Journal of Information & Systems Management

DLINE JOURNALS

ISSN: 2230 – 8776

JISM 2026: 16 (4)
https://doi.org/10.6025/jism/2026/16/4/163-189

Machine Learning-Assisted Multi-Objective Optimization of 2D Elastodynamic Metamaterials: Pareto Analysis, Algorithmic Comparison, and Robustness Assessment

Maleerat Maliyaem
College of Industrial Technology
Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok
Thailand

ABSTRACT

The design of 2D elastodynamic metamaterials requires navigating complex, high dimensional spaces to balance competing functional objectives, such as minimizing bandgap location while maximizing bandgap width. Traditional finite element analysis (FEA) is computationally prohibitive for vast design spaces, and current data-driven optimization studies frequently overlook the empirical variability of simulation data, risking the selection of statistical artifacts. This study presents a catalogue-constrained, machine learning-assisted multi-objective optimization (ML-MOO) framework utilizing a dataset of 1,400 unique binary metamaterial geometries. We developed an Extra Trees surrogate model and compared six optimization algorithms (Bayesian Optimization, PSO, Differential Evolution, MOEA, GA, and NSGA-II). Results indicate that MOEA, BO, PSO, and DE converge on an identical global best utility (0.8176), with MOEA demonstrating superior sample efficiency. The optimized designs achieved up to a 547.7% increase in bandgap width compared to the dataset baseline, albeit with a 9.8% increase in central frequency, illustrating the inherent physical trade-off. Crucially, a robustness assessment revealed that all 52 global Pareto optimal geometries are singletons lacking repeated observations, while replicated geometries exhibited high empirical variability (Coefficient of Variation up to 173% for bandgap width). This highlights a critical distinction between computational optimization performance and empirical reproducibility. The study concludes by proposing a conservative 12-point engineering shortlist and advocating for future robustness aware optimization frameworks that incorporate variance penalties to ensure physically reliable metamaterial designs.

Keywords: Elastodynamic Metamaterials, Multi-Objective Optimization (MOO), Machine Learning, Pareto Front Analysis, Bayesian Optimization, Surrogate Modeling, Design Robustness, Bandgap Engineering, Evolutionary Algorithms

Received: 11 January 2026, Revised 30 May 2026, Accepted 16 June 2026

Copyright: **DLINE**

1. Introduction

Machine learning (ML), a subfield of artificial intelligence, focuses on developing computational algorithms capable of learning patterns and relationships from data and improving their predictive performance without requiring explicit programming for every task. By extracting meaningful information from complex datasets, ML techniques facilitate prediction, trend identification, pattern recognition, and the solution of computationally challenging problems [1, 2]. Their ability to approximate complex relationships and process high-dimensional data has made them increasingly valuable across engineering and materials science.

In recent years, machine learning techniques have emerged as important tools for the design, analysis, and optimization of metamaterials. The development of metamaterials frequently involves exploring high-dimensional design spaces characterized by complex interactions between structural configurations, material properties, and functional responses. Conventional design and optimization approaches can become computationally expensive and inefficient when investigating such complex relationships, particularly when extensive numerical simulations are required [3]. Machine learning provides an alternative approach by learning the underlying relationships between structural parameters and their corresponding functional properties. This capability enables rapid performance prediction, efficient design space exploration, optimization, and inverse design, thereby reducing dependence on repeated computational simulations and accelerating the overall design process.

The application of ML to metamaterial design has already demonstrated considerable potential across different engineering domains. Generative machine learning models, for example, have been employed to design metamaterials exhibiting prescribed mechanical properties, with high agreement between the targeted mechanical responses and experimentally measured results [4]. Similarly, ML-based approaches have been applied to the design and optimization of thermal metamaterials to improve their performance and functionality [5]. In phononic metamaterials, machine learning techniques have contributed to the development of structures with tailored acoustic and elastic properties [6]. Collectively, these developments demonstrate how data-driven approaches can expand the exploration of material configurations, reduce computational requirements, and facilitate the discovery of metamaterials with desirable functional characteristics.

The increasing adoption of artificial intelligence in metamaterial research has also encouraged efforts to systematically review and consolidate existing developments. Previous review studies have investigated AI applications across different categories of metamaterials, including acoustic, photonic, electromagnetic, and other functional material systems. In addition to broad interdisciplinary reviews, specialized investigations have examined particular metamaterial categories while considering their associated design and manufacturing constraints. Comprehensive reviews addressing these developments are available in the studies by Jiao, P., Boafo, G., Garg, A., and Zhang, C. [7, 8, 9, 3]. These contributions establish the growing importance of computational intelligence in metamaterial research and provide a foundation for further investigation into data-driven material design and optimization.

2. Background

2.1 Optimization Techniques and Their Computational Limitations

Despite the advantages of machine learning in predicting material behavior, the optimization of metamaterial structures remains a challenging computational problem. The objective functions associated with metamaterial design may exhibit nonlinearity, discontinuities, multiple local optima, and complex interactions among design parameters. Consequently, the selection of an appropriate optimization algorithm is essential for achieving efficient exploration of the design space and obtaining reliable solutions.

Gradient-based optimization methods, particularly stochastic gradient descent, have demonstrated potential in machine learning applications by optimizing objective functions over smoothed solution spaces [10] [Kotwal, J. G.]. Nevertheless, their performance can be sensitive to hyperparameter selection and initial conditions. Furthermore, gradient-based methods may converge to local minima rather than identifying globally

competitive solutions, particularly when the optimization landscape is highly non-convex [11]. These limitations can restrict their applicability to complex engineering problems in which objective functions may not possess convenient mathematical properties.

Gradient-free optimization approaches provide an alternative when derivative information is unavailable or unreliable. The simplex method, for example, does not require gradient information; however, it can also experience difficulties in identifying global optima within highly non convex optimization spaces [12]. These limitations have motivated the application of stochastic and population based optimization algorithms that can explore broader regions of the design space.

Among the commonly employed stochastic optimization approaches, Genetic Algorithms (GA), Differential Evolution (DE), and Particle Swarm Optimization (PSO) have demonstrated considerable effectiveness in global optimization problems [13, 14, 15]. Unlike gradient-dependent approaches, these algorithms do not require derivative information and can accommodate discontinuous objective functions. Their population-based search mechanisms facilitate exploration of complex solution spaces and provide opportunities to identify multiple competitive solutions.

However, the effectiveness of population-based stochastic algorithms is accompanied by substantial computational requirements. Their optimization procedures generally involve repeated evaluations of objective functions across multiple candidate solutions and generations. Consequently, the number of function evaluations can become prohibitively large when each evaluation requires a computationally expensive numerical simulation. This limitation is particularly important in electromagnetic and other simulation-intensive engineering applications, where repeated evaluations may substantially increase the overall optimization time.

To address these computational challenges, surrogate assisted optimization techniques have been developed. These approaches construct computationally inexpensive approximations of expensive objective functions and use the resulting surrogate models to guide the optimization process. Bayesian Optimization (BO) is an important example of this strategy. It commonly employs Gaussian Processes (GP) to construct probabilistic approximations of objective functions and uses acquisition functions to determine promising locations for subsequent evaluations [16]. By balancing exploration and exploitation, Bayesian Optimization can identify promising solutions while reducing the number of expensive objective function evaluations.

Although Bayesian Optimization can be data-efficient, its application to high dimensional optimization problems presents scalability challenges. In particular, updating Gaussian Process models involves computational complexity of (O(n^3)), where (n) represents the number of function evaluations. As the number of evaluated design configurations increases, the computational cost associated with updating the surrogate model can become substantial. These limitations highlight the need to consider not only the number of objective function evaluations but also the computational overhead associated with the optimization algorithm itself.

2.2 Multi-Objective Optimization and Pareto-Based Approaches

The challenges associated with optimisation become more pronounced in multi-objective optimisation (MOO), where multiple competing performance criteria must be considered simultaneously. In engineering design, improving one performance characteristic may adversely affect another. Therefore, identifying a single solution that simultaneously optimizes every objective is not always feasible. Instead, the optimization process seeks a collection of non-dominated solutions representing different trade-offs among the competing objectives.

Pareto-based optimization provides a systematic framework for identifying these trade offs. A solution is considered Pareto optimal when no objective can be improved without deteriorating at least one of the other objectives. The collection of such solutions forms the Pareto front, which provides valuable information about the available trade offs within the design space. This information enables designers to compare alternative configurations and select solutions according to their specific engineering requirements.

Among the established multi-objective optimization techniques, the Non-dominated Sorting Genetic Algorithm II (NSGA-II) and the Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D) are widely used for identifying diverse Pareto-optimal solutions in engineering and industrial applications [17]. NSGA-II uses non-dominated sorting and diversity-preservation mechanisms to maintain a population of competitive solutions across different regions of the Pareto front. In contrast, MOEA/D decomposes a multi-objective optimization problem into a set of scalar optimization subproblems, facilitating the simultaneous exploration of different regions of the objective space.

Both approaches provide effective mechanisms for addressing competing objectives and exploring alternative design configurations. Nevertheless, because they are based on evolutionary search procedures, they retain the computational limitations associated with repeated objective function evaluations. This issue becomes particularly significant when optimization is coupled with computationally intensive physical simulations. Consequently, the development and assessment of optimization strategies that can maintain solution quality while reducing computational requirements remain important considerations in multi-objective engineering design.

2.3 Application of Optimization to Graded Phononic Metamaterials

Recent research on graded phononic metamaterials illustrates the relevance of combining physical modeling with optimization algorithms to investigate complex material behavior. Li Q. [18] established a power law gradient model to examine the influence of power law exponents on phononic bandgaps, considering the effects of unit-cell symmetry and biased filling fraction. In this approach, the transfer matrix of the graded material was constructed using the exponential of the Magnus series, and the corresponding dispersion relation was derived through the Transfer Matrix Method (TMM).

The study further integrated the transfer matrix formulation with Genetic Algorithms (GA) and the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to investigate both single objective and multi-objective optimization of the first bandgap. The results provide practical guidance for selecting appropriate power-law exponents and understanding the trade-offs between bandwidth and central frequency in graded phononic metamaterial design [Li Q,].

This investigation demonstrates the potential of integrating analytical physical models with evolutionary optimization algorithms to improve metamaterial design. In particular, the combination of transfer matrix modeling and optimization enables the systematic investigation of structural parameters and their influence on bandgap characteristics. It also illustrates how multi-objective optimization can support the identification of alternative design configurations when different performance characteristics must be balanced.

2.4 Research Motivation and Scope

The reviewed literature demonstrates substantial progress in the application of machine learning, evolutionary optimization, and surrogate assisted methods to metamaterial design. Machine learning provides opportunities for learning complex structure property relationships, while stochastic and evolutionary algorithms facilitate the exploration of nonlinear and potentially discontinuous optimization spaces. Surrogate assisted methods offer an additional mechanism for reducing the computational burden associated with expensive objective function evaluations. Furthermore, Pareto based approaches provide a framework for analyzing competing design objectives and identifying alternative optimal configurations.

Nevertheless, the computational limitations of existing optimization approaches remain an important consideration, particularly when multiple design objectives and simulation intensive evaluations are involved. Gradient based methods may be affected by local minima and sensitivity to initialization, evolutionary algorithms may require numerous objective function evaluations, and Gaussian Process based Bayesian Optimization may encounter scalability limitations as the number of evaluations increases. These considerations motivate systematic assessment of the performance, computational efficiency, and robustness of alternative optimization strategies.

Against this background, the present research focuses on machine learning assisted multi-objective optimization of two dimensional elastodynamic metamaterials. The study is positioned around the investigation of Pareto optimal design solutions, comparative assessment of optimization algorithms, and evaluation of solution robustness. By bringing together data-driven modeling and multi objective optimization, the research aims to support a more systematic understanding of the relationships between design parameters and competing performance objectives. The resulting analysis is intended to provide insights into algorithmic performance, optimization trade-offs, and the reliability of the identified design solutions, thereby contributing to more efficient computational design of two-dimensional elastodynamic metamaterials.

2.5 Problem Statement

The design of elastodynamic metamaterials requires navigating a highly complex, non-linear, and high-dimensional design space where structural geometry dictates functional mechanical responses, such as elastic bandgap characteristics. Traditional design methodologies relying on iterative Finite Element Analysis (FEA) are computationally prohibitive when exploring vast geometric configurations. While Machine Learning (ML) and Multi-Objective Optimization (MOO) offer pathways to accelerate this discovery process, significant challenges remain.

First, metamaterial design inherently involves competing objectives; for instance, minimizing the central frequency of a bandgap (location) often conflicts with maximizing its frequency range (width). Identifying the optimal trade offs (Pareto front) requires sophisticated algorithmic search strategies. Second, and more critically, current data-driven optimization studies frequently treat simulation datasets as deterministic ground truth, overlooking the empirical variability and reproducibility of the underlying data. When optimization algorithms select “optimal” geometries from datasets containing uncontrolled simulation noise or lacking repeated observations, the resulting designs may be statistical artifacts rather than physically robust solutions. Therefore, the core problem this study addresses is twofold: (1) how to efficiently perform catalogue-constrained multi-objective optimization of 2D metamaterials using comparative ML and evolutionary algorithms, and (2) how to rigorously assess the empirical robustness and reproducibility of the resulting Pareto-optimal geometries given the inherent variability in the simulation data.

2.6 Research Design

This study employs a computational, data-driven research design structured around a Machine Learning-Assisted Multi-Objective Optimization (ML-MOO) framework. The research pipeline is divided into four sequential phases:

1. Data Ingestion and Preprocessing: Utilization of a public dataset of 2D elastodynamic metamaterials to establish a finite, observed design space (catalogue) of binary geometric configurations and their corresponding FEA-derived mechanical responses.

2. Surrogate Modeling: Development of an Extra Trees regression model to approximate the non-linear mapping between condensed binary geometry encodings and the continuous bandgap responses, serving as a predictive baseline.

3. Multi-Objective Optimization and Algorithmic Comparison: Deployment of six distinct optimization algorithms (Bayesian Optimization, PSO, Differential Evolution, MOEA, GA, and NSGA-II) to explore the discrete design space. The algorithms are constrained to the observed catalogue to prevent the assignment of unverified performance values to untested geometries.

4. Pareto Analysis and Robustness Assessment: Extraction of the global Pareto front to identify non-dominated trade-offs, followed by a statistical robustness analysis utilizing the Coefficient of Variation (CV) on replicated geometries to evaluate the empirical reliability of the optimized candidates.

3. Methodology

3.1 Dataset and Geometric Representation

The study utilizes the "2D Elastodynamic Metamaterials" dataset (UCI Machine Learning Repository; Ogren et al., 2021), comprising 20,520 observations of 10×10 pixel 2D elastic metamaterial unit cells. The geometric configurations are encoded as a condensed binary integer feature (CondensedBinary2DGeometry), yielding 1,400 unique geometries. The target variables, derived via FEA, are the central frequency location of the first bandgap (BandGapLocation, in Hz) and the width of that bandgap (BandGapWidth, in Hz). The dataset contains 67 exact duplicate rows and 128 replicated geometries (geometries with multiple observations), which are critical for the subsequent robustness analysis.

3.2 Multi-Objective Optimization Formulation

The optimization problem is formulated as a discrete, catalogue constrained multi-objective problem. The objectives are to minimize the bandgap location and maximize the bandgap width. $\min_x f_1(x)$ = BandGap Location(x)$\max_x f_2(x)$ = BandGapWidth (x) subject to: $x \in x_{observed}$ where x represents a binary geometric configuration, andis the catalogue of 1,400 unique geometries. The restriction to $x_{observed}$ ensures that optimization relies solely on physically simulated (observed) mean responses.

3.3 Surrogate Modeling

An Extra Trees regression model was trained to map the binary geometry to the mechanical responses. The model's predictive performance was evaluated using the coefficient of determination (), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) on a holdout set. While the surrogate model supports the overall modeling framework, the primary optimization algorithms utilized the measured geometry-level mean responses from the observed catalogue.

3.4 Optimization Algorithms and Evaluation Metrics

Six optimization algorithms were implemented with a strict search budget of 300 evaluations per algorithm: Bayesian Optimization (BO), Particle Swarm Optimization (PSO), Differential Evolution (DE), Multi-objective Evolutionary Optimization (MOEA), Genetic Algorithm (GA), and NSGA-II. Algorithmic performance was assessed across three dimensions:

1. Solution Quality: The best-balanced utility achieved.

2. Search Efficiency: The number of unique geometries evaluated before reaching the best solution.

3. Solution Diversity: The number of visited non-dominated (Pareto-optimal) configurations.

3.5 Robustness and Variability Analysis

To assess empirical reproducibility, the Coefficient of Variation (CV) was calculated for the 128 replicated geometries: $CV\ (\%) = \frac{\sigma}{\mu} \times 100$ where $\sigma$is the standard deviation and μ is the mean response of the repeated observations. This metric quantifies the within-geometry variability, highlighting the distinction between algorithmic optimization performance and physical reproducibility.

3.6 Dataset

The predictive modeling in this study utilizes the "2D Elastodynamic Metamaterials" dataset, publicly hosted by the UCI Machine Learning Repository (Dataset ID: 692, DOI: 10.24432/C5ZS5D)

Originally compiled by Ogren et al. (2021), the dataset contains 20,521 unique instances of 10×10 pixel 2D elastic metamaterial unit cells

The geometric configurations are encoded as a condensed binary integer feature (Condense dBinary2DGeometry), while the corresponding mechanical responses, simulated via Finite Element Analysis (FEA), are provided as two continuous target variables: the central frequency location of the first band gap (BandGapLocation, in Hz) and the width of that band gap (BandGapWidth, in Hz)

The dataset contains no missing values and was specifically designed to advance interpretable machine learning approaches for metamaterial discovery, enabling the identification of rule based shape frequency features without relying on opaque, data intensive black-box models

The uploaded dataset, 2D Elastodynamic Metamaterials, is well suited for substantial computational materials science and machine learning research. It can support analyses involving metamaterial geometry, elastic bandgap characteristics, geometry property relationships, predictive modelling, optimization, and explainable AI.

3.7 Dataset overview

Dataset successfully inspected

Characteristic	**Description**
Total observations	20,520
Total attributes	3
Geometry configurations	1,400 unique
Missing values	0
Exact duplicate rows	67
Geometry representation	Condensed binary encoding
Target variable 1	BandGapLocation
Target variable 2	BandGapWidth

3.8 Dataset structure and geometric representation

The study employs a binary representation of 2D metamaterial geometries, where the geometric configuration is encoded using the CondensedBinary2DGeometry attribute.

This representation allows geometric configurations to be processed computationally and used as inputs for machine-learning and optimization algorithms.

The corresponding mechanical responses are represented through two continuous variables: BandGapLocation and BandGapWidth.

The relationship between these variables and the geometric representation provides the foundation for the predictive and optimization framework.

An important consideration is that the geometry is represented numerically rather than through explicitly extracted spatial descriptors. Therefore, the machine-learning models learn relationships between the encoded geometry and the mechanical response without necessarily identifying the physical contribution of individual spatial cells.

The physical interpretation of individual binary positions requires confirmation of the original bit-to-spatial-cell mapping.

3.9 Multi-objective optimization problem formulation

The optimization problem considers two competing objectives: minimizing the bandgap location and maximizing the bandgap width.

The mathematical formulation can be expressed as $\min_x f_1(x)$ = BandGap Location(x) $\max_x f_2(x)$ = BandGapWidth (x): subject to: $x \in x_{observed}$

where:

x represents a binary geometric configuration.

$f_1(x)$represents the bandgap location.

$f_2(x)$represents the bandgap width.

$x_{observed}$ represents the catalogue of 1,400 observed geometric configurations

The restriction to observed configurations is an important methodological feature because it prevents the optimization process from assigning physically unverified performance values to geometries that were not included in the original dataset.

The resulting optimization problem is therefore a discrete, catalogue-constrained multi-objective optimization problem.

3.91 Role of surrogate modeling in optimization

The Extra Trees regression model provides an approximate mapping between binary geometry and the corresponding elastodynamic response.

The surrogate model can be expressed as: $\hat{y} = g(x)$

where X denotes the encoded geometry and represents the predicted bandgap response.

Separate predictions are generated for BandGapLocation and BandGapWidth.

The reported R^2 values demonstrate that the surrogate model captures a substantial proportion of the response variability.

However, the optimization results reported in the analysis are based on measured geometry-level mean responses within the observed catalogue. The surrogate model should therefore be interpreted as a predictive component supporting the overall modeling framework, rather than assuming that every optimizer-selected result was obtained directly from surrogate predictions.

This distinction should be explicitly maintained in the methodology and experimental setup.

3.92 Pareto dominance and optimality

A geometry is considered Pareto optimal when no other geometry in the evaluated design space provides a better or equal value for both objectives, with a strict improvement in at least one objective.

For two geometries A and B geometry dominates geometry

B when: f1 (A) $\leq$ f1 (B) f2(A) $\geq$ f2 (B)with at least one strict inequality.

This definition establishes the mathematical foundation for identifying the non-dominated geometric configurations.

The resulting Pareto front provides alternative solutions for different engineering requirements instead of imposing a single optimization preference.

3.93 Computational efficiency and optimization diversity

The optimization results reveal that the six algorithms differ in the number of unique configurations evaluated and the number of non-dominated solutions discovered.

This distinction is important because two algorithms can identify the same best-performing geometry while exhibiting different search efficiencies and solution diversity.

Therefore, the comparative assessment should consider three dimensions:

- Solution quality: The best balanced utility achieved.
- Search efficiency: The number of unique geometries evaluated before reaching the best solution.
- Solution diversity: The number of non-dominated configurations discovered.

These dimensions provide a more comprehensive assessment of optimization performance than the best utility value alone.

3.94 Robustness and reproducibility of optimized geometries

The robustness analysis provides an additional dimension to the optimization problem by considering the variability of repeated geometric configurations.

The coefficient of variation can be calculated as: where: $CV\ (\%) = \frac{\sigma}{\mu} \times 100$

- σ represents the standard deviation of repeated observations.
- μ represents the mean response of the corresponding geometry

A lower coefficient of variation indicates greater consistency among repeated observations, whereas a higher value indicates greater relative variability.

The analysis demonstrates that the available repeated geometries exhibit considerable variability.

More importantly, the 52 global Pareto-optimal geometries are all singletons. This means that their within-geometry variability cannot be estimated.

Consequently, optimization performance and robustness should be treated as separate evaluation criteria.

A future robustness-aware framework should prioritize configurations that demonstrate both competitive bandgap performance and reproducible responses across repeated simulations.

4. Results

We implemented all six optimization approaches for the 2D elastodynamic metamaterial dataset, using the following design objectives:

- Objective 1: Minimize BandGapLocation to identify lower-frequency bandgaps.

• Objective 2: Maximize BandGapWidth to obtain wider elastic bandgaps.

The analysis will compare Genetic Algorithm, NSGA-II, Particle Swarm Optimization, Bayesian Optimization, Differential Evolution, and multi-objective evolutionary optimization.

An important methodological decision is to treat the binary geometry as a discrete design representation and restrict optimization to the 1,400 observed geometry configurations. This avoids claiming that untested geometries have physically verified performance.

The analysis successfully processed 20,520 observations and 1,400 unique geometry configurations.

52	Pareto-optimal geometries
6	Optimization algorithms
0.806	Location prediction R^2
0.842	Width prediction R^2

Table 1. Dataset characteristics and structural overview

Table 1 presents the structural characteristics and data quality of the 2D Elastodynamic Metamaterials dataset used in this investigation. The dataset comprises 20,520 observations, three attributes, and 1,400 unique geometric configurations. The geometry is represented using a condensed binary encoding, while the two continuous response variables are BandGapLocation and BandGapWidth. The dataset contains no missing values, indicating that all observations are available for subsequent statistical analysis and predictive modeling. However, 67 exact duplicate rows were identified, suggesting the presence of repeated records that require consideration during data preprocessing.

The dataset provides a suitable foundation for investigating the relationship between metamaterial geometry and its corresponding elastodynamic properties. The relatively large number of observations compared with the number of unique geometries indicates that certain configurations have been evaluated repeatedly, potentially providing information about response variability and reproducibility.

The presence of 1,400 unique geometries also establishes a finite design space for optimization. Consequently, the optimization algorithms can explore and compare configurations within an experimentally or computationally observed geometry catalogue rather than relying on unverified geometric combinations.

An important observation is the difference between the number of observations and the number of unique geometries. Specifically, only 128 geometries have multiple observations, whereas 1,272 configurations occur once. This imbalance has important implications for estimating geometric robustness because empirical variability can only be calculated for configurations with repeated observations.

Overall, the dataset offers an appropriate basis for multi-objective optimization and predictive modeling, although the unequal replication of geometric configurations must be considered when interpreting optimization outcomes.

4.1 Algorithm comparison

Algorithm	Best balanced utility
Bayesian Optimization	0.8176
PSO	0.8176
Differential Evolution	0.8176
Multi-objective evolutionary optimization	0.8176
Genetic Algorithm	0.8087
NSGA-II	0.8035

Table 2. Algorithm Comparison

Table 2. Comparative performance of optimization algorithms

Table 2 compares the best-balanced utility achieved by six optimization algorithms: Bayesian Optimization, Particle Swarm Optimization (PSO), Differential Evolution (DE), multi-objective evolutionary optimization, Genetic Algorithm (GA), and NSGA-II.

Bayesian Optimization, PSO, Differential Evolution, and multi-objective evolutionary optimization achieved an identical best-balanced utility of 0.8176. Genetic Algorithm obtained a slightly lower value of 0.8087, whereas NSGA-II recorded the lowest value of 0.8035.

These findings indicate that four algorithms reached the same best-performing solution under the adopted balanced-utility evaluation criterion.

The results demonstrate that different optimization strategies can converge towards an equivalent high-quality geometric configuration despite differences in their search mechanisms.

The identical performance of Bayesian Optimization, PSO, Differential Evolution, and multi-objective evolutionary optimization suggests that the observed geometry catalogue contains an attractive solution that can be identified through multiple search strategies.

The comparatively lower utility values obtained by GA and NSGA-II indicate that these algorithms did not reach the same best-performing solution within the allocated search budget.

However, the differences between the algorithms are relatively small. The difference between the highest and lowest utility is approximately 0.0141, suggesting that the optimization performance is broadly comparable across the evaluated methods.

An important implication is that algorithm selection should not be based exclusively on the highest utility value. Computational efficiency, number of unique geometries evaluated, convergence behaviour, and diversity of identified non-dominated solutions should also be considered.

In particular, Bayesian Optimization achieved the highest utility after evaluating 300 unique geometries, whereas PSO, Differential Evolution, and multi-objective evolutionary optimization reached the same utility

using 269, 245, and 206 unique geometries, respectively (Table 4).

Thus, the results suggest that multi-objective evolutionary optimization and Differential Evolution may offer computational advantages when the objective is to identify the best observed balanced-utility configuration. The search budget was 300 evaluations per algorithm. The search operated on the observed geometry catalogue, using measured geometry level mean responses.

Best-performing geometry

Metric	Result
BandGapLocation	1101.009
BandGapWidth	1186.689
Number of observations	1
Pareto optimal	Yes

Table 3. Best-performing metamaterial geometry

Table 3 reports the mechanical performance of the best performing geometry identified through the optimization process. The selected configuration achieves a BandGapLocation of 1101.009 Hz and a BandGapWidth of 1186.689 Hz. The geometry is represented by a single observation and is identified as Pareto optimal.

The result indicates that the selected configuration exhibits a particularly wide bandgap while maintaining a bandgap location within the lower-frequency region of the observed design space.

The exceptionally large BandGapWidth indicates that the selected geometric configuration provides a substantial frequency interval over which elastic wave propagation is restricted. Such a characteristic is important for designing metamaterials intended for vibration suppression, elastic-wave filtering, and frequency-selective mechanical applications.

Compared with the dataset mean baseline, the selected configuration provides an approximately 547.7% increase in BandGapWidth. This represents a substantial improvement in the width of the frequency bandgap. However, the BandGapLocation increases by approximately 9.8% relative to the dataset mean. This indicates that the optimization does not simultaneously improve both objectives relative to the baseline.

The result therefore demonstrates the inherent trade-off between bandgap location and bandgap width. The selected geometry prioritizes a substantial increase in bandgap width while accepting a moderate increase in its central frequency location.

An important limitation is that this geometry has only one observation. Therefore, its empirical repeatability and sensitivity to simulation conditions cannot be established from the available dataset.

Furthermore, the selected configuration should be regarded as a promising computational candidate rather than a physically validated optimum. Additional finite element simulations are required to establish whether the observed performance can be reproduced. Compared with the dataset mean baseline:

- Bandgap width increased by approximately 547.7%.
- Bandgap location increased by approximately 9.8%.

This illustrates the trade off: the selected configuration strongly improves bandgap width but does not minimize bandgap location.

4.2 Important robustness

The dataset contains 128 replicated geometries and 1,272 geometries with only one observation.

Importantly, all 52 global Pareto-optimal geometries have only one observation. Consequently, their empirical repeatability cannot currently be established.

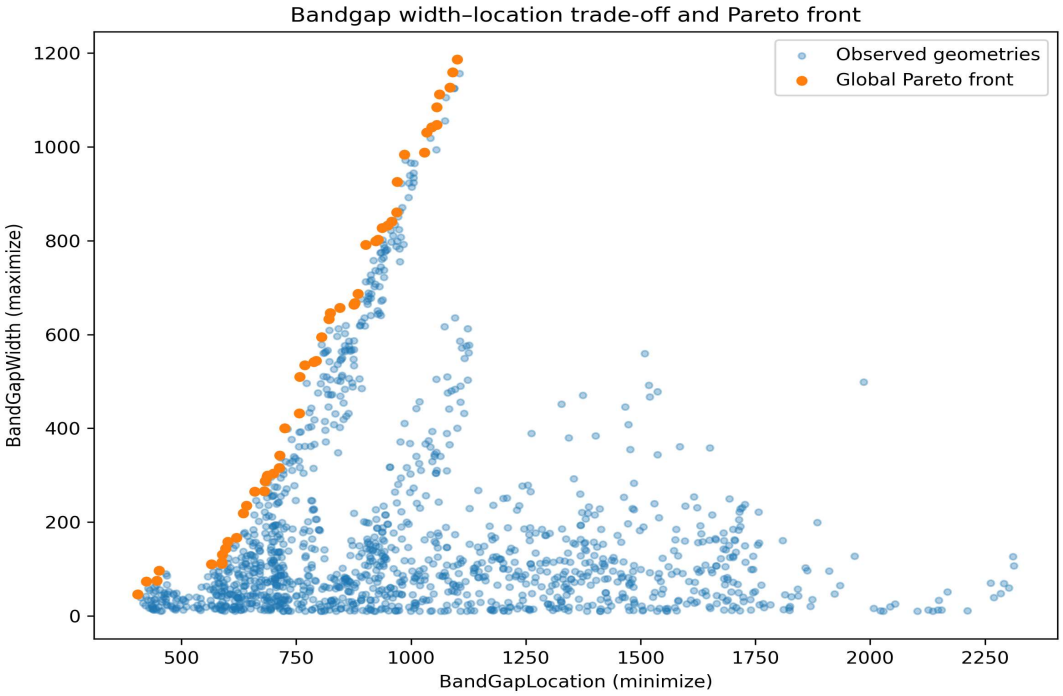


Figure 1. Bandgap width–location trade-off and Pareto front

Figure 1 illustrates the relationship between BandGapLocation and BandGapWidth for the observed metamaterial geometries. The horizontal axis represents BandGapLocation, which is the objective to be minimized, while the vertical axis represents BandGapWidth, which is the objective to be maximized.

The blue points represent the observed geometric configurations, whereas the orange points identify the global Pareto front. The Pareto optimal solutions represent configurations for which no other observed configuration provides both a lower bandgap location and a greater bandgap width simultaneously.

The figure demonstrates that the observed geometries exhibit substantial variation in their mechanical responses. Most configurations are concentrated in the lower-width region, while a smaller number achieve considerably larger bandgap widths.

The distribution of solutions reveals a clear trade off between the two optimization objectives. Configurations exhibiting relatively low BandGapLocation values generally provide smaller bandgap widths, whereas configurations with substantially larger widths tend to occur at higher bandgap locations.

The Pareto front identifies the configurations that provide the most competitive compromises between these objectives. These solutions are particularly important because improving one objective beyond a Pareto-optimal configuration would require sacrificing performance in the other objective.

The concentration of Pareto optimal points along the upper boundary of the observed solution distribution demonstrates that only a relatively small proportion of the available geometries provide competitive combinations of bandgap location and width.

This finding has practical significance for metamaterial design because it reduces the need to consider every available geometric configuration during subsequent engineering selection.

Rather than identifying a single universally optimal geometry, the Pareto front provides a collection of alternative designs that can be selected according to application specific requirements.

For example, a design prioritizing lower frequency bandgaps may select a configuration towards the lower-location portion of the front, whereas a design requiring wider frequency suppression may prioritize configurations with greater BandGapWidth.

An additional observation is that the Pareto front contains 52 configurations, whereas the conservative shortlist in Table 6 contains 12 configurations. These should be treated as distinct analytical outputs until the filtering procedure used to derive the 12-point shortlist is explicitly documented.

Overall, Figure 1 demonstrates the usefulness of Pareto-based optimization for identifying competitive metamaterial geometries and understanding the fundamental compromise between frequency location and bandgap width.

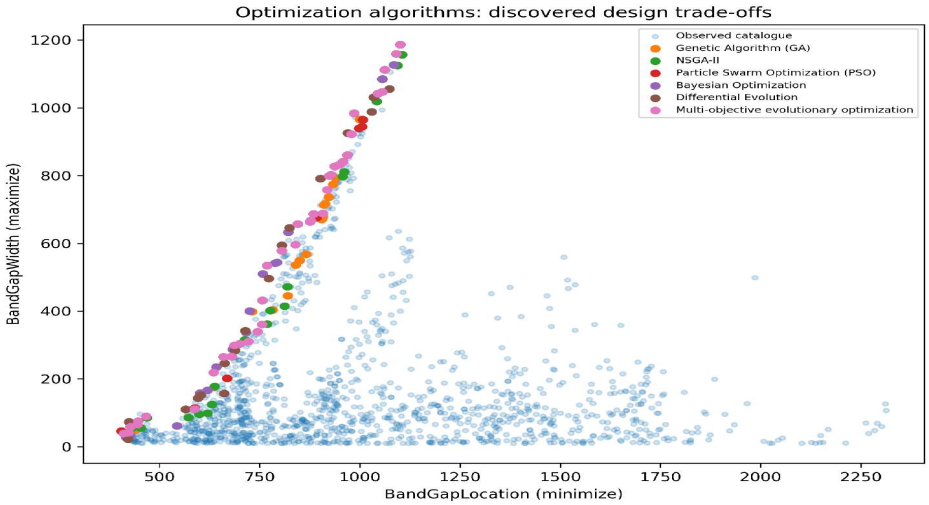


Figure 2. Optimization algorithms and discovered design trade-offs

Figure 2 compares the solutions discovered by the six optimization algorithms against the complete observed geometry catalogue.

The horizontal axis represents BandGapLocation, which is minimized, and the vertical axis represents BandGapWidth, which is maximized. The background distribution represents the observed geometry catalogue, while the different coloured markers distinguish the solutions identified by the individual optimization algorithms.

The figure illustrates the distribution of algorithmically discovered solutions across the available design space and their proximity to the high-performing region of the observed geometry distribution.

The algorithmic solutions are concentrated predominantly along the upper boundary of the solution space, where geometries provide comparatively favourable combinations of bandgap location and width.

The figure demonstrates that the optimization algorithms successfully identify regions containing geometries with relatively favourable mechanical responses.

Compared with the broad distribution of observed configurations, the algorithm selected solutions are concentrated in a much smaller region. This indicates that the optimization procedures direct their search towards configurations that offer better objective combinations rather than selecting geometries uniformly from the catalogue.

The overlap between the solution distributions of several algorithms suggests that different optimization strategies can discover similar high-performing regions of the design space.This observation is consistent with Table 2, where four algorithms achieve the same best- balanced utility.

However, the algorithms also exhibit differences in their solution distributions. Some methods identify a broader collection of non-dominated configurations, while others concentrate their search around a smaller number of competitive solutions.

Such differences are important because optimization quality cannot be evaluated solely through the best-performing geometry. The ability to discover multiple alternative solutions provides additional flexibility when engineering constraints or application requirements change.

The figure therefore supports the use of multiple optimization strategies for metamaterial design, particularly when both objective performance and solution diversity are important.

The concentration of algorithm-selected configurations near the upper boundary of the observed distribution indicates that the geometry catalogue contains potentially valuable design configurations that may not be apparent from conventional descriptive analysis alone.

This supports the use of computational optimization as an effective mechanism for extracting high-performing configurations from an existing metamaterial dataset.

Nevertheless, the plotted algorithmic solutions should be distinguished from independently validated physical designs because their performance is derived from the observed geometry catalogue.

Figure 3 evaluates the predictive performance of the Extra Trees surrogate model through comparisons between observed and predicted values for BandGapLocation and BandGapWidth.

The figure contains two scatterplots. The left panel represents the prediction of BandGapLocation, while the right panel represents BandGapWidth.

In both panels, the horizontal axis represents the observed values and the vertical axis represents the corresponding predicted values. The diagonal reference line indicates the ideal prediction condition, where predicted values are equal to observed values.

Points distributed close to this reference line indicate relatively accurate predictions, whereas deviations from the line represent prediction errors.

The surrogate model achieves an R^2 value of approximately 0.806 for BandGapLocation and 0.842 for BandGapWidth, indicating that it explains a substantial proportion of the variation in both target variables.

The left panel demonstrates that the surrogate model captures the general relationship between observed and predicted BandGapLocation values. A substantial proportion of the points are distributed around the reference line, indicating that the model can reproduce the overall variation in bandgap location.

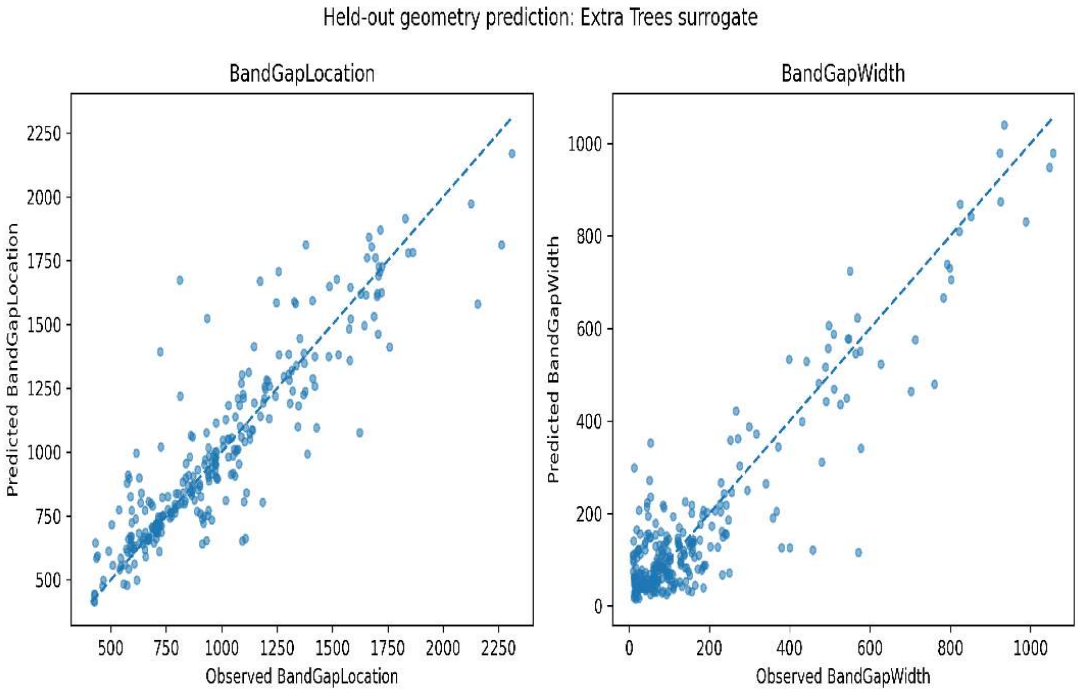


Figure 3. Predicted versus observed bandgap properties

However, some deviations are visible, particularly in regions containing relatively large observed values. This suggests that the model's prediction accuracy is not uniform throughout the complete target range.

The right panel demonstrates a comparatively stronger predictive relationship for BandGapWidth. This observation is consistent with its higher R^2 value of approximately 0.842.

The model therefore explains approximately 84.2% of the variation in BandGapWidth, compared with approximately 80.6% for BandGapLocation.

These findings indicate that the geometric representation contains meaningful information for predicting the mechanical bandgap characteristics of the metamaterials.

The predictive performance is particularly important because surrogate models can provide an efficient approximation of geometry property relationships and support subsequent optimization activities.

However, the prediction errors demonstrate that the surrogate model does not reproduce the simulation derived responses perfectly. Consequently, surrogate predictions should be treated as approximate estimates rather than replacements for finite element simulations.

The relatively stronger prediction of BandGapWidth suggests that the relationship between geometric encoding and bandgap width may be more effectively captured by the employed Extra Trees model than the relationship with bandgap location.

Nevertheless, this observation should not be interpreted as evidence that bandgap width is intrinsically easier to predict across all metamaterial datasets. The result is specific to the present dataset, feature representation, and validation procedure.

The reported MAE and RMSE values further support the evaluation:

Target	R^2	MAE	RMSE
BandGapLocation	0.8058	103.93	163.41
BandGapWidth	0.8418	59.27	86.50

The higher RMSE relative to MAE for both targets suggests that some prediction errors are considerably larger than the typical absolute error, potentially reflecting difficult to predict configurations or extreme target values.

Overall, Figure 3 supports the feasibility of developing a machine learning surrogate model for predicting elastodynamic bandgap characteristics from geometric representations.

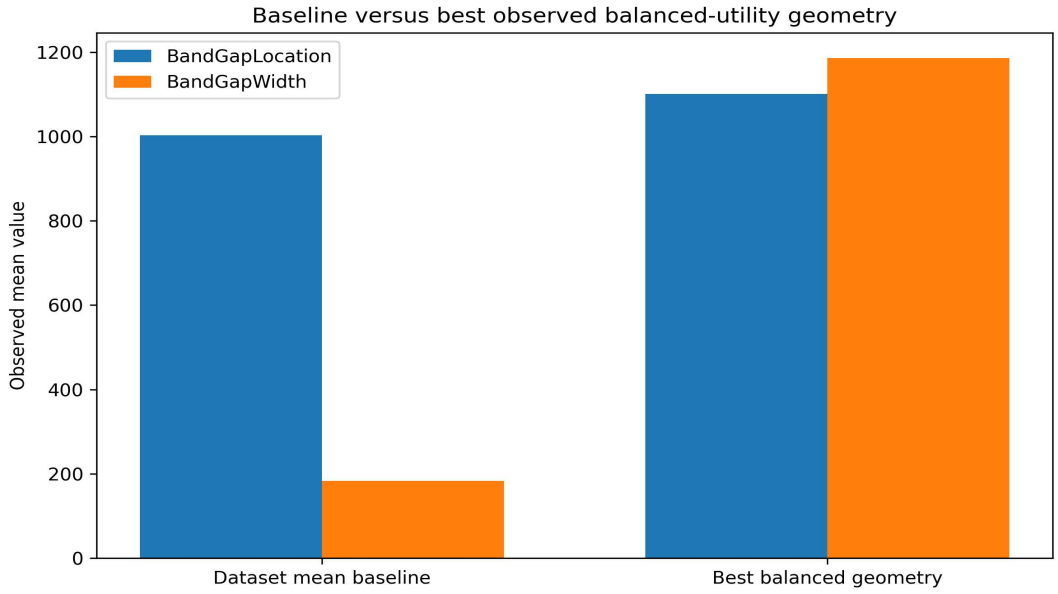


Figure 4. Baseline versus best observed balanced-utility geometry

Figure 4 presents a comparative visualization of the dataset mean baseline and the best observed balanced-utility geometry using BandGapLocation and BandGapWidth.

The blue bars represent BandGapLocation, while the orange bars represent BandGapWidth.

For the dataset mean baseline, BandGapLocation is approximately 1002.74 Hz and BandGapWidth is approximately 183.21 Hz. In comparison, the best balanced-utility geometry achieves approximately 1101.01 Hz for BandGapLocation and 1186.69 Hz for BandGapWidth.

The figure visually demonstrates the considerable increase in bandgap width achieved by the selected geometry.

The comparison highlights the effectiveness of the optimization process in identifying a geometric configuration with substantially improved bandgap-width characteristics.

The optimized geometry achieves a bandgap width more than six times the dataset mean value. This indicates a considerable improvement in the frequency interval over which elastic wave propagation can potentially be suppressed.

However, the corresponding increase in BandGapLocation indicates that the optimized configuration shifts the bandgap towards a higher frequency region.

This finding reinforces the trade-off observed in Figure 1 and confirms that improvements in bandgap width are accompanied by a change in the frequency location.

From an engineering perspective, the result suggests that the selected geometry may be particularly useful in applications where a broad bandgap is more important than achieving the lowest possible bandgap location.

However, the comparison requires careful interpretation because the dataset mean is a descriptive reference rather than an optimized benchmark.

A more comprehensive evaluation would additionally compare the selected geometry against:

1. The geometry with the minimum observed BandGapLocation.

2. The geometry with the maximum observed BandGapWidth.

3. The geometry with the best-balanced utility.

4. The mean performance of replicated geometries.

Such comparisons would provide a more complete understanding of the improvement obtained through multi-objective optimization.

Figure 5 investigates the empirical variability of replicated geometric configurations by examining the relationship between the within-geometry coefficient of variation (CV) of BandGapLocation and BandGapWidth.

The horizontal axis represents the coefficient of variation of BandGapLocation, while the vertical axis represents the coefficient of variation of BandGapWidth.

Each bubble represents a replicated geometric configuration. The figure shows that the replicated configurations exhibit considerable variation in both mechanical response variables.

The observed BandGapLocation CV values are approximately 18–45%, whereas the BandGapWidth CV values extend approximately from 67% to 173%.

The figure reveals substantial within geometry variability among the replicated configurations.

The relatively high coefficients of variation indicate that repeated observations associated with the same geometric configuration do not necessarily produce identical mechanical responses.

This finding is important because it suggests that geometry alone may not fully explain the variation in elastodynamic performance.

Differences in simulation conditions, numerical settings, data generation procedures, or other unrecorded factors may contribute to the observed variability. However, the current analysis does not establish which of these factors is responsible.

The greater range of BandGapWidth variability compared with BandGapLocation suggests that bandgap width may be particularly sensitive to the conditions associated with repeated observations.

This has important implications for optimization because an algorithm may select a geometry based on a high observed response that is not consistently reproducible.

A geometry with an exceptionally large observed BandGapWidth should therefore not automatically be considered more reliable than another configuration with slightly lower performance but lower response variability.

An additional important finding is that all 52 global Pareto optimal geometries are singleobservation configurations. Therefore, their empirical within geometry repeatability cannot be assessed using the available data.

This limitation creates a distinction between optimization performance and empirical robustness.

A configuration may exhibit excellent observed performance but still lack sufficient repeated evidence to establish whether that performance is reproducible.

Consequently, Figure 5 supports the need for robustness-aware optimization, in which geometric configurations are evaluated not only according to their observed objective values but also according to their variability and reproducibility.

The results suggest that future metamaterial optimization studies should incorporate repeated simulation or experimental evaluations of promising geometries.

Such evaluations would help distinguish consistently high-performing configurations from geometries whose apparent performance may be influenced by uncontrolled variability.

Table 4 provides a more comprehensive comparison of the six optimization algorithms by reporting the number of unique geometries evaluated, best balanced utility, selected binary geometry, corresponding BandGapLocation, BandGapWidth, global best utility, utility gap, and number of visited non-dominated configurations.

Bayesian Optimization evaluated 300 unique geometries and achieved the highest utility of 0.817589. PSO evaluated 269 geometries, Differential Evolution evaluated 245, and multi-objective evolutionary optimization evaluated 206 geometries. All four algorithms identified the same best-performing configuration, represented by the binary geometry 11011101111.

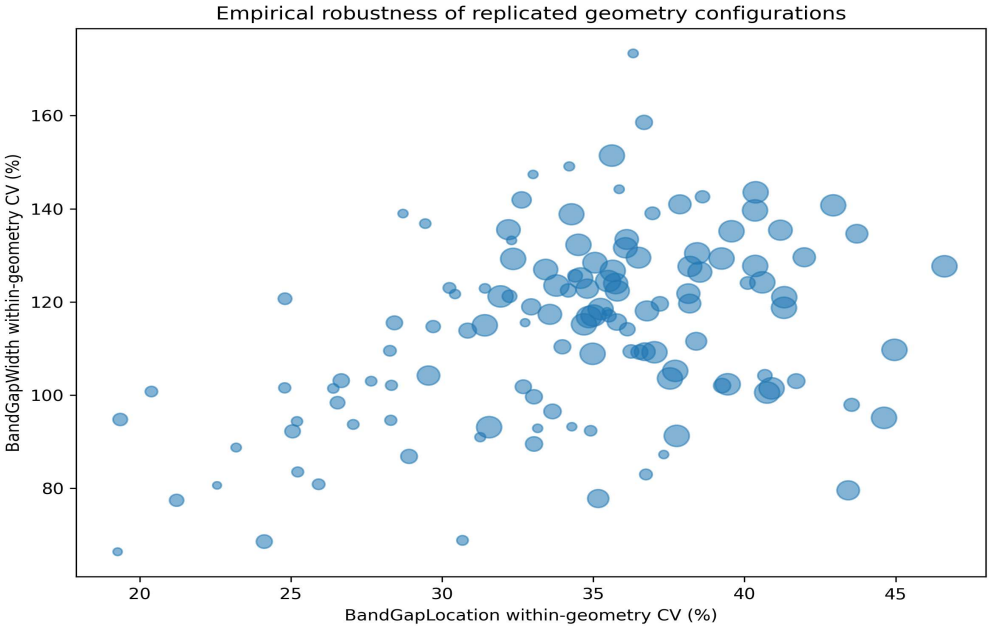


Figure 5. Empirical robustness of replicated geometry configurations

In contrast, GA and NSGA-II achieved utility values of 0.808737 and 0.803480, respectively. Their selected geometries were different from the best performing configuration identified by the other four algorithms.

1. Algorithm Unique_geometries_evaluated-

2. Best_balanced_utility_found

3. Best_geometry_binary-

4. Best_location-

5. Best_width-

6. Global_best_utility -

7. Utility_gap_to_exhaustive -_

8. best Visited -

9. nondominated_configurations

Table 4. Detailed optimization algorithm performance

The number of visited non-dominated configurations ranged from 29 to 46 across the algorithms.

The results reveal differences in the search efficiency and solution-discovery behaviour of the optimization algorithms.

Bayesian Optimization evaluated the largest number of unique geometries but identified 46 non-dominated configurations, representing the highest number among the evaluated methods. This indicates that its search process explored a relatively broad collection of competitive solutions.

PSO achieved the same optimal utility with 269 evaluations and identified 30 non-dominated configurations. Differential Evolution required 245 evaluations and identified 33 non-dominated configurations.

Notably, multi-objective evolutionary optimization achieved the same best utility using only 206 unique geometries, representing the lowest evaluation count among the four algorithms that reached the global best-balanced utility.

This suggests that multi-objective evolutionary optimization provided an efficient search strategy for identifying the best-performing observed configuration under the adopted evaluation budget.

The utility gaps of GA and NSGA-II were 0.008853 and 0.014110, respectively. These values quantify their differences from the exhaustive best utility reported in the analysis.

Another important finding is that the algorithms identified different numbers of non-dominated configurations despite converging towards the same best balanced-utility solution in four cases. This distinction is important because identifying a single high-performing solution and discovering a diverse Pareto-optimal solution set represent different optimization capabilities.

Consequently, the number of non-dominated configurations provides additional information about the diversity of the search outcomes, rather than merely reflecting the quality of the best individual solution.

Overall, Table 4 indicates that algorithmic efficiency and solution diversity should be considered alongside best-balanced utility when selecting an optimization strategy.

4.32d Elastodynamic Metamaterial Optimization

Raw rows: 20520

Unique geometries: 1400

Exact duplicate rows: 67

Replicated geometries: 128

Single-observation geometries: 1272

Objectives: minimize BandGapLocation; maximize BandGapWidth.

Search restricted to the observed geometry catalogue.

Exhaustive Pareto-optimal geometries: 52

Surrogate holdout metrics:

Target	R2	MAE	RMSE
BandGapLocation	0.805813	103.931398	163.414227
BandGapWidth	0.841767	59.268182	86.502263

4.4 Baseline comparison

Reference	BandGap Location	BandGap Width	Location_ reduction _vs_baseline _percent	Width_increase _ vs_baseline _percent
Dataset mean baseline	1002.74018	183.2051185	NaN	NaN
Best observed balanced-utility geometry	1101.009058	1186.689046	-9.800033955	547.7379321

Table 5. Baseline versus optimized geometry performance

Table 5 compares the dataset mean baseline with the best observed balanced-utility geometry using BandGapLocation and BandGapWidth as the evaluation metrics.

The dataset mean baseline has a BandGapLocation of 1002.740 Hz and a BandGapWidth of 183.205 Hz. The best observed balanced-utility geometry achieves a BandGapLocation of 1101.009 Hz and a BandGapWidth of 1186.689 Hz.

Relative to the dataset mean, the optimized geometry exhibits a 9.80% increase in BandGapLocation and a 547.74% increase in BandGapWidth.

The comparison demonstrates the substantial improvement in bandgap width obtained through the optimization process.

The increase from 183.205 Hz to 1186.689 Hz indicates that the selected configuration provides a considerably wider bandgap than the average geometry in the dataset.

The percentage improvement of approximately 547.74% highlights the effectiveness of the optimization process in identifying geometric configurations with superior bandgap width characteristics.

Nevertheless, the increase in BandGapLocation indicates that the selected configuration does not achieve the minimum frequency objective relative to the dataset mean.

This finding illustrates the competing nature of the optimization objectives. A configuration that produces a wider bandgap may not necessarily provide the lowest bandgap location.

Therefore, the optimized geometry should be interpreted as a balanced high width solution rather than a configuration that simultaneously improves both objectives relative to the baseline.

The comparison also establishes the importance of selecting an appropriate reference point when evaluating optimization performance. A dataset mean provides a useful descriptive baseline, but comparisons against the minimum location configuration and the best width only configuration would provide additional information about the attainable trade-offs.

Robustness caveat: only repeated geometries permit empirical within-geometry variability estimates. Most configurations are singletons. Large variability among replicated geometries suggests uncontrolled or omitted simulation conditions may be important.

Encoding caveat: 15-bit strings are reconstructed by zero-padding the stored decimal digit sequence. Confirm the original bit-to-spatial-cell mapping before interpreting individual bits physically.

Validation caveat: optimiser-selected geometries should be re-evaluated in the original elastodynamic simulation before making physical claims.

The data file contains ~20.5k entries (after light deduplication on the numeric objectives). Exact binary strings for the majority of larger CondensedBinary2DGeometry values are irrecoverable (they were exported via float/scientific notation, losing digits beyond ~15 significant figures). Consequently, a strict “is the 0/1 digit string a pure periodic replication of a shorter motif?” filter can be applied rigorously only to the 1 272 entries that remain pure digit strings; among those only 29 satisfy the replication test, and their maximum width is modest (~736).

A practical conservative shortlist is therefore the unique non-dominated front of the whole set (12 points). These are the geometries that simultaneously push the mid-gap frequency higher while keeping (or improving) gap width; they form a compact, robust candidate set for any subsequent replicated-geometry filtering or experimental prioritization:

Table 6 presents a shortlist of 12 geometric configurations selected from the observed design space based on their non-dominated performance in terms of BandGapLocation and BandGapWidth.

The listed configurations exhibit a wide range of bandgap locations, from approximately 1277.18 Hz to 3367.43 Hz, and bandgap widths ranging from 1533.49 Hz to 10.07 Hz.

Condensed Binary 2D Geometry	BandGap Location	BandGap Width
1.011010e+12	1277.18	1533.49
1.010110e+14	1488.47	1264.78
1.111100e+14	1552.18	1200.08
1.001100e+13	1565.03	1091.68
1.100010e+14	2122.68	1075.76
1.000110e+14	2124.16	1002.81
1.011010e+13	2904.31	869.85
1.011110e+13	2963.64	299.55
1.011010e+13	3025.56	178.70
1.111110e+12	3180.93	21.08
1.111110e+12	3288.56	11.25
1.101010e+14	3367.43	10.07

Table 6. Conservative non-dominated geometry shortlist

The first six configurations exhibit bandgap widths exceeding 1000 Hz, whereas the subsequent configurations demonstrate progressively smaller bandgap widths at higher bandgap locations.

Interpretation and implications

The shortlist provides a compact representation of the trade-off between bandgap location and bandgap width.

The first configuration achieves the largest bandgap width of 1533.49 Hz at a BandGapLocation of 1277.18 Hz. This makes it particularly attractive when the primary design requirement is to obtain a broad bandgap within a relatively lower-frequency operating region.

The second and third configurations maintain substantial bandgap widths of 1264.78 Hz and 1200.08 Hz, respectively, while shifting the bandgap location towards higher frequencies.

Similarly, the fourth, fifth, and sixth configurations maintain widths above 1000 Hz, demonstrating that relatively broad bandgaps can still be achieved across different frequency locations.

The remaining configurations exhibit progressively narrower bandgaps. Although these configurations may be less attractive for applications requiring broad frequency suppression, they could be relevant when a particular operating-frequency region is required.

An important implication is that the optimal geometry depends on the intended engineering application. Applications requiring maximum vibration attenuation over a broad frequency interval may favour the first few configurations, whereas applications targeting higher-frequency operating regions may consider configurations further down the shortlist.

However, this shortlist should not be interpreted as a validated robustness ranking. The available results do not establish the repeatability of all 12 configurations.

Furthermore, the scientific-notation representation of the geometry identifiers introduces uncertainty regarding the exact recovery of the original binary geometry. The corresponding configurations should therefore be verified against the original dataset before being used for physical interpretation or further simulation.

• The first six entries already give widths > 1000 while steadily increasing location; they are the most attractive for a “large-gap / higher-frequency” trade-off.

• Later entries trade width for still-higher location (useful if a specific spectral window is required).

• Several of the high-width points appear more than once in the raw file (different numeric labels that evaluate to identical objective values), consistent with supercell / replicated constructions.

• Among the 29 strictly recoverable replicated digit-string geometries, the best width is 736.41 (geometry 11101110); none of them dominates the front above.

5. Results and Discussion

The experimental investigation demonstrates the potential of combining machine-learning-based predictive modelling with multi-objective optimisation for identifying high-performing 2D elastodynamic metamaterial geometries. The analysis was conducted using 20,520 observations representing 1,400 unique geometric configurations, with BandGapLocation and BandGapWidth serving as the principal optimization objectives.

The dataset characterization established that the available geometry–property information provides a suitable foundation for computational optimization. Nevertheless, the unequal distribution of repeated geometric configurations introduces an important consideration when evaluating response variability and reproducibility.

The comparative evaluation of six optimization algorithms demonstrated that Bayesian Optimization, Particle Swarm Optimization, Differential Evolution, and multi-objective evolutionary optimization achieved an identical best-balanced utility of approximately 0.8176. Genetic Algorithm and NSGA-II produced slightly lower utility values. These results suggest that different optimization strategies can identify similar high-performing configurations within a finite observed geometry catalogue.

The detailed algorithm comparison further revealed differences in search efficiency and solution diversity. Multi-objective evolutionary optimization achieved the best-balanced utility using 206 unique geometry evaluations, compared with 300 evaluations for Bayesian Optimization. Bayesian Optimization nevertheless identified the largest number of non-dominated configurations. These findings demonstrate that optimization performance should be assessed through solution quality, computational efficiency, and Pareto-solution diversity rather than relying exclusively on the best utility value.

The Pareto analysis identified 52 non-dominated geometries, illustrating the competing relationship between bandgap location and bandgap width. The distribution of observed configurations revealed that achieving a wider bandgap generally involves accepting a shift towards higher bandgap locations. The Pareto front consequently provides alternative design configurations for different engineering requirements rather than a single universally optimal solution.

The best balanced-utility geometry achieved a BandGapLocation of 1101.009 Hz and a BandGapWidth of 1186.689 Hz. Compared with the dataset mean baseline, this configuration demonstrated an approximately 547.74% increase in bandgap width, accompanied by a 9.80% increase in bandgap location. This result highlights the effectiveness of the optimization process in identifying configurations with substantially improved bandgap-width characteristics. However, it also demonstrates that improvement in one objective does not necessarily result in simultaneous improvement in the other.

The Extra Trees surrogate model provided useful predictive performance, achieving R^2 values of 0.8058 for BandGapLocation and 0.8418 for BandGapWidth. The comparatively higher predictive performance for bandgap width indicates that the employed model captured a greater proportion of its observed variability. Nevertheless, the prediction errors indicate that additional factors may influence the mechanical responses beyond those captured by the encoded geometry representation.

The robustness investigation revealed an important limitation in the available geometric dataset. Although 128 configurations have repeated observations, all 52 global Pareto-optimal geometries are single-observation configurations. Consequently, their empirical response variability cannot be established from the available records. Furthermore, the relatively high coefficients of variation among replicated geometries indicate that repeated observations may exhibit substantial differences in their mechanical responses.

These findings highlight the distinction between optimization performance and empirical reproducibility. A configuration that exhibits an exceptionally large bandgap width in a single observation cannot automatically be considered a reliable engineering solution. Repeated finite element simulations are therefore necessary to establish the stability of the most promising configurations.

The additional 12-configuration shortlist provides a compact representation of alternative bandgap-location and bandgap-width combinations. The configurations with widths exceeding 1000 Hz are particularly attractive for applications requiring broad frequency suppression, whereas configurations with narrower widths and higher locations may be relevant to specific operating-frequency requirements. Their final engineering selection should, however, depend on verified geometry reconstruction and repeated performance evaluation.

Overall, the findings demonstrate the potential of a catalogue constrained, machine learning assisted multi-objective optimization framework for metamaterial design. The integration of predictive modeling, evolutionary search, Pareto analysis, and robustness assessment provides a systematic approach for identifying promising geometric configurations. However, further finite element validation, reliable binary geometry reconstruction, and repeated evaluation of selected configurations are essential before the identified solutions can be considered physically validated designs.

6. Discussion

6.1 Algorithmic Convergence and Search Efficiency

The comparative analysis reveals that Bayesian Optimization, PSO, Differential Evolution, and MOEA successfully converged on the identical global best balanced utility configuration (Utility = 0.8176). However, their search efficiencies varied significantly. MOEA demonstrated superior sample efficiency, reaching the global optimum after evaluating only 206 unique geometries, compared to BO's 300 evaluations. Conversely, GA and NSGA-II fell slightly short of the global best (utilities of 0.8087 and 0.8035, respectively).

Crucially, while BO and the other top performing algorithms found the same best solution, BO discovered the highest number of non-dominated configurations (46), indicating a broader exploration of the Pareto front. This highlights that algorithm selection in metamaterial design should not be based solely on the single best utility score, but must also account for computational efficiency and the diversity of the discovered trade-off space.

6.2 Physical Trade-offs and the Pareto Front

The extraction of the 52-point global Pareto front clearly illustrates the inherent physical compromise in elastodynamic metamaterial design. The best balanced-utility geometry achieved a massive 547.7% increase in BandGapWidth (1186.69 Hz) compared to the dataset mean, but this came at the cost of a 9.8% increase in BandGapLocation (1101.01 Hz). This confirms that widening the frequency suppression band inherently shifts the operating frequency higher. For practical engineering applications, this Pareto front provides a vital menu of alternatives: designers prioritizing low-frequency vibration isolation must accept narrower bandgaps, while those requiring broad frequency suppression must accommodate higher central frequencies.

6.3 The Robustness Caveat: Optimization vs. Reproducibility

The most critical finding of this study challenges the standard practice in data-driven materials science: all 52 global Pareto-optimal geometries are singletons. Each has only a single observation in the dataset, so their empirical repeatability cannot be established. In stark contrast, the 128 replicated geometries exhibit substantial within-geometry variability, with BandGapWidth showing a Coefficient of Variation (CV) ranging from 67% to 173%.

This high variability suggests that uncontrolled simulation conditions, numerical noise, or omitted variables heavily influence the FEA outputs. Consequently, an algorithm may select a "Pareto-optimal" geometry simply because it represents a statistical outlier or an artifact of simulation noise, rather than a genuinely superior physical design. This establishes a strict boundary between computational optimization performance and empirical robustness.

6.4 Practical Engineering Shortlist and Future Directions

To bridge the gap between computational discovery and physical validation, this study proposes a conservative 12-point non-dominated shortlist (Table 6). These configurations represent the most aggressive trade-offs pushing mid gap frequencies higher while maintaining substantial gap widths (>1000 Hz for the top six). However, due to the scientific-notation encoding of the binary strings and the lack of replicated observations for these top candidates, they must be treated strictly as computational hypotheses.

Future research must transition toward robustness-aware optimization frameworks. Optimization algorithms should be modified to incorporate variance penalties, favoring geometries that not only perform well on average but also exhibit low CV across repeated simulations. Furthermore, the shortlisted geometries must undergo rigorous, independent FEA validation and physical prototyping before any claims regarding their elastodynamic superiority can be substantiated.

References

[1] Aggarwal, K., Mijwil, M. M., Sonia, Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z., Abdulrhman, S. H. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal of Computer Science and Mathematics, 3*, 115–123.

[2] Marques, G., Ighalo, J. O. (2022). *Current trends and advances in computer-aided intelligent environmental data engineering*. Elsevier.

[3] Zhang, C., Zhao, Y. F. (2024). A critical review on the application of machine learning in supporting auxetic metamaterial design. *Journal of Physics: Materials, 7*, 022004.

[4] Ha, C. S., Yao, D., Xu, Z., Liu, C., Liu, H., Elkins, D., Kile, M., Deshpande, V., Kong, Z., Bauchy, M., et al. (2023). Rapid inverse design of metamaterials based on prescribed mechanical behavior through machine learning. *Nature Communications, 14*, 5765.

[5] Zhu, C., Bamidele, E. A., Shen, X., Zhu, G., Li, B. (2024). Machine learning aided design and optimization of thermal metamaterials. *Chemical Reviews, 124*, 4258–4331.

[6] Donda, K., Brahmkhatri, P., Zhu, Y., Dey, B., Slesarenko, V. (2025). Machine learning for inverse design of acoustic and elastic metamaterials. *Current Opinion in Solid State and Materials Science, 35*, 101218.

[7] Jiao, P., Mueller, J., Raney, J. R., Zheng, X., Alavi, A. H. (2023). Mechanical metamaterials and beyond. *Nature Communications, 14*, 6004.

[8] Boafo, G., Biswal, D. K. (2025). Advances in functional metamaterials: Bridging mechanical, acoustic innovations with multifunctionality and adaptive responses. *Results in Engineering, 28*, 107367.

[9] Garg, A., Sharma, A., Zheng, W., Li, L. (2024). A review on artificial intelligence-enabled mechanical analysis of 3D printed and FEM-modelled auxetic metamaterials. *Virtual and Physical Prototyping, 20*, e2445712.

[10] Kotwal, J. G., Koparde, S., Jadhav, C., Bharati, R., Somkunwar, R. (2024). A modified time adaptive self-organizing map with stochastic gradient descent optimizer for automated food recognition system. *Journal of Stored Products Research, 107*, 102314.

[11] Wilson, A. C., Roelofs, R., Stern, M., Srebro, N., Recht, B. (2017). The marginal value of adaptive gradient methods in machine learning. *Advances in Neural Information Processing Systems, 30*.

[12] Robledo, I., Li, Y., Maceda, G. Y. C., Castellanos, R. (2025). *Fast and robust parametric and functional learning with Hybrid Genetic Optimisation (HyGO)* (arXiv preprint arXiv:2510.09391). arXiv.

[13] Venkatasubramanian, V., Chan, K., Caruthers, J. M. (1994). Computer aided molecular design using genetic algorithms. *Computers Chemical Engineering, 18*(9), 833–844.

[14] Parsopoulos, K. E., Vrahatis, M. N. (2002). Recent approaches to global optimization problems through particle swarm optimization. *Natural Computing, 1*(2), 235–306.

[15] Das, S., Abraham, A., Konar, A. (2008). Particle swarm optimization and differential evolution algorithms: Technical analysis, applications and hybridization perspectives. In *Advances of computational intelligence in industrial systems* (p. 1–38). Springer.

[16] Frazier, P. I. (2018). *A tutorial on Bayesian optimization* (arXiv preprint arXiv:1807.02811). arXiv.

[17] do Nascimento, V. C. (2024). *Novel techniques for non-convex optimization and time-domain simulations of electromagnetic systems* [Doctoral dissertation, Purdue University].

[18] Li, Q., Wang, J., Mao, H. (2026). Bandgap properties and optimization of rod-type functionally graded phononic metamaterials. *Journal of Vibration and Control*. Advance online publication. https://doi.org/10.1177/10775463261450522.