2026 IEEE/SICE International Symposium on System Integration (SII)
January 11-14, 2026. Cancún, México

Application of Sequential Approximation Optimization with reduced simulation numbers for integrated system optimization

Kaito Toyoshima, Yoshiharu Iwata, Hidefumi Wakamatsu

Abstract— **Simulation plays a crucial role in optimizing complex system designs. While individual subsystems may have relatively short processing times, when the entire system is integrated, the significant time and computational resources required remain a considerable challenge. On the other hand, system integration introduces effects that cannot be easily formalized, making integrated simulation indispensable. To address this challenge, this study proposes a methodology to enhance accuracy with limited data by augmenting the training dataset utilizing the characteristics of the Integration Neural Networks (INN) within the framework of Sequential Approximation Optimization (SAO). The INN is a surrogate model that integrates a deductive neural network, which handles scientifically formulable phenomena, and an inductive neural network, which addresses non-formulable phenomena. In this context, the expressive constraints of the deductive neural network serve to restrict variations between training points. Consequently, the surrogate model can attain high accuracy even with a limited quantity of training data. By leveraging this attribute, the proposed method achieves improved solution accuracy in SAO with minimal data points by adding training points surrounding the predicted optimal solution. In conventional SAO, the initial training dataset starts with 10 times the number of design variables. In contrast, this study hypothesized that optimization with fewer simulation runs could be achieved by strategically replacing random, low-quality initial training data with data selected based on intermediate learning results. Specifically, we investigated whether the total amount of training data required could be reduced by decreasing the number of initial training data points and increasing the number of data points added during the optimization process. The results of the experiments showed that stable solutions within the error range were obtained for all cases. Furthermore, it was demonstrated that a balance between data reduction and optimization stability could be successfully achieved by setting the number of initial training samples to three times the number of design variables (21 samples). As a result, the number of simulation runs required to construct the training dataset was reduced by approximately 50.2% compared to the conventional SAO method, demonstrating significant potential for shortening the design cycle in product optimization.**

I. INTRODUCTION

Systems composed of multiple integrated subsystems are commonly found across various fields of modern engineering. When evaluating such complex systems, it is not sufficient to assess each subsystem's performance individually; a comprehensive evaluation of the overall system behavior, which emerges from the interaction among subsystems, is essential. However, as more elements are combined, the overall system structure tends to become increasingly complex. Consequently, simple formulations that could be applied within individual subsystems often become ineffective due to the influence of boundary conditions between subsystems. As a result, verifying the performance of such sophisticated and highly accurate product systems requires extensive time and resources for prototyping and simulation. On the other hand, the time available for design is limited due to time-to-market constraints, which often forces designers to rely on their intuition and experience. This reliance can lead to situations where the final design solution cannot be considered truly optimal [1]-[3]. Therefore, there is a pressing need for effective methods to resolve the trade-off between time constraints and solution accuracy. Generally, the optimization process proceeds through repeated cycles of design variable updates and performance evaluations. In the context of product design, however, the evaluation phase consumes a substantial portion of the total time. To reduce the time required for optimization, two possible approaches are: decreasing the number of iterations or shortening the time needed for each evaluation. In practice, achieving both simultaneously is quite challenging. Various optimization methods have been proposed to reduce the number of iterations. Although optimization methods with fast convergence have been explored, techniques that can handle multimodal problems—such as Genetic Algorithms—often incur high computational costs [4]. Meanwhile, surrogate models using machine learning have recently gained attention as a promising approach to reduce evaluation time [5][6]. These models can deliver results in a significantly shorter time compared to actual simulations, thereby enabling substantial reduction in evaluation time during the design process. However, to function as sufficiently accurate surrogate models, general machine learning models typically require a large volume of training data (i.e., simulation results). In practical design practice, the time required for generating such data through simulation tends to dominate the overall design schedule. Reducing the amount of training data would mitigate this issue, but it also risks degrading the surrogate model's prediction accuracy, potentially leading to suboptimal or invalid design solutions.

Therefore, as a method for reducing the amount of training data required for constructing a surrogate model, active learning has been proposed. Active learning sequentially adds training data to regions in which the accuracy of the approximator is low, and it enables the surrogate model to maintain its accuracy even with a small number of data. However, if the objective is optimization, high accuracy of the surrogate model over the entire design space is not necessary, because what is required in optimization is the optimal

Kaito Toyoshima, Yoshiharu Iwata, and Hidefumi Wakamatsu are with Dept. of Materials and Manufacturing Science, Graduate School of Eng., The University of Osaka, Yamadaoka 2-1, Suita, Osaka 565-0871, Japan (e-mail: {kaito.toyoshima, wakamatu, iwata} @mapse.eng.osaka-u.ac.jp).

solution and its objective value. Therefore, in this study, we focus on Sequential Approximate Optimization (SAO) [7]–[9], a method that enhances the accuracy of the surrogate model specifically in the neighborhood of the derived optimal solution by adding training data around that solution, thereby enabling the optimization to be completed with a small amount of data. Furthermore, we employ an SAO technique that adopts Integration Neural Networks (INN) [10][11] as the surrogate model. INN combines a physics-inspired neural network with a conventional deep-learning network, and by leveraging the constraints imposed by physical equations, it enables the construction of a highly accurate surrogate model even when only a limited number of training data are available.

It has been reported that SAO using INN can reach the optimal solution with fewer training samples than conventional Bayesian optimization methods [12]. Nevertheless, in the context of practical application to real-world design problems, further reduction in the number of training samples is still needed. Previous research on SAO using INN [10][11] showed that the prediction error between the surrogate model and the simulation results for the optimal design solution significantly decreased after just two data additions, suggesting that the additional learning data, which took advantage of the INN's features, was of high quality in the training data. In conventional implementations, SAO uses an initial training dataset with 10 times the number of design variables. However, because SAO builds a surrogate model through the process of iteratively adding high-quality training data based on past learning results, we hypothesize that it is not necessary to prepare an initial dataset as large as 10 times the number of design variables. Therefore, this paper aims to investigate whether the total amount of training data required in SAO can be reduced by decreasing the number of initial training data samples.

II. Power Devices as a Case of System Integration

In this study, a power device is selected as the target system because its overall performance cannot be accurately evaluated through a simple additive assessment of individual components. It is well known that when the thermal resistance of each component is simply summed to construct a model, a significant error is introduced. Even when employing a 45-degree heat-spreading model that partially accounts for the thermal interface between components, the resulting prediction still exhibits an error of approximately 20%. Therefore, it has been recognized that appropriate evaluation of such a system requires the use of detailed simulations or prototype-based validation.

The power device considered in this study is a system in which multiple heterogeneous materials and components are integrated into a multilayered structure, as illustrated in Fig. 1. In this paper, the model used in the previous study [12] is adopted as the target model, and its CAD representation is shown in Fig. 2. Below, we will extract only the minimum information necessary from the previous research to proceed with the subsequent analysis and provide an overview.

An automotive power device is a component that performs switching operations to appropriately control the power flow to the engine, motor, and battery. At the same time, it is a major

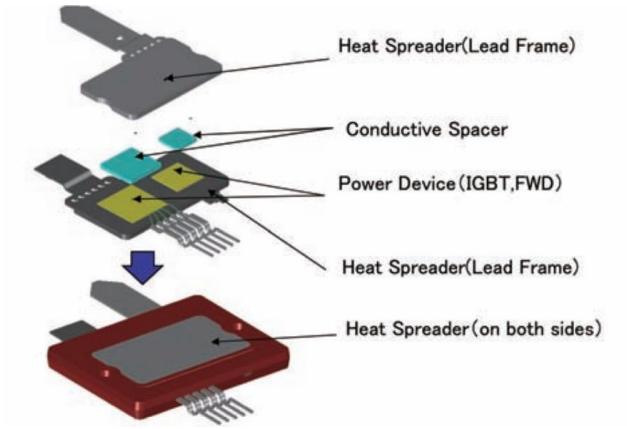

Figure 1. Image of power device [13].

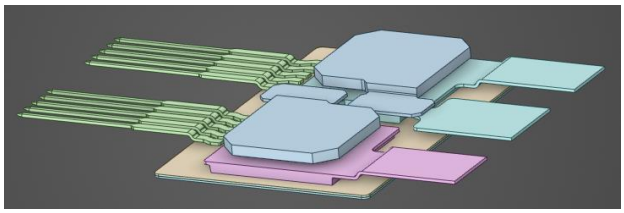

Figure 2. Simplified model of the power device used in the simulation.

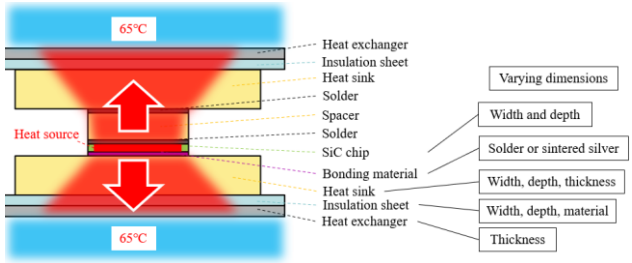

Figure 3. Power device stack structure and heat dissipation path.

TABLE 1. Information of optimization problem.

Objective variable	Total weight(g)
Constraint	Max temperature < 175°C
Design variables	SiC chip size : 5.0 to 12.0 (mm) Bonding material : Sn-Pb or Ag Heat sink size : 20.0 to 50.0 (mm) Heat sink thickness : 2.0 to 5.0 (mm) Insulation sheet thermal size : 1.3 to 6.5 (mm) Insulation sheet thermal conductivity : 15 or 18 or 20 (W/mK) Heat exchanger thickness : 1.0 to 3.0 (mm)

heat-generating component due to the large currents that flow through it. There are two main types of automotive power devices: single-sided cooling and double-sided cooling. In this study, we deal with the double-sided cooling type.

As shown in Fig. 3, the automotive power device is composed of multiple components—such as a heat exchanger, insulating sheet, heat sink, SiC elements, bonding material,

and spacers—that operate in close functional and structural coordination. Each power device contains two heat-generating SiC elements, with one embedded within each of the laminated structures. Additionally, due to the layered arrangement of components with different widths and depths, heat spreads within the structure. However, the commonly assumed 45-degree heat spreading angle does not always occur, since the angle of heat diffusion is affected by changes in thermal conductivity. Heat flows in the direction of the arrows, with 65°C cooling water supplied from the heat exchangers on both sides.

In the prior study [10][11], a mathematical formulation of the maximum temperature under steady-state conditions after current is applied to the target power device was developed, and the system was modeled using an INN. It was demonstrated that, by using training data based on a three-level factorial design, the model could accurately predict the thermal behavior of the power device.

III. Identification of Issues in SAO Using INN

A. Optimization Problem

In this paper, we use a simplified optimization problem aimed at reducing the weight of an automotive power device as a case study. The objective variable is the total mass of the power device, and the goal is to minimize this mass. The design variables consist of five continuous component dimensions and two discrete material selections, making a total of seven design variables. The constraint condition requires that the maximum temperature under steady-state conditions—after current is applied to the power device—must remain below a specified threshold of 175°C. Details of the optimization problem are summarized in Table 1. The range of each design variable is based on actual product designs, with variations introduced around their nominal values, and includes candidate materials that were under consideration.

B. Prior Findings and the Motivation for Reducing Initial Training Data

In the previous study [12], results such as those shown in Fig. 4 were obtained. Fig. 4 illustrates the relationship between the total number of training data and the error between the surrogate model's predicted value and the simulation value at the optimal solution. It can be seen that the error is rapidly reduced by the first two additions of training data, and that by improving the selection of additional training data, the convergence to the optimal solution is shortened. From this result, it can be considered that the data addition process in SAO consists of a phase that performs rough optimization by exploring the approximate region of the optimal solution, and a phase that improves the accuracy of the optimal solution. In other words, the previous study [12] examined the shortening of the accuracy improvement phase. On the other hand, the reduction of training data, taking advantage of the short rough optimization phase, has not been examined. Therefore, this study focuses on reducing the initial training data and optimizing the trade-off with the increase in additional training data required during rough optimization.

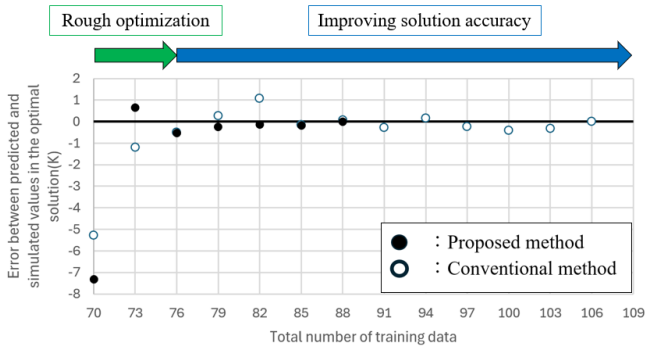

Figure 4. Separation of the optimal solution search process and search stages using SAO [12].

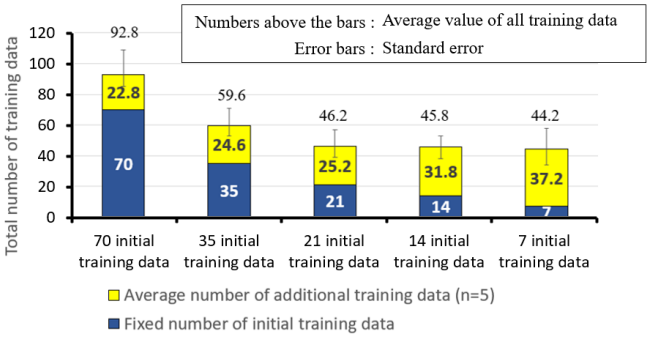

Figure 5. Results of the total number of training data for each trial of SAO from 7 to 70 initial training data.

IV. Investigation of Reducing Total Training Data through Reduction of Initial Training Data in SAO

A. Experimental Conditions

In this experiment, we employ the same case study as used in the previous research. Regarding the reduction of initial training data, we consider several ratios relative to the number of design variables: five times (half of the conventional tenfold rule), three times, twice, and finally the same number as the design variables. Specifically, SAO is performed using initial training datasets consisting of 70, 35, 21, 14, and 7 samples, respectively. The initial training data are generated using Latin Hypercube Sampling.

When the number of training data is reduced, there is a risk that the results may exhibit greater variance. This variability can be attributed to random elements inherent in both the SAO process and the surrogate model construction. While such variance has minimal impact when a sufficient number of training samples is used—allowing single-run evaluations to suffice—it may become significant when the amount of initial training data is substantially reduced. To appropriately evaluate this effect, we perform multiple SAO runs for each initial training data size, generating new initial datasets each time with different random number sequences. This allows us to account for potential biases in the training data and differences in the initialization of model parameters due to random weight generation.

Accordingly, for each initial training data size, SAO is executed five times, and the results are evaluated based on the averaged performance.

B. Experimental Results

Fig. 5 shows the average total number of training data over five SAO runs for each initial training data size. The blue bars represent the number of initial training data, while the yellow bars indicate the average number of additional training data required in SAO. Since the number of initial training data is fixed, it is shown as an integer, whereas the number of additional training data is an average and may appear as a decimal. Error bars represent the range (minimum to maximum) of total training data used across the five SAO runs for each initial data size.

The results show that reducing the initial training data from 70 to 21 samples results in only a small increase in the number of additional training data—from 22.8 to 25.2. However, further reducing the initial training data from 21 to 7 samples leads to a significant increase in additional data—from 25.2 to 37.2. Consequently, the total number of training data in SAO using 21 and 7 initial samples differs by only 2.0 samples.

C. Discussion

From Fig. 5, it can be observed that the SAO initialized with 7 training data points achieved the greatest reduction in the total number of training data. However, when compared to the SAO initialized with 21 data points, the difference in the total reduction amounted to only 2.0 data points. Furthermore, judging from the increase in additional training data observed in the SAO with 7 initial data points, the effect of accuracy improvement in the surrogate model brought by additional training data appears to be weaker than that in the SAO with 21 initial data points. In contrast, when reducing the number of initial training data from 70 to 21, the impact of additional data on model accuracy was more significant. Therefore, it is reasonable to assume that SAO starting with fewer than 21 initial data points would, in many cases, require a larger total number of training data compared to the present trials.

In addition, Fig. 4 demonstrates that the discrepancy between the predicted value of the surrogate model at the optimal solution and the actual simulation result is dramatically reduced during the early stages of SAO, as additional training data are incorporated. This result highlights that, in SAO employing INN, the additional training data provide highly informative contributions. Building on this finding, Fig. 6 illustrates our hypothesis: as the number of initial training data decreases below a certain level, the amount of additional training data required to identify the global optimum increases sharply. With a sufficiently large number of initial training data, the global optimum can naturally be located at an early stage. Even if the number of initial training data is moderately reduced, the delay in reaching the global optimum remains relatively small. However, if the initial training data are excessively reduced, the search for the global optimum is likely to be significantly delayed. To verify this hypothesis, Fig. 7 summarizes the number of additional training data required until convergence to the global optimum for each setting of initial training data. Five trials were performed for each amount of initial training data, and the upper and lower limits of the number of additional training data required for rough optimization for each SAO are shown as error bars. The curve connecting the average values is shown as a gray dashed line. This curve remains almost flat

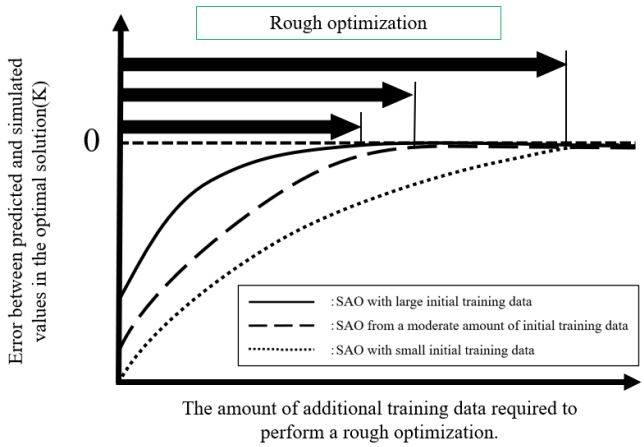

Figure 6. Hypothesis on the change in the number of additional training data required for global search depending on the initial training data.

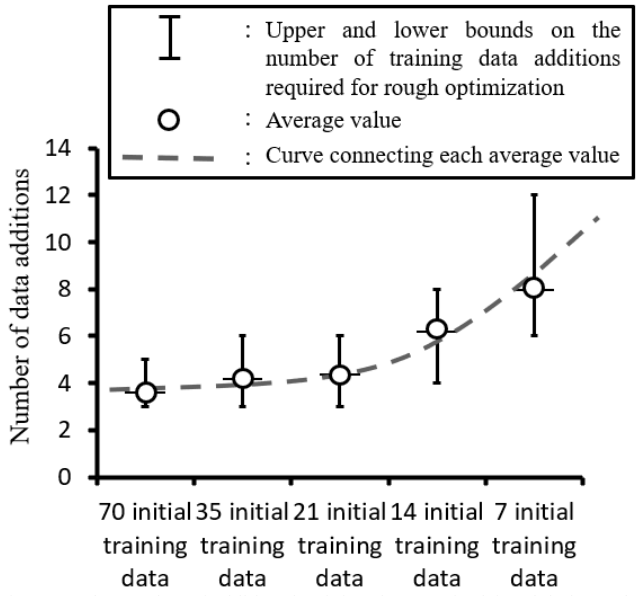

Figure 7. The number of additional training data required for global search depending on the initial training data.

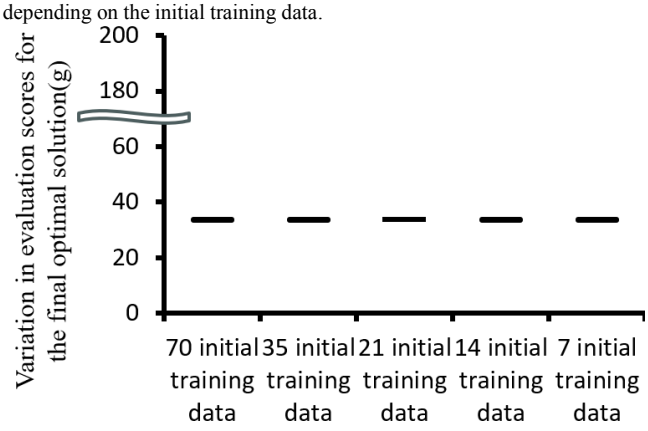

Figure 8. Variation in evaluation scores for the final optimal solution in SAO from each initial training data.

between 70 and 21 initial training data points, whereas below 14 data points, the slope steepens, indicating a rapid increase in the required additional data. These results suggest that SAO initialized with 21 or more training data provides both smaller variance and fewer additional data requirements, while SAO

with 14 or fewer initial data points results in greater variance and larger additional data requirements. Thus, SAO with 21 initial data points is considered the most appropriate setting.

We next examined the stability of design variables and the validity of the optimization process itself. Fig. 8 shows the distribution of evaluation scores of the final optimal solutions obtained by SAO with different numbers of initial training data. Although intermediate values during the optimization process occasionally reached around 200 g, the final solutions consistently converged to approximately 33.5 g across all cases, confirming that SAO successfully performed optimization regardless of the number of initial data. Furthermore, Fig. 9 presents the distributions of the most variable design variables among the optimal solutions obtained in each SAO. The two variables showing notable variation were the SiC chip size and the thickness of the heat exchanger, with the SiC chip size exhibiting particularly large variability. The average values vary for each initial training data set, likely due to the small number of trials (5 for each set). However, considering the significant digits, this difference has little impact on the evaluation value. Therefore, it does not contradict the results in Fig. 8. However, we evaluated the degree of variation because we thought it was necessary to evaluate the stability of the solution convergence. For the SiC chip size, the standard deviations were 0.013 mm and 0.011 mm when the initial training data set was 7 and 14, respectively, indicating significant variation. On the other hand, when the initial training data set was reduced from 10 to 3 times the design variables, the standard deviations were 0.005 mm, 0.008 mm, and 0.007 mm, confirming that the variation in the optimal solution position was somewhat suppressed. In summary, the analyses from Figs. 8 and 9 indicate that while reasonable optimal solutions can still be obtained even with fewer initial data, from the perspective of solution quality, it is desirable to secure at least three times as many initial training data as the number of design variables.

Based on these findings, we conclude that SAO with 21 initial training data points (i.e., three times the number of design variables) represents the most balanced setting among the conditions investigated, successfully achieving both data reduction and stability. Moreover, the prediction accuracy around the optimal solution and its vicinity did not deteriorate significantly, while the required total training data were reduced by approximately 50.2% compared with SAO initialized with 70 training data points.

By using the above-mentioned methods for optimization, we will confirm whether system optimization that takes into account the complexity of the model due to system integration can be properly performed using finite element method (FEM) simulation, which can be performed in a short time due to the extremely small number of components. Fig. 10 presents the results for a particular SAO optimal solution obtained by three different evaluation approaches: (a) thermal prediction under the conventional assumption of 45° heat spreading, (b) FEM simulation, and (c) INN prediction trained with FEM simulation data. Method (a) is essentially a one-dimensional heat flow model in which each component is evaluated independently and the overall thermal resistance is assumption (a), the predicted temperature was 147.4 °C, while FEM simulation (b) yielded 174.9 °C, since the spreading angle

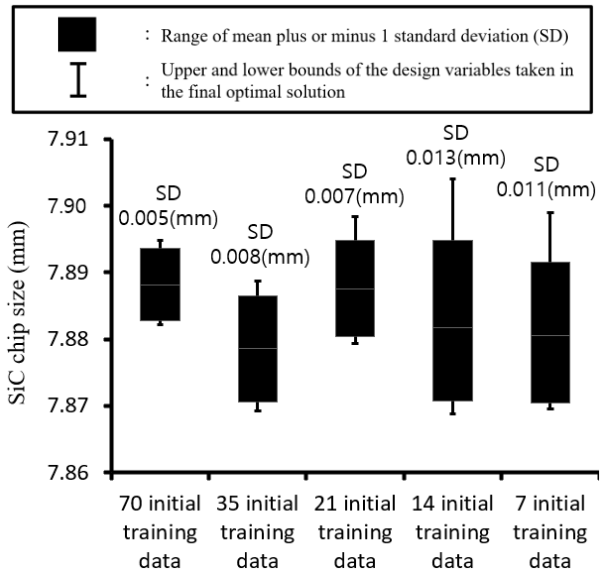

Figure 9. The variability of the design variables that showed the greatest variation among the optimal solutions obtained by SAO using each initial training data set.

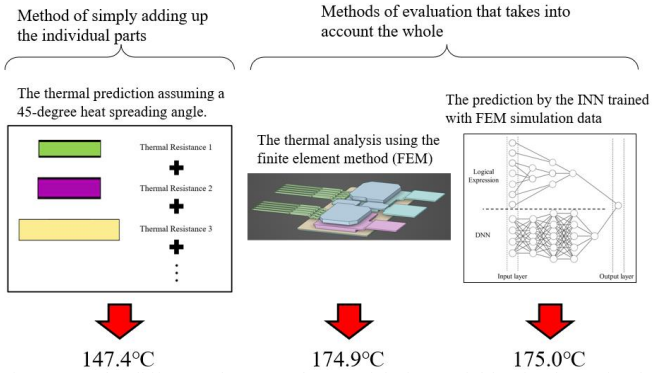

Figure 10. The difference between the set of design variables for the optimal solution obtained by running SAO, evaluated using a model that evaluates each variable individually, and evaluated using a model that considers the entire solution.

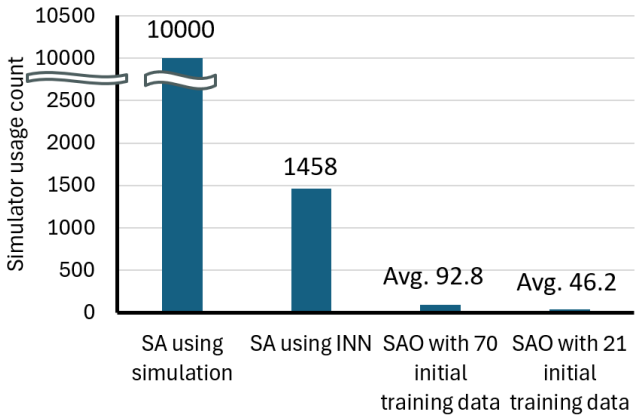

Figure 11. The number of times the simulator was used to complete the optimization for each of the four methods.

deviates from 45° due to anisotropic thermal conductivities of components. The INN prediction (c) gave 175.0 °C, which closely matched the FEM simulation result. These results clearly demonstrate that the widely used “45° heat spreading assumption” produces substantial errors compared with FEM analysis. Consequently, employing INN trained with FEM

simulation data as a surrogate prediction model is considered more appropriate for reliable thermal evaluation.

Fig. 11 compares the total number of simulator evaluations required to complete optimization using four approaches: simulated annealing with direct simulations, SA using INN,SAO with 70 initial training samples, and SAO with 21 initial training samples. In simulated annealing with direct simulations, if the number of cooling steps is 100 and 100 trials are conducted at each temperature, a total of 10,000 simulator evaluations are required. However, by employing the INN, this number can be reduced to $3^6 \times 2 = 1458$ (Bonding material has only two levels), which corresponds to approximately 15% of the original requirement. Furthermore, SAO starting from 70 initial training samples required only 92.8 simulations in total, representing more than a 95% reduction. Notably, SAO starting from 21 initial training samples—investigated in this study—required only 46.2 simulations in total, cutting the simulation count by more than half again.

These results demonstrate that, in implementing SAO for product optimization, using INN in combination with reduced initial training data has the potential to further reduce the total amount of training data required and significantly shorten the design time.

V. Conclusion

This study examined how efficiently the optimal solution can be reached in the optimization of complex systems caused by system integration by applying the SAO method using surrogate models. As a result, the following conclusions were obtained:

- By optimizing the balance between reducing the initial training data and increasing the additional training data, a hypothesis was proposed that the number of simulations and the computation time could be further reduced, and it was clarified that even if the initial training data were reduced to 21, the required additional training data did not increase significantly.
- Furthermore, as a result of analyzing the number of additional training data until the end of rough optimization, it was clarified that when the initial training data are 21 or more, the number of additional training data is small and its variation is also small, while when the initial training data are 14 or fewer, the number of additional training data increases and its variation also becomes larger. From this, it can be determined that 21 initial training data are appropriate.
- In SAO starting from any number of initial training data, an evaluation value of 33.5 g for the device weight was obtained, confirming that the solution quality is stable. However, although the variable that showed the largest variation among the design variables in the optimal solution was nearly constant, when the initial training data were 21, the standard deviation was 0.007 mm, while when the initial training data were 14, the standard deviation was 0.013 mm, revealing a large difference. This indicates that a smaller standard deviation is more adaptable and appropriate for a wider variety of optimization problems. From the perspective of optimization, it can also be judged that SAO starting from 21 initial training data is appropriate.
- A prediction model was constructed by individually evaluating each component and simply summing them, as well as by comprehensively evaluating the entire system considering interactions between components based on FEM simulation results, and the evaluation results of each were compared. As a result, the former showed a large error compared with the FEM simulation results, while the latter remained at an extremely small error, clarifying that an optimal solution with high accuracy can be obtained in a short time. This clearly demonstrates the effectiveness of optimization in systems that have been complexly system-integrated.

References

[1] J. Zheng, Z. Li, L. Gao, G. Jiang, “A parameterized lower confidence bounding scheme for adaptive metamodel-based design”, Engineering Computations, 33-7 (2016), pp.2165-2184.

[2] H. Liu, Y.S. Ong, J. Cai, “A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design”, Structural and Multidisciplinary Optimization, 57-1 (2018), pp. 393-416.

[3] B.J. Bichon, M.S. Eldred, L.P. Swiler, S. Mahadevan, J.M. McFarland, “Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions”, American Institute of Aeronautics and Astronautics, 46-10 (2008), pp. 2459-2468.

[4] Parviz Mohammad Zadeh, Mohsen Sayadi, Amirreza Kosari, “An efficient metamodel-based multi-objective multidisciplinary design optimization framework”, Applied Soft Computing, No. 74 (2019), pp. 185-215

[5] X. Li, Q. Yang, Y. Wang, X. Han, Y. Cao, L. Fan, J. Ma, “Development of surrogate models in reliability-based design optimization: A review”, Mathematical Biosciences and Engineering, 18-5 (2021), pp. 6386-6409.

[6] I. Negrin, M. Kripka, V. Yepes, “Metamodel-assisted design optimization in the field of structural engineering: A literature review”, Structures, 52 (2023), pp. 609-631.

[7] Y. Yun, “Metamodels and multi-objective optimization methods”, Design & Systems Conference, 55-9 (2011), pp.367-373.

[8] Q. Lin, Q. Zhou, J. Hu, Y. Cheng, Z. Hu, “A Sequential Sampling Approach for Multi-Fidelity Surrogate Modeling-Based Robust Design Optimization”, Journal of Mechanical design, 144(11) (2022) 111703.

[9] Y. Xiong, W. Chen, K.L. Tsui, “A new variable-fidelity optimization framework based on model fusion and objective-oriented sequential sampling”, Journal of Mechanical design, 130(11) (2008) 111401.

[10] Y. Iwata, K. Fujishiro, H. Wakamatsu, “Simulation Approximators Using linear and Nonlinear Integration Neural Networks”, Systems, Control and Information, 67-8 (2023), pp. 243-250.

[11] Y. Iwata, H. Wakamatsu, “Investigation of Methods for Expanding the Application of Integration Neural Network and Their Appropriate Structures”, Proceedings of the 26th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)(2025)

[12] K.Toyoshima, T.Takao, Y.Iwata, H.Wakamatsu, “Balancing Accuracy and Data Reduction Using Record of Additional Data in Sequential Approximate Optimization”, Proceedings of the 26th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)(2025)

[13] Y. Sakamoto, “Assembly Technologies of Double-sided Cooling Power Modules”, Denso technical review, 16(2011).https://www.denso.com/jp/ja/-/media/global/business/innovation/review/16/16-doc-8-ja.pdf