Article [10] considers the problem of optimising the design of a pharmaceutical plant by distributing reliability among heterogeneous components, taking into account design constraints. The authors solve the problem by using three natural artificial intelligence algorithms and a penalty function to work with constraints. The relevance of this work to our research lies in the fact that it demonstrates the effectiveness of applying artificial intelligence algorithms to complex reliability distribution problems with constraints, which directly echoes our approach to optimal redundancy. Article [11] demonstrates the application of an improved PSO to optimal reservation, particularly in series-parallel systems, which is also considered in our study. This confirms the relevance of combining metaheuristic methods with theoretical approaches in reservation tasks, which directly correlates with our work. This work [12] uses a gradient method to solve the problem of reserve allocation in a series-parallel system. They formalise the reliability function and apply a gradient optimiser to maximise it under constraints. This is directly relevant to our task: it shows that gradient methods can work in optimal reservation and provide a basic reference point, especially if the reliability function is smooth.

The paper [13] emphasises the importance of the series-parallel structure as a basic model for analysing the reliability of complex systems. The authors applied the Lagrange multiplier method, which provides real solutions for the main parameters, demonstrating the practical effectiveness of this approach in optimal redundancy problems. The article [14] considers the problem of optimal distribution of reserve modules in large serial-reserved systems, taking into account budget constraints. To solve the problem, the Lagrange multiplier method (analytical approach) was used, as well as other algorithms, including evolutionary ones. This is relevant to our work, since Dual Ascent also uses Lagrange multipliers and is a combination of approaches, and this all confirms the feasibility of using Dual Ascent in combination with a heuristic algorithm to achieve more accurate solutions in the optimal reservation problem.

Article [15] discusses the Dual Ascent method, which is relevant to our work in which this method is used. It states that Dual Ascent is an effective method for solving linearly constrained convex optimisation problems, as well as the importance of choosing the step size in it. In this work [16] the Dual Ascent method is applied in the context of resource allocation in networks with distributed architecture. Dual Ascent proves its relevance when it is necessary to ensure a balance between cost and performance. This demonstrates that Dual Ascent is a real modern method with applications in similar constrained problems.

3. Problem statement

3.1. General provisions

The system under consideration consists of several series of subsystems (stages or nodes), each of which may contain one or more parallel elements. This architecture is called a series-parallel structure. It reflects a broad class of real technical and organisational systems: power supply systems, computing clusters, transport nodes, etc.

To ensure reliability in such structures, a redundancy approach is widely used, i.e. the duplication of individual system elements. If at least one element of the subsystem is working properly, then the entire subsystem is considered operational. If all elements of the subsystem fail, this leads to the failure of the entire node and, consequently, to a decrease in the overall reliability of the system.

The reliability of the system depends both on the probability of fault-free operation of individual elements and on the number of parallel reserves installed in each subsystem. At the same time, the number of reserves cannot be arbitrary, as there are resource constraints: financial budget, equipment weight, dimensions, energy consumption.

Thus, the task of optimal reservation arises — it is necessary to distribute available resources among subsystems in such a way as to maximise the probability of trouble-free operation of the entire system.

3.2. Formalization

Let the system consist of m series of subsystems (stages). For subsystem i, the following is specified:

- probability of failure-free operation of a single element $p_i \in (0,1)$;
- cost of one element $c_i > 0$;
- number of reserves $k_i \geq 1$, which is determined by the solution.

Reliability of a subsystem with k_i parallel elements:

$$R_i(k_i) = 1 - (1 - p_i)^{k_i} \tag{1}$$

This means that the subsystem fails only if all of its elements fail simultaneously.

The overall reliability of a series-parallel system is equal to the product of the reliabilities of all subsystems, since they are connected in series [17]:

$$R(k) = \prod_{i=1}^{m} R_i(k_i) = \prod_{i=1}^{m} \left(1 - (1 - p_i)^{k_i}\right) \tag{2}$$

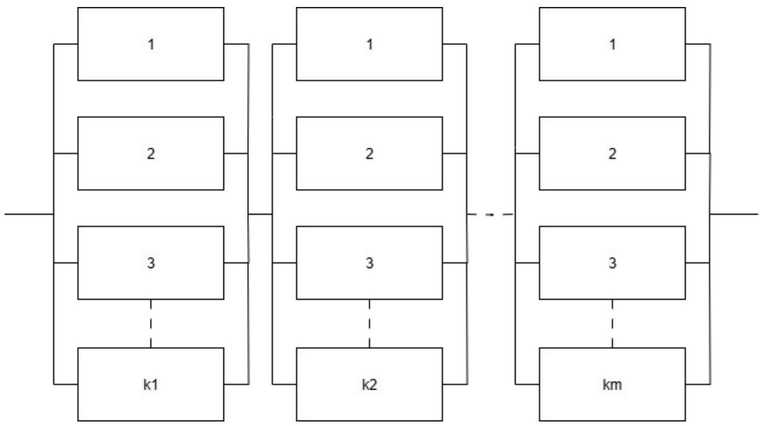

Figure 1: Series-parallel configuration.

3.3. Objective function

The task is formulated as maximising reliability with a given constraint:

$$\max_k R(k) \quad \text{provided that} \quad \sum_{i=1}^{m} c_i k_i \leq B \tag{3}$$

Alternative approach – minimising costs while achieving a specified level of reliability R_{min}:

$$\min_k C(k) \quad \text{provided that} \quad R(k) \geq R_{min} \tag{4}$$

In this work, the focus is on the task of maximising reliability within a limited budget, as this is most typical scenario in system design practice.

Specifics of the task:

1. Nonlinearity of the objective function. The reliability function $R(k)$ is nonlinear and has the property of diminishing marginal effect: each new reserve adds less to reliability than the previous one.
2. Discreteness of variables. The values of k_i must be integers, which complicates the use of classical optimisation methods for continuous problems.

3. NP complexity. The problem of optimal reservation usually belongs to the class of NP-hard. With a large number of subsystems, a brute force of possible options becomes practically impossible.
4. Practical importance. Even a slight increase in system reliability can be critical for security, power supply, or data centre operation. Therefore, the task of optimal redundancy has not only theoretical but also high practical value.

4. Methods for solving

4.1. General provisions

Optimisation problems, such as optimal reservation in series-parallel systems, are characterised by non-linearity of the objective function, discreteness of variables, and hard constraints. Therefore, both classical numerical methods and metaheuristic methods capable of working effectively in complex solution spaces are used to solve them. Various approaches are used to solve the optimal reservation problem.

One of the classic approaches is gradient methods, which are based on using derivatives of the objective function to determine the direction of movement towards the optimum. They work well for continuous and smooth problems, but in cases of discrete variables and constraints, they require special modifications.

Another important approach is Dual Ascent, which allows constraints to be taken into account effectively. This method is useful when constraints have a linear structure, such as budget constraints. Dual Ascent gradually adjusts the multipliers, bringing the solution closer to the optimal one, and is therefore suitable for practical problems with resource constraints.

Metaheuristic methods, including Particle Swarm Optimisation, also play an important role. Their advantage is that they do not require information about derivatives and are capable of effectively exploring large and complex solution spaces. This makes them particularly suitable for tasks involving nonlinear or many local extrema.

Hybrid approaches, which combine the strengths of different methods, constitute a separate direction. For example, the use of Dual Ascent to accurately account for the constraint multiplier in combination with PSO as an internal optimiser allows both the fulfilment of conditions to be guaranteed and high search quality to be achieved in complex solution spaces.

Thus, in the following sections, a sequence of key methods applied in this work will be considered.

4.2. Gradient methods

Gradient methods are among the most common numerical methods for solving optimisation problems. Their key idea is to use the direction of the gradient of the objective function to find the point at which this function takes on its maximum or minimum value. In the case of maximisation, the step is taken in the direction of the gradient, and in the case of minimisation, it is taken in the opposite direction.

Let 's have a function $f(x)$, where $x \in R^n$. The gradient is defined as a vector of partial derivatives:

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) \tag{5}$$

The classic gradient descent algorithm (for minimisation) is defined by the following formula [18]:

$$x^{(t+1)} = x^{(t)} - \alpha \, \nabla f\left(x^{(t)}\right) \tag{6}$$

where $\alpha > 0$ is the step size (learning rate). To maximise, the formula with the opposite sign is used.

Advantages:

- Simple implementation and interpretation.
- Good efficiency for smooth and convex functions.
- Possibility of applying modifications (e.g. stochastic gradient descent).

In optimal reservation problems, the reliability function is nonlinear and defined for integer variables k_i. The use of gradient methods in their "pure" form is complicated: the variables are discrete, which requires a special approach (e.g., rounding of values), constraints are imposed, which requires appropriate approaches (projections), the objective function is saturating (additional reserve gives an increasingly smaller increase in reliability), which can lead to slow convergence.

Despite these limitations, gradient methods remain an important instrument, as they allow for initial analysis of function behaviour and sensitivity to variable changes, and can be used as a basic component in further algorithms.

Thus, gradient methods play a fundamental role in solving reliability optimisation problems and serve as a starting point for more complex methods, such as Dual Ascent.

4.3. Dual Ascent

The Dual Ascent method belongs to the class of numerical optimisation methods with constraints, which allows solving primal and dual problems. Its main idea is to move from directly solving a problem with constraints to its dual form, where constraints are taken into account using multipliers. This allows optimisation to be performed in two steps: first, the inner problem without constraints (for fixed multipliers) is solved (e.g., the gradient descent method), and then the multipliers themselves are gradually updated to ensure that the constraints are satisfied.

Let 's consider the following problem:

$$\max_x f(x) \quad \text{provided that} \quad g(x)=0 \tag{7}$$

Forming the Lagrangian:

$$L(x,u)=f(x)+u^T g(x) \tag{8}$$

where u is the Lagrange multiplier responsible for enforcing the budget constraint.

Dual Ascent Algorithm [19]:

1. Initialise the multiplier $u^{(0)}$ and $x^{(0)}$.
2. Repeat the following steps, where t = 1, 2, … denotes the step number:

    a. Update the step size τ_t.
    b. For the current $u^{(t)}$, solve the internal problem:

$$x^{(t+1)} \in \operatorname*{argmin}_x L\left(x, u^{(t)}\right) \tag{9}$$

    c. Update the multiplier:

$$u^{(t+1)} = u^{(t)} + \tau_{t+1} g\left(x^{(t+1)}\right) \tag{10}$$

Features:

- Dual Ascent is useful when the main constraint has a simple linear form (e.g., budget).
- Allows to effectively take into account limitations and adjust the search for the optimal solution.
- The update step τ_t is often made variable: large at the beginning for fast movement and smaller at the end for stability.

In our work, the goal is to maximise system reliability within cost restrictions. Therefore, internal optimisation naturally takes the form of argmax rather than argmin. At the same time, when describing the method itself, we adhere to the classical form with minimisation, as accepted in the literature, and emphasise that the formulation may vary depending on the setting of the task.

In the formulation, the problem has a clear budget constraint: $\sum c_i k_i = B$. Dual Ascent allows to work effectively with this type of conditions. Internal optimisation finds the combination of reserves that best increases reliability for a given penalty u, while external factor updating gradually adjusts the solution to exact budget compliance. Another advantage of this approach is that it can be combined with other methods (e.g., PSO for internal optimisation, as discussed later in this paper).

4.4. Particle swarm optimization

Particle Swarm Optimisation is a computational optimisation method inspired by the collective behaviour of living organisms, such as flocks of birds or schools of fish [20]. Its main idea is that a set of agents (particles) moves in the space of possible solutions, adjusting their trajectory based on their own experience and the experience of their neighbours. PSO is a metaheuristic method because it does not require significant assumptions about the optimisation problem and is capable of searching large spaces of possible solutions. The PSO method is a swarm intelligence method because it mimics collective behaviour, where particles exchange information and adapt to the environment to find optimal solutions.

Let 's say there is a population of N particles, each of which is characterised by:

- position $x_i(t)$,
- velocity $v_i(t)$,
- the best solution found for this particle p_i,
- the globally best solution for the population g.

The update rules are as follows [21]:

$$v_i^{(t+1)} = w\, v_i^{(t)} + c_1 r_1 \left(p_i - x_i^{(t)}\right) + c_2 r_2 \left(g - x_i^{(t)}\right) \tag{11}$$

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)} \tag{12}$$

where w is the inertia coefficient that controls the "memory" of speed; c_1, c_2 are the coefficients of the "cognitive" and "social" components; r_1, $r_2 \sim Rnd(0,1)$ are random coefficients that add stochasticity.

Advantages of PSO:

- Does not require information about the derivatives of the objective function.
- Works well for problems with nonlinear or complex search spaces.

- Has the ability to avoid local optima due to random perturbations.
- Simple implementation and few parameters to configure.

In our task, the objective function is the probability of fault-free operation of the system, which has a nonlinear and complex form. PSO, thanks to its ability to effectively explore the space of possible configurations, is well suited for the role of an internal optimiser. Particles can represent different distributions of reserves $k = (k_1, k_2, \ldots, k_m)$, and the swarm gradually finds the configuration that maximises the reliability of the system for a given budget.

4.5. Hybrid approach

Our work uses a hybrid method that combines Dual Ascent and Particle Swarm Optimisation. The idea is that Dual Ascent is responsible for updating the u multiplier and enforcing the budget restriction, while PSO is responsible for solving the internal optimisation Lagrangian problem.

Scheme of operation:

1. The Lagrangian is formulated, which in the problem has the form:

$$L(k,u)=R(k)-u\left(\sum_{i=1}^{m} c_i k_i - B\right) \tag{13}$$

where $R(k)$ is the reliability of the system, c_i is the cost of subsystem element i, B is the budget, and u is the multiplier.

2. For a fixed value of u, it is necessary to solve the internal maximisation problem using PSO:

$$k^{(t+1)} \in \underset{k}{argmax}\, L\left(k, u^{(t)}\right) \tag{14}$$

Therefore, with a fixed u, iteration or iterations are performed using the PSO method. Each particle in PSO represents a possible configuration of reserves k.

3. The step size τ_t can be adaptive, for example, decreasing with increasing number of iterations to ensure stable convergence.

4. After receiving $k^{(t)}$, the u multiplier is updated according to the Dual Ascent rule:

$$u^{(t+1)} = u^{(t)} + \tau_t \left(\sum_{i=1}^{m} c_i k_i^{(t)} - B \right) \tag{15}$$

5. Move again with the new fixed value u to optimisation using the PSO method. The process is repeated until the stop condition is reached.

Advantages of this approach:

- Dual Ascent ensures that budget constraints are taken into account.
- PSO effectively explores the solution space and finds good configurations even in complex conditions.
- The combination of methods allows to simultaneously control resource constraints and achieve high system reliability.
- The approach is simply scalable for large systems.

In the optimal reservation problem, this approach naturally fits the structure of the problem: Dual Ascent acts as a “regulator” that determines the multiplier and acts externally, while PSO searches for the optimal amount of reserves in each subsystem in the internal problem. The result is an approach that simultaneously ensures optimal resource allocation and a high probability of fault-free system operation.

5. Experimental research

To demonstrate the performance of the proposed hybrid approach, a numerical experiment was conducted with the task of optimal reservation in a series-parallel structure. The aim of the experiment is to compare the effectiveness of the Dual Ascent + PSO hybrid approach with the classic Dual Ascent variant, where the internal maximisation task is solved by conventional gradient descent. Thus, it is possible to see not only the quality of the final solution, but also the speed at which it is achieved and the stability of the results in different runs.

The same optimal reservation problem formulation was used for both methods. The values of the probabilities of failure-free operation of elements, costs, and budget were set as follows: $P = (0.92, 0.90, 0.95)$, $C = (20, 3, 10)$, $B = 150$.

The initial value of the multiplier was set as $u = 0.1$. For gradient descent in the inner problem, a step size $\alpha = 10$ was chosen. For the hybrid method with using PSO, the swarm parameters were determined: population size *swarm_size* = 10, inertia coefficient $w = 0.7$, cognitive and social coefficients $c_1 = 1.4$, $c_2 = 1.4$.

An important feature of the hybrid approach is that PSO requires search limits. In our task, they are defined naturally: for each element, the minimum value of the number of reserves is $k_i^{min} = 1$, since at least one element must be present, and the maximum value is limited by the budget, i.e. $k_i^{max} = \frac{B}{C_i}$, since the number of reserves at one stage cannot exceed the total budget divided by the cost of one element of this stage. This property makes the optimal reservation problem very convenient for applying PSO, since the search space constraints are easy to set and have a clear engineering rationale.

The same mechanism for updating the u multiplier was used for both methods. Instead of a constant step size τ, an adaptive pattern was chosen, which ensures a reduction in the step size during iterations:

$$\tau_t = max\left(\tau_{min}, \tau_0 \cdot (decay)^t\right) \tag{16}$$

where τ_{min} is the minimum permissible value, τ_0 is the initial step, and decay is the reduction coefficient. $\tau_0 = 0.001$, $decay = 0.99$, $\tau_{min} = 10^{-8}$. Both methods worked with the same values for these parameters.

Each method was run for 10000 iterations, and the best result was selected from all intermediate solutions. Since in the case of PSO the final result may depend on the initial generation of the swarm, this method was run 100 times, after which the results were averaged. Due to the stochastic nature of the PSO algorithm, slight variations between runs are expected, and averaging helps to ensure stable and reliable outcomes. Considering the stochastic nature of the method, the analysis focused on the quality, convergence, and stability of the obtained solutions rather than on runtime performance. For reliability of the assessment, the standard deviation was also calculated. All configurations obtained by both methods after completion of the calculations were rounded down to ensure integer values for the number of reserves. The reliability function was evaluated using these adjusted integer values.

The results showed that the hybrid Dual Ascent + PSO method achieved better results compared to the conventional Dual Ascent with gradient descent. The hybrid method found the optimal configuration (4,6,5), while the classical method found the solution (4,5,4). The reliability of the obtained systems was as follows:

$$R_{hyb} = 0.9999577275540725, \text{ versus } R_{clas} = 0.9999427907280976 \tag{17}$$

The difference in numerical terms may seem insignificant, but it has practical significance, as it relates to the level of fault tolerance in critical systems.

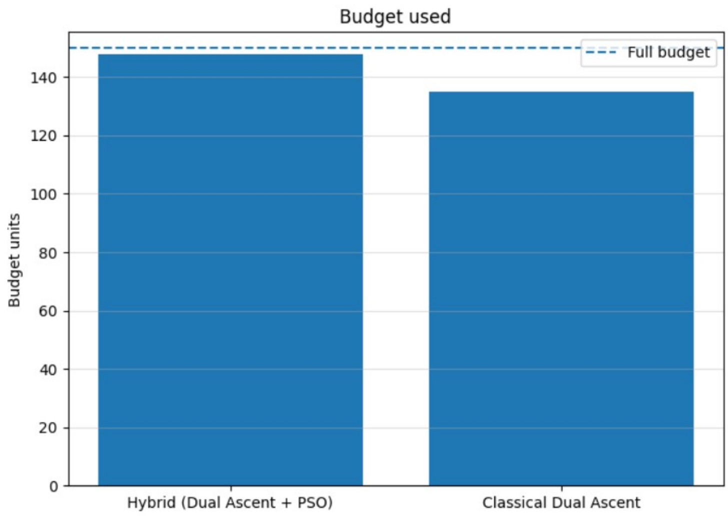

Figure 2: Budget used.

In addition, the hybrid method made more efficient use of the available budget, spending 148 units out of a possible 150, while the classical method spent only 135. Using almost the entire budget is an advantage, since the remaining resources require an additional brute force search to find even better options. However, this search is much simpler, since the remaining budget in such cases is small ($B_{remaining} \ll B_{initial}$). It is important to note that even after such additional analysis, the configurations obtained did not exceed the reliability result found by the hybrid approach.

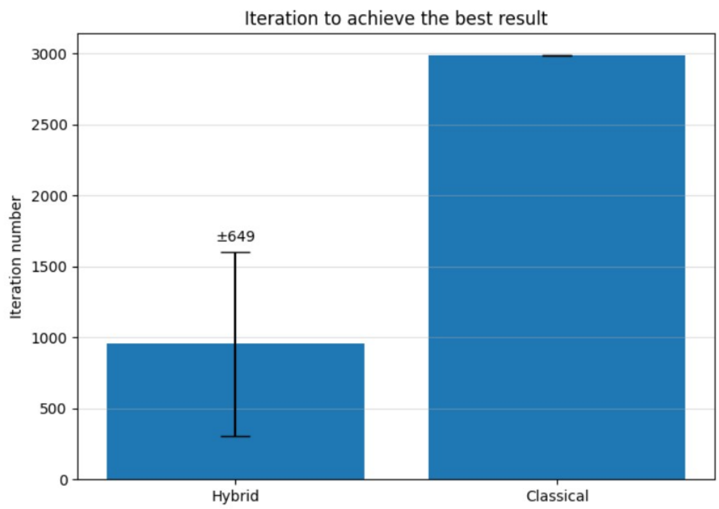

Figure 3: Iteration to achieve the best result.

The behaviour of the methods in terms of the speed of reaching the best solution is also interesting. The hybrid approach found the optimal configuration much earlier – on average at the 954th iteration, with a standard deviation of 649, while the classic Dual Ascent with gradient descent achieved its best result only at the 2989th iteration. This means that the hybrid method not only provides better quality, but also finds the optimal solution first.

Another important feature is the stability of the results obtained. All 100 independent runs of the hybrid algorithm resulted in the same optimal system configuration. This demonstrates the high reliability of the method: despite the stochastic nature of PSO, the final solution proved to be stable and repeatable.

In summary, it can be concluded that the hybrid Dual Ascent + PSO approach outperforms the classic Dual Ascent with gradient descent in terms of solution quality, iteration speed, and stability of results. This confirms the feasibility of using metaheuristic algorithms within internal optimisations for optimal reservation tasks, where both the fulfilment of constraints and effective search in complex solution spaces are important.

6. Conclusions

Optimal reserving is one of the key tasks of reliability theory and practical design of critical IT systems. The correct distribution of reserves affects not only overall fault tolerance but also the efficient use of limited resources. The task is complicated by its non-linear nature, working with discrete variables and budget constraints. Therefore, the choice of effective optimisation methods is extremely important from both a theoretical and applied point of view.

This paper considers and compares two approaches to solving the problem of optimal redundancy in a series-parallel structure. The first is the classic Dual Ascent method, which uses gradient descent for the internal optimisation problem. The second is the proposed hybrid method, where the internal problem is solved using the metaheuristic method Particle Swarm Optimisation. PSO is a form of swarm intelligence that models the collective behaviour of agents and does not require the calculation of objective function derivatives.

The numerical experiment demonstrated the advantages of the hybrid approach. It provided a better level of system reliability (0.9999577 vs. 0.9999428), more complete use of the available budget (148 vs. 135 units), and the first to find the optimal solution (on average at the 954th iteration vs. the 2989th in the classical approach). It is also important that all independent runs of the hybrid method led to the same optimal configuration, which indicates its stability and reliability.

Thus, the Dual Ascent + PSO hybrid approach has proven its effectiveness in the task of optimal reservation. It allows to simultaneously control the implementation of resource constraints and ensure high-quality search in a complex solution space, which makes it a promising instrument for use in the design of systems where reliability is critical.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] Francisco G. Montoya, Raúl Baños Navarro. Optimization Methods Applied to Power Systems. Volume 1. 382 p. https://doi.org/10.3390/books978-3-03921-131-9.

[2] Sun, Shiliang & Cao, Zehui & Zhu, Han & Zhao, Jing. (2019). A Survey of Optimization Methods from a Machine Learning Perspective. 10.48550/arXiv.1906.06821.

[3] Li D, Xu P, Gu J, Zhu Y. A Review of Reliability Research in Regional Integrated Energy System: Indicator, Modeling, and Assessment Methods. Buildings. 2024; 14(11):3428. https://doi.org/10.3390/buildings14113428.

[4] Modarres, M., Kaminskiy, M.P., & Krivtsov, V. (2016). Reliability Engineering and Risk Analysis: A Practical Guide, Third Edition (3rd ed.). CRC Press. https://doi.org/10.1201/9781315382425.

[5] Viknesh Andiappan, Michael Francis D. Benjamin, Raymond R. Tan, Denny K.S. Ng, Design, optimisation and reliability allocation for energy systems based on equipment function and operating capacity, Heliyon, Volume 5, Issue 10, 2019, e02594, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2019.e02594.
[6] Morozov, V., Deineha, V., Kovalchuk, D., (2024), Multi-Criteria Method for Comparing the Effectiveness of Gradient Descent Modifications on Benchmark Functions, CEUR Workshop Proceedings, 3909, pp. 123–137.
[7] Guan, B., Li, Z., Coit, D. W., & Li, Y. F. (2025). Review of the redundancy allocation problem to optimize system reliability. Engineering Optimization, 57(1), 44–68. https://doi.org/10.1080/0305215X.2024.2447078.
[8] W. Kuo and V. R. Prasad, "An annotated overview of system-reliability optimization," in IEEE Transactions on Reliability, vol. 49, no. 2, pp. 176-187, June 2000, doi: 10.1109/24.877336
[9] Marouani, H., Al-mutiri, O. (2022). Optimization of Reliability–Redundancy Allocation Problems: A Review of the Evolutionary Algorithms. Computers, Materials & Continua, 71(1), 537–571. https://doi.org/10.32604/cmc.2022.020098.
[10] Aqel I, Mellal MA. Optimal reliability allocation of heterogeneous components in pharmaceutical production plant. Int J Interact Des Manuf. 2023 Mar 1:1-10. doi: 10.1007/s12008-023-01256-1. Epub ahead of print. PMID: 40479467; PMCID: PMC9975440.
[11] Marouani, Haykel. (2021). Optimization for the Redundancy Allocation Problem of Reliability Using an Improved Particle Swarm Optimization Algorithm. Journal of Optimization. 2021. 1-9. 10.1155/2021/6385713.
[12] Ashraf, Zubair & Shahid, Mohammad & Ahmad, Faisal. (2021). Gradient Based Optimization Approach to solve Reliability Allocation System. 337-342. 10.1109/ICCCIS51004.2021.9397197.
[13] Surapati, Ramadevi & Akiri, Sridhar & Velampudi, Srinivasa & Nagarjuna, Vasili B V & Swethan, Potluri & Kapu, Bhavani. (2025). Enhancing Reliability of Series-Parallel Systems: A Novel Mathematical Model for Redundancy Allocation. European Journal of Pure and Applied Mathematics. 18. 5672. 10.29020/nybg.ejpam.v18i1.5672.
[14] Leon, Florin & Caşcaval, Petru & Badica, Costin. (2020). Optimization Methods for Redundancy Allocation in Large Systems. Vietnam Journal of Computer Science. 7. 10.1142/S2196888820500165.
[15] Yuan Shen, Chang Liu, Yannian Zuo, Xingying Zhang. A modified self-adaptive dual ascent method with relaxed stepsize condition for linearly constrained quadratic convex optimization. Journal of Industrial and Management Optimization, 2023, 19(5): 3637-3649. doi: 10.3934/jimo.2022101.
[16] Yang, Y.; Wen, F.; Yang, J.; Liu, H.; Liu, D.; Xin, S.; Fan, H.; Wu, C. Real-Time Power Regulation of Flexible User-Side Resources in Distribution Networks via Dual Ascent Method. Energies 2024, 17, 4890. https://doi.org/10.3390/en17194890.
[17] Muhuri, Pranab & Ashraf, Zubair & Lohani, Q. M. Danish. (2018). Multi-objective Reliability-Redundancy Allocation Problem with Interval Type-2 Fuzzy Uncertainty. IEEE Transactions on Fuzzy Systems. 26. 1339-1355. 10.1109/TFUZZ.2017.2722422.
[18] Morozov, Viktor & Deineha, Vladyslav & Yeremenko, Bohdan. (2025). Development of energy consumption prediction models based on gradient descent methods. 1-7. 10.1109/SIST61657.2025.11139299.
[19] Overview of Dual Ascent, Adam Rumpf, Illinois Institute of Technology, Department of Applied Mathematics, February 19, 2019.
[20] J. Kennedy and R. Eberhart, "Particle swarm optimization," Proceedings of ICNN'95 - International Conference on Neural Networks, Perth, WA, Australia, 1995, pp. 1942-1948 vol.4, doi: 10.1109/ICNN.1995.488968.
[21] Pereira, Gonçalo. (2011). Particle Swarm Optimization.