Proceedings of CONF-CDS 2026 Symposium: Machine Learning and Neural Network Applications in Engineering
DOI: 10.54254/2755-2721/2026.AST36219

An Integrated Framework for Industrial Product Design, Simulation and Optimization Using Deep Surrogate Modeling and Multi-Objective Evolutionary Optimization

Siqi Dou

Beijing Jingcheng International School, Beijing, China
741491407@qq.com

Abstract. In response to the contradiction of multi-performance conflict and high cost of high-fidelity simulation evaluation in the integrated design of industrial products, this paper proposes a design-simulation-optimization integration framework that couples a deep surrogate model and multi-objective evolutionary optimization: a deep neural network learns the rapid mapping of parameterized designs to multi-physics performance to replace expensive simulation, and then NSGA-II searches for the Pareto front in the engineering-feasible space. On the public building energy-efficiency benchmark data, the coefficients of determination of the deep surrogate for heating and cooling loads reach 0.994 and 0.982 respectively, and the GPU evaluation speed is about 4.44 million times per second; the surrogate-driven optimization quantitatively portrays the trade-off front of energy consumption and daylighting, and identifies key design drivers such as relative compactness and window-to-wall ratio. The results show that the framework can support the integrated design of data-driven industrial products at a very low computing cost, providing an efficient paradigm for reproducible design decision-making.

Keywords: Industrial product design, Integrated design optimization, Deep surrogate model, Multi-objective evolutionary optimization, NSGA-II

1. Introduction

The design of industrial products is shifting from the traditional model of relying on experience and trial and error to a quantitative decision-making model with data and simulation as the core. The integrated design of modern industrial products (such as mechanical components, equipment structures and building enclosures) requires weighing structure, energy consumption, cost and comfort simultaneously at the early stage of the scheme, and the local optimization of any single objective may damage the overall performance of the system [1].

The key to achieving integrated design lies in the accurate evaluation of candidate schemes. High-fidelity simulations such as finite element analysis, computational fluid dynamics and energy consumption simulation give reliable results, but a single evaluation often takes minutes to hours, while a systematic search of the design space requires thousands of evaluations, making the computational cost unbearable [2, 3]. Surrogate models alleviate this contradiction by fitting the

DOI: 10.54254/2755-2721/2026.AST36219

mapping from design variables to performance on a small number of simulation samples, replacing expensive simulation with cheap millisecond-level approximations; deep neural networks further provide high-precision and differentiable surrogates [4, 5, 6].

Real industrial product design also faces an inherent conflict between objectives: improving one performance often sacrifices another, with no single solution optimal for all objectives simultaneously. Multi-objective evolutionary optimization can approach the Pareto front, which is composed of a set of mutually non-dominated compromise schemes, in one run, providing designers with a basis for trade-off. Among them, NSGA-II is widely used because of its elitism and crowding-distance sorting mechanism [7]. However, most studies examine surrogate modeling and multi-objective optimization separately, or focus on single-link algorithm improvement, lacking a closed loop of "design-simulation-optimization" integrated for industrial products that can be reproduced on public real data.

In response to the above shortcomings, this paper proposes and verifies an integrated design-simulation-optimization framework for industrial products, the main contributions of which are as follows. First, we build an integrated closed loop with a deep surrogate model as the core and coupled with multi-objective evolutionary optimization, realizing automated decision-making support from parametric design to the Pareto front. Second, we quantitatively verify the accuracy and computational efficiency of the deep surrogate relative to traditional methods on a public real engineering data set. Third, the surrogate-driven NSGA-II reveals the performance trade-off front and key design drivers of a typical product, providing a reproducible paradigm for the integrated design of data-driven industrial products.

2. Related work

Surrogate modeling has been accumulated for a long time in engineering design optimization. Early methods are represented by the response surface method, Kriging and radial basis functions, which fit the performance response of the low-dimensional design space through polynomials or kernel functions, and have been systematically summarized in aerospace and other fields [1, 3]. With the rise of problem dimensions and nonlinearity, deep neural networks have gradually become a powerful tool for surrogate modeling: their differentiable response surface is easy to couple with optimizers and can achieve large-scale design evaluation through batch inference [8]. On this basis, physics-informed machine learning that embeds physical laws into the network has further improved the generalization ability under small-sample conditions [9].

In terms of multi-objective optimization, NSGA-II and its variants have become standard tools for engineering design [7]. In order to reduce the evaluation burden of expensive simulation, surrogate-assisted evolutionary algorithms embed cheap surrogates into the evolutionary search loop, and only call real simulation when necessary [4]. In recent years, related methods have been systematically reviewed [10], and decomposition-based multi-objective algorithms with deep models such as convolutional neural networks as surrogates have emerged [11].

At the level of specific product performance prediction, machine learning has been widely used in building and equipment energy consumption modeling. Taking public building energy-efficiency data as an example, ensemble learning methods achieve high accuracy in heating and cooling load prediction [12], and automated machine learning further reduces the labor cost of modeling [13]. In the field of equipment manufacturing, surrogate-assisted multi-objective optimization is used in the design of complex industrial products such as electrical machines, showing significant efficiency advantages [14]. In general, existing work advances surrogate accuracy or optimization algorithms

DOI: 10.54254/2755-2721/2026.AST36219

in isolation, but rarely integrates parametric design, deep surrogates and multi-objective optimization under a unified, reproducible framework, which is the gap this paper addresses.

3. Methodology

3.1. Overview of the framework

The integrated design-simulation-optimization framework proposed in this paper includes three coupling links (Figure 1). First, the adjustable design elements of industrial products are encoded into parametric design vectors $\boldsymbol{x} = (x_1, \ldots, x_d)$, whose values are constrained by the engineering-feasible domain. Second, a deep surrogate model f_θ establishes a rapid mapping of design vectors to multi-physics performance, replacing expensive high-fidelity simulation. Finally, multi-objective evolutionary optimization searches the design space to obtain a set of mutually non-dominated compromise schemes for decision-makers, closing the loop from design input to optimization decision.

3.2. Deep surrogate model

The surrogate model is implemented by a multilayer perceptron, whose input is a d -dimensional standardized design vector. After three hidden layers (with 64, 64 and 32 neurons respectively, and ReLU as the activation function), multiple performance predictions are output as $\widehat{\boldsymbol{y}} = f_\theta\left(\boldsymbol{x}\right)$. The model parameters θ are determined by minimizing the mean square error on the training samples:

$$L\left(\theta\right) = \frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{y}_i - f_\theta(\boldsymbol{x}_i)\|_2^2 \tag{1}$$

where N is the number of samples and $\boldsymbol{y}_i$ is the true performance of the i -th sample. The training uses the Adam optimizer together with an early-stopping strategy to suppress overfitting. The surrogate accuracy is evaluated on held-out data by the coefficient of determination R^2, the root mean square error RMSE and the mean absolute error MAE:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \qquad \mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_i (y_i - \hat{y}_i)^2} \tag{2}$$

3.3. Multi-objective evolutionary optimization

After obtaining a reliable surrogate, the integrated design is expressed as a multi-objective optimization problem:

$$\min_{\boldsymbol{x}\in\Omega} \boldsymbol{F}\left(\boldsymbol{x}\right) = [f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x})] \tag{3}$$

where Ω is the engineering-feasible design space, and f_1 and f_2 are conflicting performance objectives whose values are given by the surrogate model. For design schemes $\boldsymbol{x}^a$ and $\boldsymbol{x}^b$, if $\boldsymbol{x}^a$ is not inferior to $\boldsymbol{x}^b$ on all objectives and is better on at least one objective, then $\boldsymbol{x}^a$ dominates $\boldsymbol{x}^b$; the solutions that are not dominated by any scheme constitute the Pareto-optimal front. This paper adopts NSGA-II for solution, which maintains the diversity of the solution set while preserving convergence through non-dominated sorting and crowding distance [7]. To avoid generating

DOI: 10.54254/2755-2721/2026.AST36219

geometrically infeasible designs, the adjustable variables are constrained within the discrete base geometries and continuous parameter ranges covered by the real samples. The optimization quality is measured by the hypervolume (HV) indicator; a larger HV indicates that the front is closer to the ideal and more widely distributed.

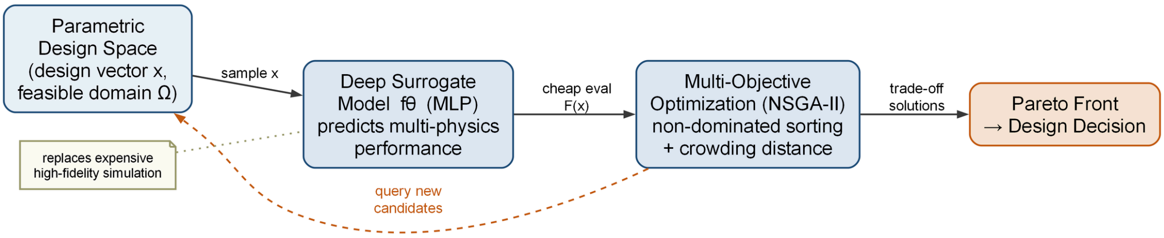

Figure 1. The integrated design-simulation-optimization framework for industrial products

4. Experiments and results

4.1. Data and experimental settings

The experiment adopts the public building energy-efficiency benchmark data set (Energy Efficiency, ENB2012) [15]. The data set is generated by physical simulation and contains 768 samples; the input consists of 8 parametric design variables describing the enclosure structure (relative compactness, surface area, wall area, roof area, overall height, orientation, window-to-wall ratio and glazing area distribution), and the output is the two interrelated energy-consumption performances of heating load Y_1 and cooling load Y_2. As a typical industrial product that can be manufactured in bulk, the mapping structure from design variables to multi-physics performance of building enclosures is highly compatible with this framework. All experiments use a fixed random seed, the surrogate accuracy is evaluated by 5-fold cross-validation, and the optimization is completed on a graphics processing unit (NVIDIA RTX 5080).

4.2. Surrogate model accuracy

Table 1 compares the prediction accuracy of the deep surrogate and four baseline models on the two objectives. The deep surrogate achieves $R^2 = 0.994$ (RMSE of 0.76) on the heating load and $R^2 = 0.982$ (RMSE of 1.28) on the cooling load, which is significantly better than linear regression (R^2 of 0.914 and 0.885 respectively) and support vector regression, and is comparable to strong baselines such as gradient boosting and random forests. More importantly, the deep surrogate provides a smooth and differentiable response surface and can reach about 4.44 million design evaluations per second on the graphics processing unit, providing key computational-efficiency support for subsequent large-scale optimization search.

4.3. Key design drivers

Sensitivity analysis based on permutation importance shows that energy-consumption performance is mainly dominated by a few design variables: relative compactness contributes about 50%, followed by the window-to-wall ratio (about 17%) and the overall height (about 16%), which together explain more than 80% of the performance variation. This result is consistent with the laws of building physics, and also provides a basis for designers to focus on key variables at the early stage of the scheme.

DOI: 10.54254/2755-2721/2026.AST36219

4.4. Multi-objective optimization and trade-off front

Reflecting the performance-comfort trade-off common in industrial product design, this paper sets minimization of total energy load (heating plus cooling) against maximization of daylighting potential (characterized by the window-to-wall ratio) as conflicting objectives. The surrogate-driven NSGA-II completes 30,000 evaluations in about 1 second and obtains a Pareto front composed of 120 mutually non-dominated schemes (Figure 2). The front clearly portrays the trade-off relationship: as the window-to-wall ratio increases from 0 to 40%, the achievable total energy consumption increases from 16.8 to 29.2 kWh/m²; in other words, raising the daylighting demand to the maximum will increase the achievable minimum energy consumption by about 74%. The minimum-energy scheme on the front reduces the total load by about 64% compared with the average design of the samples, while the "knee-point" scheme takes into account both daylighting and energy efficiency at a window-to-wall ratio of about 23%. Compared with random search using the same number of evaluations, the hypervolume achieved by NSGA-II is comparable, indicating that the surrogate-driven search has stably converged to the global trade-off front; in a higher-dimensional and larger-scale real product design space, the efficiency advantage of the evolutionary algorithm is expected to be more prominent [10, 14].

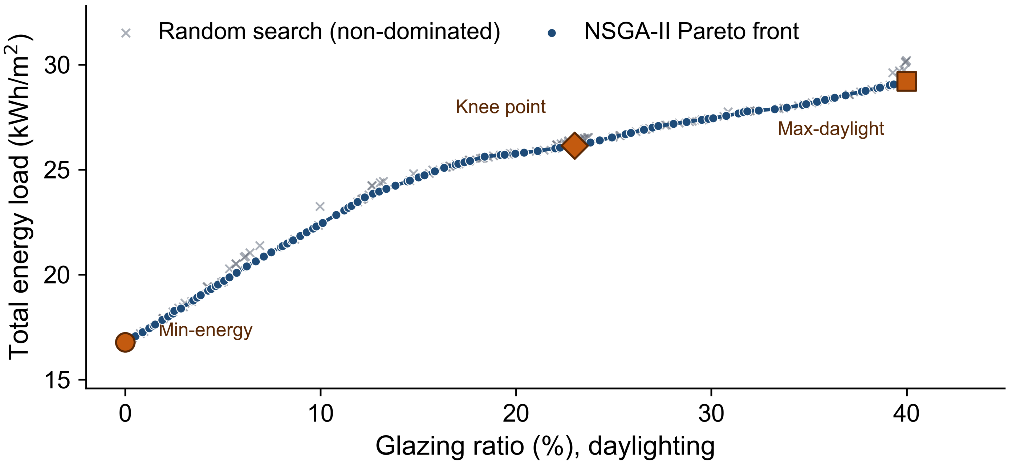

Figure 2. Pareto front of the energy-daylighting trade-off obtained by surrogate-driven NSGA-II, versus random search of equal budget

Table 1. Surrogate accuracy on the ENB2012 benchmark (5-fold cross-validation means)

Model	R^2 (heating)	RMSE (heating)	R^2 (cooling)	RMSE (cooling)
MLP (proposed)	0.994	0.76	0.982	1.28
Linear	0.914	2.94	0.885	3.21
SVR (RBF)	0.941	2.43	0.918	2.72
Random Forest	0.998	0.48	0.961	1.86
Grad. Boosting	0.998	0.47	0.974	1.52

R^2: coefficient of determination (higher is better); RMSE in kWh/m² (lower is better).

5. Discussion

The results show that connecting parametric design, deep surrogates and multi-objective optimization quantifies multi-performance trade-offs and identifies key drivers at very low cost in the early design stage, supporting data-driven integrated design of industrial products. The deep surrogate is the core enabler of this closed loop: its sub-microsecond evaluation speed makes almost exhaustive design exploration possible, which is difficult to achieve when relying on one-by-one high-fidelity simulation; at the same time, its differentiable, batch-parallel and joint multi-performance output characteristics make it naturally adaptable to evolutionary optimization and subsequent gradient-based optimization [6, 8].

There are still some limitations in this study. First, the daylighting potential is approximated by the window-to-wall ratio, and has not incorporated real optical simulation such as glare and illuminance distribution, which is a simplified characterization of the comfort objective. Second, the design space of the benchmark used in this paper is relatively compact, and random search is already competitive, so the scalability advantage of the evolutionary algorithm needs to be verified on larger-scale problems. Third, there is uncertainty in the extrapolation of the surrogate in the sparse regions of the training samples, and embedding physical constraints into the network is expected to alleviate this problem [9]. Fourth, the verification is based on a single product case, and the generalization of the framework to other product domains such as machinery and equipment needs to be further tested.

6. Conclusion

For the integrated design needs of industrial products, this paper proposes and verifies a design-simulation-optimization integration framework that couples a deep surrogate model with multi-objective evolutionary optimization. On public real engineering data, the deep surrogate replaces expensive simulation with a coefficient of determination of about 0.99 and an evaluation speed of millions of times per second; the surrogate-driven NSGA-II reveals the Pareto trade-off front between energy consumption and daylighting, and locates key design drivers such as relative compactness, providing a practical and reproducible paradigm for data-driven integrated design of industrial products. Future work will extend to high-dimensional product design spaces, embed physical and manufacturing constraints, and deeply integrate with real optical, structural and other multi-physics simulation as well as multidisciplinary optimization.

Declaration of interest

The author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contribution

Siqi Dou: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization.

References

[1] Wang, G. G., & Shan, S. (2007). Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical Design, 129*(4), 370–380. https://doi.org/10.1115/1.2429697

DOI: 10.54254/2755-2721/2026.AST36219

[2] Forrester, A. I. J., & Keane, A. J. (2009). Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences, 45*(1–3), 50–79. https://doi.org/10.1016/j.paerosci.2008.11.001

[3] Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Tucker, P. K. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences, 41*(1), 1–28. https://doi.org/10.1016/j.paerosci.2005.02.001

[4] Jin, Y. (2011). Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation, 1*(2), 61–70. https://doi.org/10.1016/j.swevo.2011.05.001

[5] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

[6] Regenwetter, L., Nobari, A. H., & Ahmed, F. (2022). Deep generative models in engineering design: A review. *Journal of Mechanical Design, 144*(7), 071704. https://doi.org/10.1115/1.4053859

[7] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation, 6*(2), 182-197. https://doi.org/10.1109/4235.996017

[8] Woldseth, R. V., Aage, N., Bærentzen, J. A., & Sigmund, O. (2022). On the use of artificial neural networks in topology optimisation. *Structural and Multidisciplinary Optimization, 65*(10), 294. https://doi.org/10.1007/s00158-022-03347-1

[9] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics, 3*(6), 422-440. https://doi.org/10.1038/s42254-021-00314-5

[10] Liu, S., Wang, H., Peng, W., & Wen, Y. (2024). Surrogate-assisted evolutionary algorithms for expensive combinatorial optimization: A survey. *Complex & Intelligent Systems, 10*(4), 5933-5949. https://doi.org/10.1007/s40747-024-01465-5

[11] Zhang, T., Li, F., Zhao, X., Qi, W., & Liu, T. (2022). A convolutional neural network-based surrogate model for multi-objective optimization evolutionary algorithm based on decomposition. *Swarm and Evolutionary Computation, 72*, 101081. https://doi.org/10.1016/j.swevo.2022.101081

[12] Chaganti, R., Rustam, F., Daghriri, T., de la Torre Díez, I., Vidal Mazón, J. L., Rodríguez, C. L., & Ashraf, I. (2022). Building heating and cooling load prediction using ensemble machine learning model. *Sensors, 22*(19), 7692. https://doi.org/10.3390/s22197692

[13] Zhang, C., Tian, X., Zhao, Y., & Lu, J. (2023). Automated machine learning-based building energy load prediction method. *Journal of Building Engineering, 80*, 108071. https://doi.org/10.1016/j.jobe.2023.108071

[14] Liu, L., Li, Z., Kang, H., Xiao, Y., Sun, L., Zhao, H., Zhu, Z. Q., & Ma, Y. (2025). Review of surrogate model assisted multi-objective design optimization of electrical machines: New opportunities and challenges. *Renewable and Sustainable Energy Reviews, 215*, 115609. https://doi.org/10.1016/j.rser.2025.115609

[15] Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings, 49*, 560–567. https://doi.org/10.1016/j.enbuild.2012.03.003