

Article

Pareto-Active-Region-Guided Sequential Surrogate Modeling for CFD-Based Multi-Objective Optimization of Liquid-Cooled Battery Thermal Management Systems

Zhanming Luo ¹, Lei Wang ^{1,2,*} and Deyong Song ¹

¹ Underwater Engineering Research and Development Department, China Ship Scientific Research Center, Wuxi 214082, China; zmluo702@163.com (Z.L.); songdy@cssrc.com.cn (D.S.)

² School of Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

* Correspondence: wangl702@163.com

Abstract

Computational fluid dynamics (CFD)-driven optimization of engineering systems is often constrained by high computational cost, particularly when surrogate models must be constructed from limited simulation samples. Although surrogate-assisted multi-objective optimization can substantially reduce CFD evaluations, local prediction errors in decision-sensitive Pareto regions may alter feasibility classification and engineering recommendations near active constraints. To address this issue, this study proposes a Pareto-active-region-guided sequential surrogate modeling framework (PAR-SSM) for multi-objective optimization of liquid-cooled battery thermal management systems. Starting from 15 face-centered central composite design (FCCD) samples, the framework selectively introduces additional high-fidelity CFD evaluations into Pareto-active and constraint-sensitive regions, yielding a 21-sample refined surrogate model. Rather than uniformly improving global prediction accuracy, PAR-SSM directs the limited CFD budget toward regions where surrogate errors can directly influence engineering decisions. After model freezing, three independent Fluent cases were used exclusively for validation, yielding mean absolute deviations of 0.098 °C for maximum temperature and 0.341 °C for temperature difference, while also revealing residual feasibility risk near active constraint boundaries. Application to an autonomous underwater vehicle (AUV) battery module showed that the $N = 3$ configuration dominated the nominally constrained Pareto set and provided a favorable thermal-hydraulic trade-off under low auxiliary energy consumption. Overall, PAR-SSM provides a decision-oriented strategy for balancing computational cost and optimization credibility in CFD-intensive, constrained multi-objective design.

Keywords: computational fluid dynamics (CFD); surrogate modeling; sequential design optimization; Pareto-active-region-guided optimization; multi-objective optimization; battery thermal management; autonomous underwater vehicle (AUV)

Academic Editor: Krzysztof Rogowski

Received: 5 August 2026

Revised: 19 August 2026

Accepted: 19 August 2026

Published: 21 August 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Lithium-ion battery-based energy systems are increasingly adopted in high-performance mobile platforms because of their high energy density and power capability. However, increasing energy demands, compact system integration, and prolonged high-load operation have intensified thermal management challenges. Autonomous underwater

vehicles (AUVs) represent a typical application where battery thermal behavior directly affects mission endurance and operational reliability. Their battery modules are usually installed within confined pressure-resistant hulls, where limited onboard energy, restricted internal convection, and long-duration operation can accelerate heat accumulation. Previous studies on underwater platforms have identified energy availability, compact system integration, and energy-efficient operation as key factors governing AUV endurance [1–3]. Excessive battery temperature accelerates degradation and increases safety risks, whereas non-uniform temperature distribution promotes cell inconsistency and uneven aging. Therefore, battery thermal management influences not only safety and lifetime but also the effective energy available for propulsion and payload operation.

Existing battery thermal management systems (BTMSs) mainly include air cooling, liquid cooling, phase-change-material cooling, heat pipes, and hybrid configurations. Zhang et al. [4] systematically reviewed the principal thermal-management approaches for lithium-ion batteries. Fu et al. [5] further summarized the characteristics of liquid-cooling systems, while Ghalkhani and Habibi [6] discussed the integration of battery thermal management with advanced battery-management strategies. Recent reviews have also highlighted the importance of efficient thermal control under fast-charging conditions [7] and compact battery configurations [8]. Compared with air cooling and passive thermal-buffering approaches, liquid cooling provides relatively high heat-transfer capability, controllable operating conditions, and compact integration. For AUV-specific applications, Li et al. [9] experimentally investigated an anisotropic expanded-graphite/paraffin composite thermal-management strategy, while Li et al. [10] subsequently introduced a metal-foam/paraffin hybrid configuration for confined AUV battery packs. These studies highlight the need for application-specific thermal management under the confined installation conditions of AUVs.

Recent liquid-cooling studies have increasingly focused on structural and operating-parameter optimization. Wang et al. [11] investigated a bio-inspired liquid-cooling plate and demonstrated the influence of channel arrangement on heat-transfer performance and hydraulic resistance. Zhao et al. [12] analyzed distributed liquid-cooling plates for cylindrical batteries and highlighted the role of flow allocation in temperature control. From a broader numerical perspective, Chen and Hu [13] reviewed CFD modeling approaches for pack-level battery thermal-management systems. At the parameter-optimization stage, Ding et al. [14] quantified the influence of channel parameters on battery temperature distribution, while Xu et al. [15] optimized a cold-plate splitter to improve flow distribution and thermal uniformity. More recently, CFD has been integrated with multi-objective optimization. Liu et al. [16] optimized a bionic cold plate considering coupled thermal and hydraulic responses, and Aydin et al. [17] combined CFD with a multi-objective genetic algorithm for immersion-cooling design. More specifically for AUV applications, Li et al. [18] combined CFD, response-surface modeling, and multi-objective optimization for a Tesla-type minichannel structure, while Li et al. [19] proposed a speed-adaptive ram liquid-cooling strategy under varying AUV operating conditions. These studies demonstrate the capability of CFD for physically interpretable thermal-management optimization; however, direct CFD-based optimization remains computationally expensive because each candidate design requires an independent high-fidelity evaluation.

Surrogate-assisted optimization has therefore been introduced to reduce the computational burden of CFD-intensive design. Data-driven thermal models have been developed for rapid prediction of battery thermal behavior [20]. Ye et al. [21] combined a BP neural-network surrogate with NSGA-II to balance maximum temperature, temperature non-uniformity, and pressure drop, whereas Cho et al. [22] employed sequential multi-objective Bayesian optimization to reduce the number of required high-fidelity evaluations. Surrogate-assisted strategies have also demonstrated computational advantages in

other thermo-fluid process systems [23]. NSGA-II remains a widely adopted benchmark for obtaining diverse trade-off solutions in multi-objective optimization [24]. More generally, engineering metamodeling studies have shown that response-surface, Kriging/Gaussian-process, and neural-network models differ substantially in nonlinear approximation capability, interpretability, data requirements, and local predictive behavior [25]. A qualitative comparison of representative surrogate and sequential optimization approaches is summarized in Table 1.

Table 1. Qualitative comparison of representative surrogate and sequential optimization approaches.

Approach	Main Strengths	Main Limitations	Typical Refinement Focus
Response-surface models	Low computational cost; good interpretability; well suited to small CFD datasets.	Limited flexibility for strong local nonlinearities; accuracy may deteriorate in sparsely sampled local regions.	Global response approximation and design-space trend identification.
Kriging/Gaussian-process models	Flexible nonlinear approximation; predictive uncertainty can be estimated explicitly.	Greater modeling complexity; performance depends on kernel and hyperparameter selection.	Global approximation, uncertainty reduction, and adaptive sampling.
Neural-network surrogates	Strong nonlinear representation capability; flexible for complex input–response relationships.	Generally requires more training data; lower interpretability; increased risk of overfitting in small-sample settings.	Global nonlinear mapping and surrogate-assisted optimization.
Sequential Bayesian optimization	Adaptive and sample-efficient for expensive evaluations; can balance exploration and exploitation.	Performance depends on the acquisition function and probabilistic model; implementation is comparatively complex.	Expected improvement, uncertainty reduction, or Pareto-front approximation.
PAR-SSM (this study)	Directs limited CFD evaluations toward Pareto-active and constraint-sensitive regions that govern engineering decisions.	Does not provide formal uncertainty quantification or robust optimization; boundary candidates still require independent CFD verification.	Local engineering decision reliability, including feasibility, dominance, and candidate ranking.

As summarized in Table 1, the distinguishing feature of PAR-SSM is not a universally more flexible surrogate architecture, but a shift in the refinement objective from global approximation toward local engineering decision reliability.

The remaining challenge is that global surrogate accuracy does not necessarily guarantee engineering decision reliability. In multi-objective optimization, final design recommendations are usually selected from limited regions near the Pareto front, knee regions, or active constraint boundaries. These decision-sensitive regions may be insufficiently represented by initial CFD samples even when the surrogate exhibits satisfactory global accuracy. Consequently, local prediction errors may alter feasibility classification, change dominance relationships, shift the predicted Pareto front, or affect the ranking of candidate solutions. Under a limited CFD budget, uniform improvement of global surrogate accuracy may allocate expensive simulations to regions with little influence on the final engineering decision, while critical Pareto-active regions remain insufficiently resolved. Sequential surrogate optimization methods based on expected improvement, Bayesian optimization, scalarization, and uncertainty reduction have been developed for expensive multi-objective problems [26–29]. However, many existing strategies in CFD-based engineering optimization primarily emphasize global exploration, uncertainty reduction, or overall Pareto-front approximation, without explicitly considering whether local prediction errors will influence engineering feasibility and solution ranking. Therefore, a decision-oriented refinement strategy is needed to allocate additional high-fidelity simulations according to their potential influence on feasibility, dominance, and final engineering selection.

For AUV battery modules, this issue has direct system-level significance because coolant circulation consumes the same limited onboard energy required for propulsion, navigation, communication, and sensing. Excessive cooling power may reduce mission endurance, whereas insufficient cooling capability may compromise battery safety and reliability. Therefore, this study proposes a Pareto-active-region-guided sequential surrogate modeling framework (PAR-SSM) for reliable CFD-based multi-objective optimization of a liquid-cooled AUV battery module under limited computational budgets. The maximum battery temperature, module temperature difference, and hydraulic power consumption are selected as competing objectives representing thermal safety, temperature uniformity, and auxiliary energy demand, respectively. Unlike uniform refinement or generic uncertainty-driven sampling, PAR-SSM shifts the refinement focus from global surrogate accuracy toward local decision reliability and selectively allocates additional CFD evaluations to predicted Pareto-active regions and adjacent constraint boundaries, where surrogate errors have the greatest impact on solution feasibility, dominance, and ranking.

The main contributions of this study are as follows:

- (1) A Pareto-active-region-guided sequential surrogate modeling strategy is proposed to improve the decision reliability of CFD-based BTMS optimization under limited sample and computational budgets, with additional CFD resources preferentially directed toward decision-sensitive regions rather than global model refinement.
- (2) A decision-oriented refinement mechanism is developed to selectively allocate additional high-fidelity CFD evaluations toward Pareto-active regions and active constraint boundaries, rather than uniformly improving surrogate accuracy across the entire design space.
- (3) The proposed framework is demonstrated on an AUV liquid-cooled battery module considering maximum temperature, temperature difference, and hydraulic power consumption as competing objectives, and the reliability of the obtained trade-off solutions is verified through independent high-fidelity CFD simulations performed after the final surrogate model is frozen and excluded from subsequent model updating.

2. Materials and Methods

2.1. Equivalent Battery Module and Optimization Problem Definition

All CFD simulations were performed using ANSYS Fluent 2021 R1 (ANSYS, Inc., Canonsburg, PA, USA), while surrogate-model construction, statistical analysis, Pareto-front screening, and multi-objective optimization were implemented in Python 3.8.20 (Python Software Foundation, Wilmington, DE, USA).

The research object is the internal liquid-cooling cold-plate thermal-management design of a power battery module for a parent-child AUV. Because the focus of this study is the optimization of the cold-plate structure and operating parameters rather than the fine-scale electrochemical-thermal coupling process inside individual cells, a module-level equivalent modeling method was used to represent the battery heat source. Under the fixed 2C discharge condition, the heat generation of the battery module was treated as a volumetric heat source. While maintaining the same total heat-generation rate, the thermal load was applied to the equivalent heat-source region of the battery module. Specifically, the 75 cylindrical cells in a single module were equivalent to five heat-generating bars. This treatment reduces the complexity of transient CFD calculations while retaining the dominant heat-transfer relationship among the battery module, the cold plate, and the coolant. Verification of the equivalent heat-source treatment and the mesh-independence assessment are provided in Appendix A to support the numerical reliability of the CFD model.

The equivalent volumetric heat source of the battery module can be expressed as follows:

$$q''' = \frac{Q_{\text{gen}}}{V_{\text{eq}}} \quad (1)$$

where Q_{gen} is the total heat-generation rate under the specified discharge condition, V_{eq} is the volume of the equivalent heat-source region, and q''' is the volumetric heat-generation rate. The heat-generation treatment and equivalent-model verification are further described in Appendix A.1. All candidate designs were compared under the same 2C discharge condition and identical heat-source boundary. Therefore, this equivalent model is suitable for the relative performance evaluation of different liquid-cooling structures. Heat conduction in the solid domains and convective heat transfer in the fluid domain were obtained by CFD calculation. The temperature field and pressure-drop response at $t = 1800$ s were uniformly extracted as the data basis for subsequent surrogate modeling and optimization analysis.

The number of channels, channel width, and inlet velocity were selected as design variables. The channel number N is a discrete structural variable with values of 3, 5, and 7. The channel width w is a continuous geometric variable in the range of 4–8 mm, and the inlet velocity v is a continuous operating variable in the range of 0.05–0.20 m s^{−1}. The equivalent module and the main design variables are shown in Figure 1.

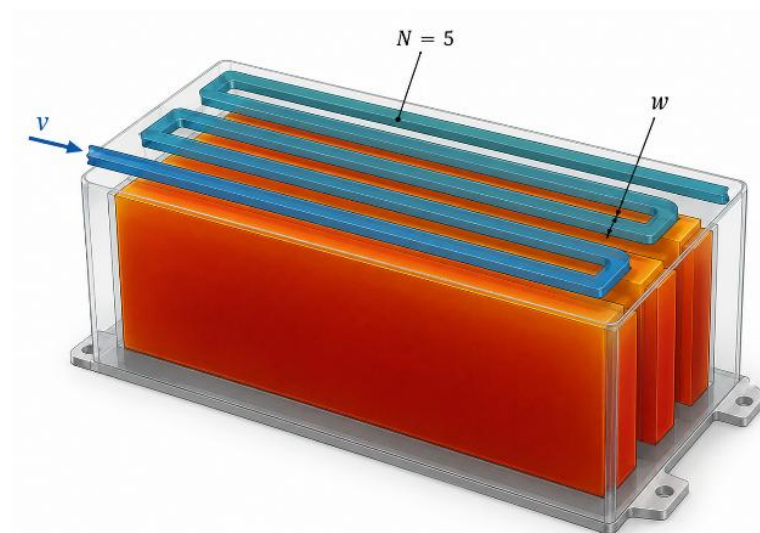


Figure 1. Schematic of the equivalent battery module with a liquid-cooling plate and the main design variables.

Let the design vector be defined as $x = (N, w, v)$. The optimization objectives were to minimize the maximum temperature, maximum temperature difference, and hydraulic power:

$$\min_x F(x) = [T_{\text{max}}(x), \Delta T(x), P_{\text{hyd}}(x)] \quad (2)$$

where

$$\Delta T = T_{\text{max}} - T_{\text{min}} \quad (3)$$

Because the cooling structure uses a single total inlet and a single total outlet, the total inlet flow rate is defined as

$$Q = vwh \quad (4)$$

where h is the channel height. The hydraulic power is defined as

$$P_{\text{hyd}} = \Delta p Q \quad (5)$$

Here, P_{hyd} represents the hydraulic power required for coolant transport and is not equivalent to the actual electrical power of a pump. In this study, $T_{\text{max}} \leq 50\text{ }^{\circ}\text{C}$ and $\Delta T \leq 10.5\text{ }^{\circ}\text{C}$ were used as nominal engineering screening constraints to determine the thermal feasibility of candidate designs. These limits are treated as case-specific nominal screening criteria rather than uncertainty-inclusive safety guarantees for all AUV operating conditions. Table 2 summarizes the design variables, objectives, constraints, and calculation definitions used in this study.

Table 2. Design variables, objectives, constraints, and calculation definitions.

Category	Symbol/Item	Value or Range	Type	Definition/Calculation
Design variable	Channel number	$N = 3, 5, 7$	Discrete structural variable	Number of liquid-cooling channels
Design variable	Channel width	$w = 4\text{--}8\text{ mm}$	Continuous geometric variable	Width of each liquid-cooling channel
Design variable	Inlet velocity	$v = 0.05\text{--}0.20\text{ m s}^{-1}$	Continuous operating variable	Coolant inlet velocity
Fixed geometric parameter	Channel height	$h = 0.005\text{ m}$	Constant	Fixed channel height for flow-rate and hydraulic-power calculation
Thermal response	Maximum temperature	T_{max}	Objective	Peak temperature of the battery module at $t = 1800\text{ s}$
Thermal response	Maximum temperature difference	$\Delta T = T_{\text{max}} - T_{\text{min}}$	Objective	Temperature-uniformity indicator of the module
Hydraulic response	Pressure drop	Δp	Intermediate response	Pressure difference between outlet and inlet
Hydraulic objective	Hydraulic power	$P_{\text{hyd}} = \Delta p Q$	Objective	Coolant transport power; not pump electrical power
Flow-rate definition	Total inlet flow rate	$Q = vwh$	Derived variable	Single total inlet and single total outlet; no multiplication by N
Optimization objective	Multi-objective function	$\min(T_{\text{max}}, \Delta T, P_{\text{hyd}})$	Objective vector	Simultaneous minimization of peak temperature, temperature non-uniformity, and hydraulic cost
Engineering constraint	Maximum-temperature limit	$T_{\text{max}} \leq 50\text{ }^{\circ}\text{C}$	Constraint	Nominal thermal-safety screening criterion
Engineering constraint	Temperature-difference limit	$\Delta T \leq 10.5\text{ }^{\circ}\text{C}$	Constraint	Nominal temperature-uniformity screening criterion
CFD condition	Discharge condition	2C discharge	Fixed boundary	Same heat-generation condition for all candidate designs
Response extraction time	Transient endpoint	$t = 1800\text{ s}$	Fixed output time	Temperature and pressure responses extracted at the end of transient simulation

2.2. Surrogate-Model Construction and Pareto-Active-Region-Guided Sequential Refinement

To avoid high-density transient CFD calculations over the entire design space, an initial global surrogate model was first constructed based on 15 face-centered central composite design (FCCD) samples. The surrogate model was used to establish the mapping between the design variables (N , w , v) and the thermo-fluid responses, including T_{max} , ΔT , and Δp . The pressure-drop response was modeled as $\ln(\Delta p)$ to reduce the influence of magnitude variation on fitting stability.

Before response-surface fitting, the three design variables were transformed into coded variables to improve numerical scaling and provide a consistent representation of the three-level structural and continuous factors:

$$x_N = \frac{N-5}{2}, x_w = \frac{w-6}{2}, x_v = \log_2 \left(\frac{v}{0.10} \right) \quad (6)$$

where ($x_N = -1, 0, 1$) correspond to ($N = 3, 5, 7$), respectively, while ($x_v = -1, 0, 1$) correspond to inlet velocities of 0.05, 0.10, and 0.20 m s⁻¹. An indicator variable I_3 was additionally introduced to allow the ($N = 3$) structural layer to exhibit a local width response different from those of the other channel-number levels:

$$I_3 = \begin{cases} 1, & N = 3, \\ 0, & N \neq 3. \end{cases} \quad (7)$$

Although (N) was represented by the coded terms x_N and (x_N^2) during response-surface fitting, its admissible values remained restricted to the three discrete structural levels ($N = 3, 5, 7$) throughout the subsequent optimization.

Candidate response-surface structures containing linear, quadratic, interaction, and structure-specific terms were evaluated separately for each thermo-fluid response. The final V21 response models can be expressed as

$$\hat{T}_{\max} = \beta_0 + \beta_1 x_N + \beta_2 x_w + \beta_3 x_v + \beta_4 x_N^2 + \beta_5 x_N x_w + \beta_6 I_3 x_w + \beta_7 I_3 x_w^2 + \beta_8 x_w x_v + \varepsilon_T \quad (8)$$

$$\widehat{\Delta T} = \gamma_0 + \gamma_1 x_N + \gamma_2 x_w + \gamma_3 x_v + \gamma_4 x_N^2 + \gamma_5 x_N x_w + \gamma_6 x_w x_v + \varepsilon_{\Delta T} \quad (9)$$

$$\ln(\widehat{\Delta p}) = \eta_0 + \eta_1 x_N + \eta_2 x_w + \eta_3 x_v + \eta_4 x_N^2 + \eta_5 x_N x_w + \eta_6 x_w^2 + \eta_7 x_w x_v + \varepsilon_p \quad (10)$$

The $\hat{T}_{\max}$ model therefore permits an ($N = 3$)-specific linear and quadratic width response, whereas the $\widehat{\Delta T}$ model retains the global width-velocity interaction and the $\ln(\widehat{\Delta p})$ model additionally contains a global quadratic width term. Model structure was selected using leave-one-out cross-validation LOO-cv rather than by minimizing the fitting residual of individual enrichment points. LOO- Q^2 was used as the primary selection criterion; when multiple candidate structures fell within a predefined near-optimal Q^2 tolerance, the model containing fewer parameters was preferred to control model complexity and reduce the risk of overfitting. The adopted Q^2 tolerances were 0.020 for $\hat{T}_{\max}$ and $\widehat{\Delta T}$, and 0.002 for $\ln(\widehat{\Delta p})$.

The initial surrogate is not intended to provide the final engineering design directly; rather, it serves to capture the dominant response trends and to provide a predictive basis for identifying subsequent Pareto-active regions.

After the initial surrogate model was established, a non-design point was used for external checking, and the initial predicted Pareto front was extracted. For a multi-objective minimization problem, a candidate belongs to the nondominated solution set if it cannot be simultaneously improved in all objectives by any other candidate. All nondominated solutions constitute the Pareto set. In this study, the design region close to the low-energy branch and thermal-constraint boundary of the initially predicted Pareto front was defined as the Pareto-active region. In the implemented workflow, the Pareto-active region was identified according to three engineering features: the predicted low-hydraulic-power Pareto branch, proximity to the nominal thermal constraints, and competition among the discrete structural levels. Additional CFD points were therefore selected to examine whether local surrogate errors in these regions could alter feasibility classification, Pareto dominance, or branch ranking, rather than according to a fixed geometric-distance threshold from the Pareto front.

Unlike blind sample enrichment over the entire design space, this study adopted a Pareto-active-region-guided sequential refinement strategy. This strategy concentrates

the limited CFD budget on checking the local credibility of the initial surrogate model in decision-sensitive engineering regions. If the checking process indicates local thermal-response deviations, the informative CFD samples are incorporated into the training set and the surrogate model is refined. The sequential refinement followed a study-specific staged procedure. After the E0 external check, A1–A4 were used to examine the initially identified Pareto-active and structural-competition regions, while A5 was subsequently introduced as a low-flow, narrow-channel anchor point. After completion of these planned checking and refinement stages, the resulting 21-sample model was designated as V21 and frozen. Thus, the present termination was based on completion of the predefined refinement stages rather than on an automatic statistical convergence criterion.

The final sequentially refined surrogate model, V21, was frozen before the subsequent optimization and validation steps; after freezing, no additional Fluent results were used to update the model. The overall sequential modeling and refinement workflow is shown in Figure 2, and the detailed implementation procedure of PAR-SSM is summarized in Algorithm 1.

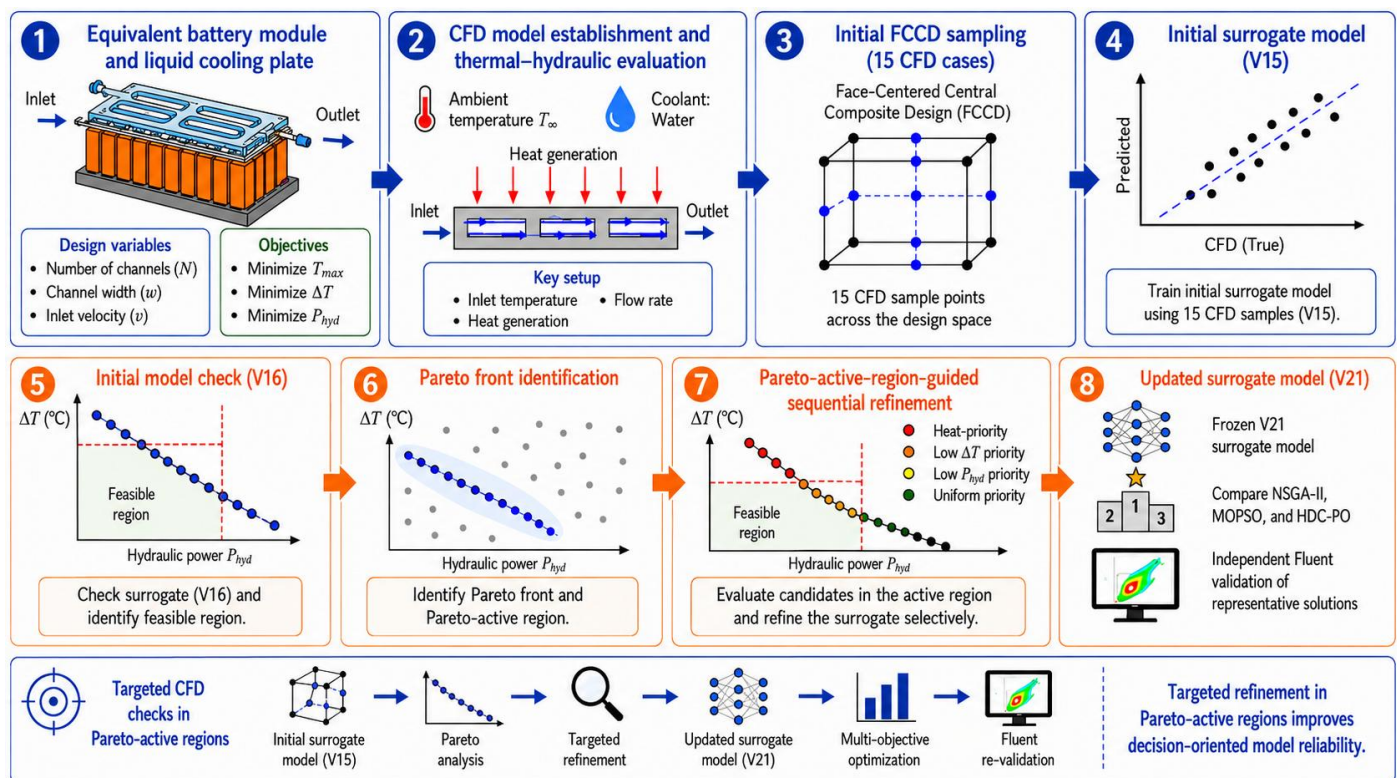


Figure 2. Workflow of the Pareto-active-region-guided sequential surrogate-modeling strategy. Arrows indicate the sequential progression of the workflow; the blue and orange frames distinguish the initial-modeling and Pareto-active refinement stages, respectively, while the colored points indicate different refinement priorities. The five-pointed star and ranking symbol denote the selection and comparison of representative solutions before independent Fluent re-validation.

Algorithm 1. Pseudocode of the PAR-SSM framework

Input: Initial 15-case FCCD dataset D15, design space Ω , and nominal thermal constraints.

Phase A: Initial surrogate construction

1. Construct the initial surrogate models using the 15 FCCD samples.
2. Perform the E0 non-design-point external check and incorporate E0 into the training dataset to obtain the 16-sample dataset and the corresponding V16 surrogate model.
3. Generate the surrogate-predicted Pareto front using V16.

Phase B: Pareto-active-region identification and sequential refinement

4. Identify decision-sensitive regions according to:
 - (a) the low-hydraulic-power Pareto branch;
 - (b) proximity to the nominal thermal constraints; and
 - (c) competition among the discrete structural levels.
5. Select A1–A4 as targeted CFD checking points within the identified decision-sensitive regions.
6. Perform high-fidelity CFD calculations for A1–A4 and compare the CFD responses with the current surrogate predictions.
7. Incorporate the informative checked points into the training dataset and refit the response-specific surrogate models using LOO- Q^2 -based model selection.
8. Introduce A5 as a low-flow, narrow-channel local anchor point and update the surrogate models after CFD checking.
9. Form the final 21-sample V21 surrogate model and freeze the model structures and coefficients.

Phase C: Multi-objective optimization and independent validation

10. Construct the reference Pareto front and perform multi-objective optimization using the frozen V21 surrogate model.
11. Select representative engineering candidates from the obtained trade-off solutions.
12. Perform independent Fluent validation using F1–F3; these cases are excluded from all subsequent surrogate-model updating.

Output: Frozen V21 surrogate models, engineering-constrained reference Pareto front, and independently validated representative candidates.

The core advantage of this method is that surrogate-model refinement is not designed merely to reduce the global average error, but to improve the credibility of the final engineering decision region. For the liquid-cooled battery module, the region that truly affects the design recommendation is often located near the low-energy Pareto branch and the thermal-constraint boundary. Therefore, concentrating additional samples in these regions is computationally more efficient and engineering-oriented than uniform enrichment over the full design space.

2.3. Reference Pareto Front, Algorithm Comparison, and Independent Validation

After the final V21 surrogate model was frozen, a high-density scan was performed in the surrogate-model space to construct the reference Pareto front under nominal engineering constraints. The scan considered three discrete structural levels, $N = 3, 5$, and 7 , and the variables w and v were grid-searched within each structural level. For each structural level, a 301×301 uniform grid was generated over the channel width w and the coded velocity variable x_v , resulting in $3 \times 301 \times 301 = 271,803$ surrogate-model evaluations in total. Subsequently, thermally feasible points were filtered using $T_{\max} \leq 50^\circ\text{C}$ and $\Delta T \leq 10.5^\circ\text{C}$, and the engineering-feasible nondominated solution set was further extracted.

It should be emphasized that this reference Pareto front was obtained from the V21 surrogate-model space rather than from the true global Pareto front of the original CFD model. Its role is twofold: first, to identify the competition among thermal safety, temperature uniformity, and hydraulic power for different structural levels; and second, to serve as a benchmark for evaluating the approximation capability of multi-objective algorithms.

This study compared NSGA-II, MOPSO, and HDC-PO in their ability to approximate the reference front under an identical surrogate-model evaluation budget. NSGA-II is a classical elitist nondominated sorting genetic algorithm proposed by Deb et al. [24], whereas MOPSO originates from the multi-objective particle-swarm framework in which Coello Coello et al. [30] introduced Pareto dominance into particle-swarm optimization. HDC-PO is defined in this study as a mixed-variable optimization framework consisting

of discrete-structure stratification, intra-layer continuous multi-objective search, and cross-layer global nondominated filtering. More specifically, HDC-PO first separates the mixed-variable problem into the three discrete structural layers ($N = 3, 5, 7$) and performs continuous Pareto search within each layer. Layer-wise local archives and a global nondominated archive are maintained simultaneously. After the initial search stage, the remaining evaluation resources are redistributed according to the Pareto contribution of the individual structural layers, followed by cross-layer global nondominated filtering. This formulation avoids treating the discrete channel-number variable as an artificial continuously varying optimization dimension.

All three algorithms called the same V21 surrogate model and used the same evaluation budget, archive size, and number of repeated runs. For a controlled comparison, all three algorithms used a population or swarm size of 120,100 generations or iterations, a total evaluation budget of 12,120 surrogate-model calls per run, and an external archive capacity of 120. Each algorithm was independently repeated 30 times under the same reference front and normalization settings.

Algorithm performance was evaluated using IGD, HV, and final feasible ratio. IGD measures the distance between the algorithm-generated solution set and the reference Pareto front; a smaller value indicates better front approximation. HV evaluates the coverage capability of the nondominated solution set in the objective space; a larger value indicates better convergence and diversity. The final feasible ratio was defined as the proportion of candidate solutions in the final population or swarm that simultaneously satisfied the two nominal thermal constraints. Zitzler et al. [31] systematically discussed the performance assessment of multi-objective optimizers and provided the basis for comparing algorithms with set-quality indicators. The purpose of algorithm comparison is not to replace the reference Pareto front or directly determine the final engineering design, but to evaluate the efficiency and stability of different search strategies in identifying useful trade-off solutions under a limited budget. The mixed discrete–continuous search procedure of HDC-PO is summarized in Figure 3.

The mixed discrete–continuous search procedure of HDC-PO

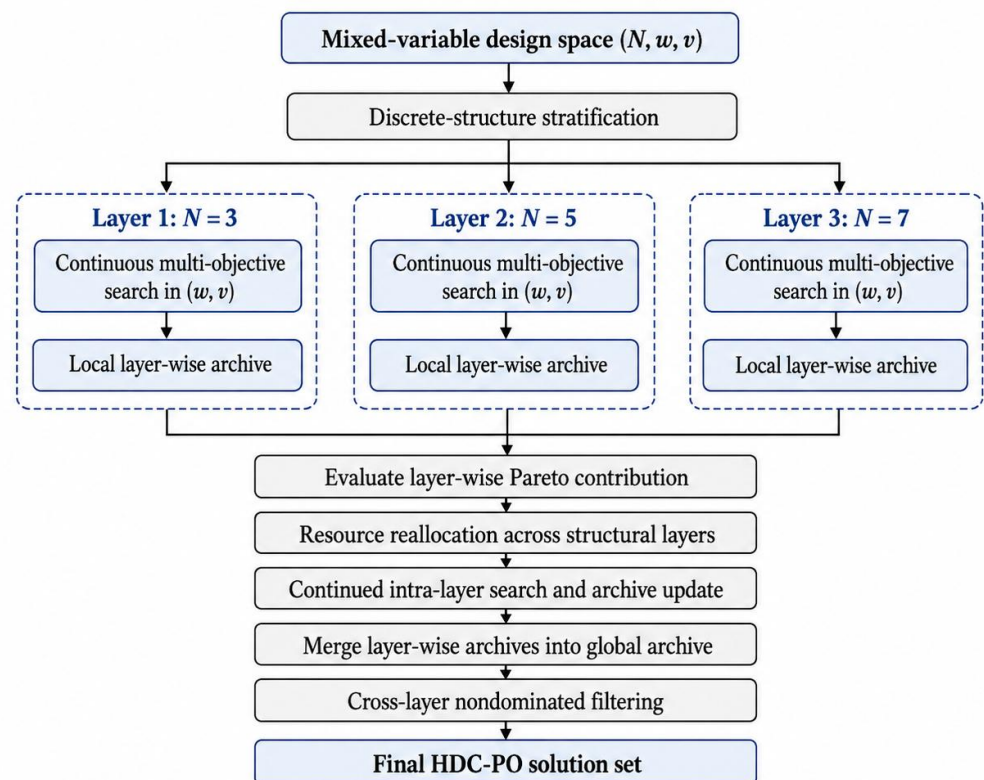


Figure 3. Mixed discrete–continuous search procedure of HDC-PO.

Finally, representative candidate designs were selected for independent Fluent validation after the V21 surrogate model had been completely frozen. The independent validation points were not used in any model training, and their CFD results were not fed back into the surrogate model. This step was used to determine whether boundary candidates screened by the surrogate model still satisfied the nominal thermal constraints after full CFD recalculation, thereby supporting the final engineering recommendation.

3. Results and Discussion

3.1. Formation and Performance Evaluation of the V21 Sequentially Refined Surrogate Model

The initial 15 FCCD samples covered the main boundary and center regions of the design space and provided a basis for identifying the global trends of the thermo-fluid responses of the liquid-cooling cold plate. However, for a multi-objective optimization problem, global trends do not necessarily indicate reliable prediction in the engineering decision region. The final recommended or compared candidate designs are usually concentrated near the Pareto front, especially near the low-energy branch and the thermal-constraint boundary. Therefore, the initial surrogate model was not directly used for final engineering judgment. Instead, its prediction results were first used to identify the Pareto-active region, and targeted CFD checks were then performed to correct local deviations. To clarify the roles of different CFD samples in initial modeling, sequential refinement, and final independent validation, Table 3 summarizes the sample roles, training assignment, and validation boundaries.

Table 3. Sample roles and training-validation boundaries.

Sample Group	Sample ID	Cases	Representative Condition	Initial Role	Included in Fi- nal V21 Train- ing?	Final Independent Fluent Vali- dation?	Function in the Study
Initial design samples	CCD-1- CCD-15	15	Face-centered CCD samples over the N , w , v design space	Initial surrogate-model training	Yes	No	Establish baseline global response trends
Initial external check	E0	1	$N = 3$, $w = 4.0$ mm, $v = 0.10$ m s ⁻¹	External check for the initial 15-sample model	Yes	No	Verify initial-model behavior in a non-design regular region before dataset expansion
Pareto-active check	A1	1	$N = 3$, $w = 4.0$ mm, $v = 0.068$ m s ⁻¹	Check point near the $N = 3$ Pareto-active branch	Yes	No	Examine local prediction reliability near the low-energy/thermal branch
Pareto-active check	A2	1	$N = 5$, $w = 4.0$ mm, $v = 0.090$ m s ⁻¹	Check point near the $N = 5$ competitive branch	Yes	No	Evaluate the competing $N = 5$ branch in the active region
Pareto-active check	A3	1	$N = 3$, $w = 4.8$ mm, $v = 0.055$ m s ⁻¹	Key local check point	Yes	No	Reveal local thermal-response deviation of the initial surrogate in the $N = 3$ low-flow, local-width region
Pareto-active check	A4	1	$N = 5$, $w = 4.8$ mm, $v = 0.090$ m s ⁻¹	Check point for $N = 5$ width effect	Yes	No	Support intermediate refinement of the $N = 5$ local response
Low-energy anchor	A5	1	$N = 3$, $w = 4.0$ mm, $v = 0.055$ m s ⁻¹	Anchor point unseen by the intermediate model	Yes	No	Anchor the low-flow, narrow-channel branch and support final V21 construction
Final independent validation	F1	1	$N = 3$, $w = 4.0$ mm, $v = 0.060$ m s ⁻¹	Independent Fluent validation after V21 freezing	No	Yes	Validate the balanced trade-off branch
Final independent validation	F2	1	$N = 3$, $w = 4.0$ mm, $v = 0.075$ m s ⁻¹	Independent Fluent validation after V21 freezing	No	Yes	Validate the thermal-priority branch and assess ΔT boundary risk
Final independent validation	F3	1	$N = 5$, $w = 8.0$ mm, $v = 0.068$ m s ⁻¹	Independent Fluent validation after V21 freezing	No	Yes	Validate the wide-channel, low- ΔT competing branch

As shown in Table 3, the 15 CCD samples were used mainly to establish the initial global response trends. E0 and A1–A5 were stage-wise checking and model-updating samples; after being checked, they were incorporated into the V21 training set. In contrast, F1–F3 were newly calculated Fluent cases after the V21 model had been frozen. They were not involved in any model training and were used only for final independent validation. Accordingly, the final V21 model was constructed from 21 training cases, whereas F1–F3 constituted a separate three-case validation set. This classification of sample roles prevents model-refinement samples from being mixed with final validation samples and thus keeps the independent-validation boundary of V21 clear.

The sequential checking results indicate that the initial surrogate exhibited reasonable predictive capability in well-sampled regions, but retained a risk of underestimating the peak temperature in the $N = 3$, low-flow, local-width region. This phenomenon suggests that, for the liquid-cooled battery module, the heat-transfer capacity under low-flow conditions approaches the constraint boundary. Small changes in channel width may alter the heat-transfer matching between the coolant and the heat-source region, leading to nonlinear local thermal responses. In other words, the low-energy region is not a smooth and inherently safe area, but rather a narrow window in which thermal safety and hydraulic cost are highly sensitive. To further visualize the sample-updating process and the training-validation boundary, Figure 4 presents the sequential surrogate-model refinement workflow used in this study.

Sample Roles and Training–Validation Boundaries in Sequential Surrogate Refinement

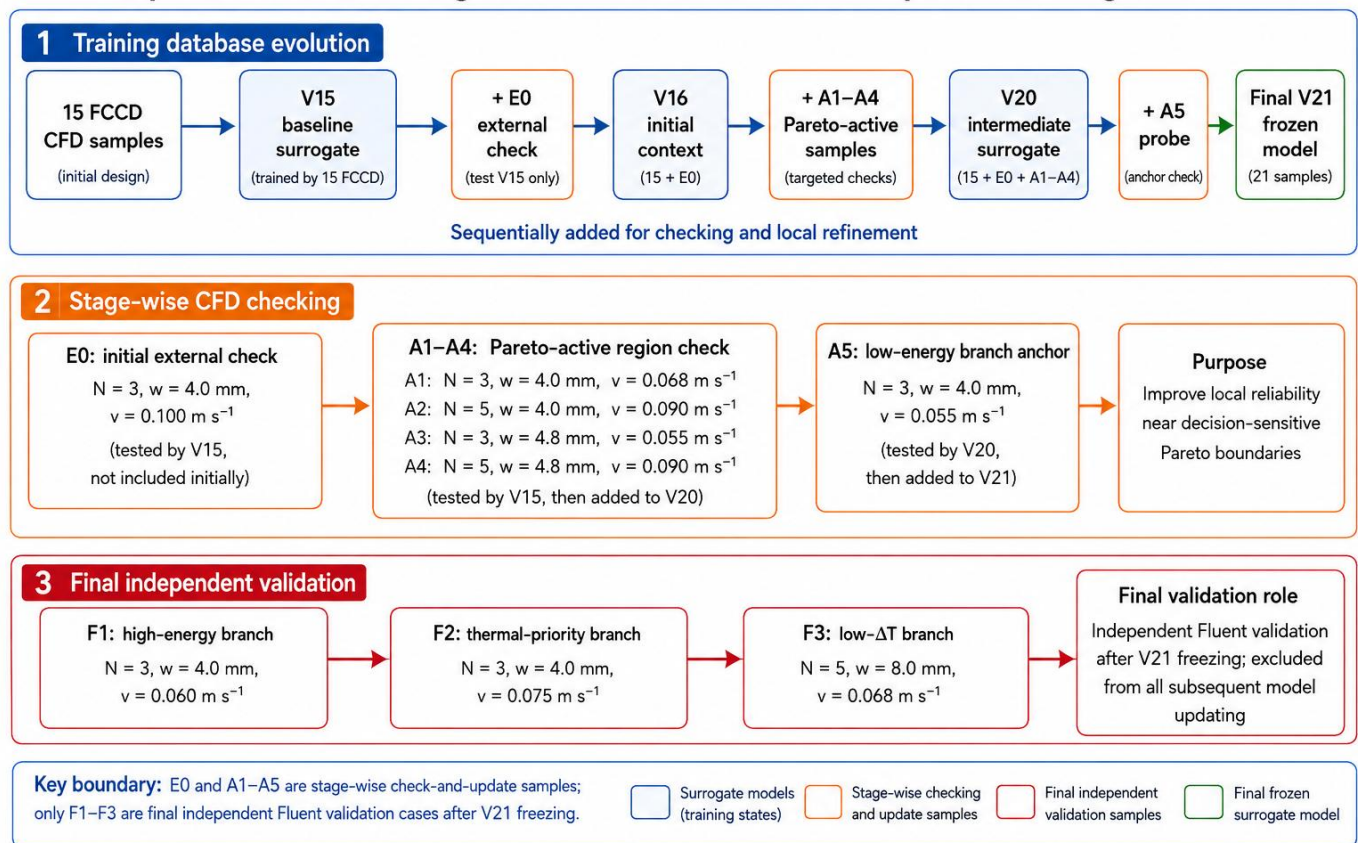


Figure 4. Sample roles and training-validation boundaries during sequential surrogate refinement.

After E0 and A1–A4 were incorporated into training, the intermediate refined model improved the local description capability in the Pareto-active region. A5 was further placed at $N = 3$, $w = 4.0$ mm, and $v = 0.055$ m s⁻¹ to anchor the low-flow narrow-channel branch. After A5 was incorporated into the V21 training set, the LOO- Q^2 -based model-selection procedure favored a $T_{\max}$ model containing $N = 3$ -specific linear and quadratic width terms together with a width-velocity interaction. This process demonstrates that the surrogate-model refinement in this study was not a simple increase in sample number, but a targeted allocation of limited CFD budget to the region most likely to affect engineering judgment. The local width response of the $N = 3$ low-flow branch is shown in Figure 5.

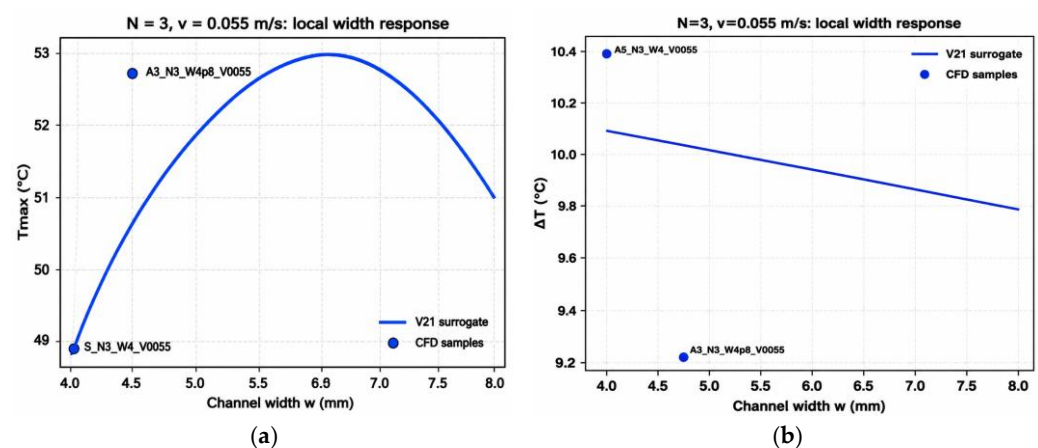


Figure 5. Local width response at $N = 3$ and $v = 0.055$ m s⁻¹: (a) maximum temperature ($T_{\max}$); (b) maximum temperature difference (ΔT).

The performance of the final V21 model is listed in Table 4. The $T_{\max}$ model achieved an R^2 of 0.928 and an LOO- Q^2 of 0.592, indicating that it can reasonably describe the overall variation trend of peak temperature and provides moderate out-of-sample predictive capability within the sampled design domain under the present small-sample setting. The ΔT model achieved an R^2 of 0.780 and an LOO- Q^2 of 0.559, indicating that temperature-uniformity prediction is more difficult than the prediction of peak temperature and pressure drop. This is reasonable because ΔT is not a single-point response; rather, it is jointly determined by the highest and lowest temperatures within the module and is therefore more susceptible to local flow distribution and heat-transfer-path variations. In contrast, the $\ln(\Delta p)$ model achieved an R^2 of 0.998 and an LOO- Q^2 of 0.994, suggesting that the pressure-drop response is mainly governed by channel geometry and inlet velocity and is therefore more regular and more stable for surrogate modeling.

Table 4. Structure and fitting/cross-validation performance of the final V21 surrogate models.

Response	Selected Model Structure	R^2	Adjusted R^2	LOO- Q^2	LOO-RMSE	LOO-MAE
$T_{\max}$	M3 with $N = 3$ -specific linear/quadratic width terms and w - v interaction	0.928	0.880	0.592	1.255	0.734
ΔT	M3 with global w - v interaction	0.780	0.686	0.559	0.422	0.359
$\ln(\Delta p)$	M3 with global quadratic width term and w - v interaction	0.998	0.996	0.994	0.056	0.046

Note: LOO denotes leave-one-out cross-validation. The pressure-drop response was modeled as $\ln(\Delta p)$.

Therefore, V21 can serve as the basis for response-surface analysis, Pareto screening, and algorithm comparison. Nevertheless, candidates close to the $\Delta T = 10.5$ °C constraint boundary still require independent Fluent recalculation for final judgment. This result also demonstrates the necessity of the sequential modeling strategy: the surrogate model is not a replacement for CFD, but an engineering screening tool for efficiently identifying candidate branches and guiding limited CFD validation.

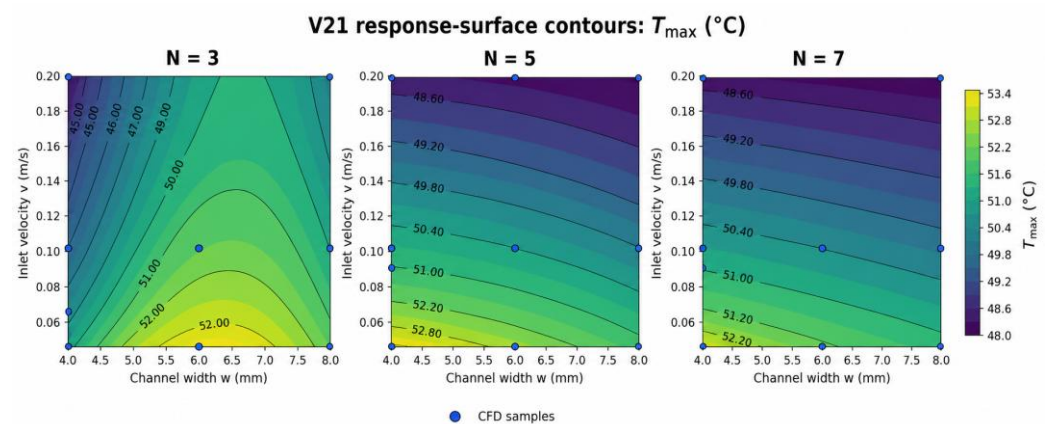
3.2. V21 Response-Surface Characteristics and Thermo-Fluid Coupling Mechanisms

The response-surface results of the frozen V21 model show that the effects of N , w , and v on the thermal response and hydraulic cost of the battery module are not simple monotonic superpositions; instead, they present typical thermo-fluid coupling characteristics. Increasing the inlet velocity v generally enhances the convective heat-transfer capability of the coolant and thereby reduces $T_{\max}$. However, a higher velocity also increases the pressure drop and causes P_{hyd} to increase rapidly. The influence of channel width w is more complicated. At a prescribed inlet velocity, increasing the channel width simultaneously increases the coolant flow area and total flow rate, while changing the hydraulic diameter, wetted-perimeter-to-flow-area ratio, and local flow distribution. The resulting competition between increased coolant heat-capacity transport and changes in local convective heat-transfer effectiveness can therefore produce a non-monotonic thermal response rather than a simple monotonic dependence on channel width. Variation in channel number N further changes the channel layout and heat-transfer coverage, causing different structural levels to compete in terms of thermal safety, temperature uniformity, and energy consumption. The response-surface contours obtained from the V21 model are shown in Figure 6.

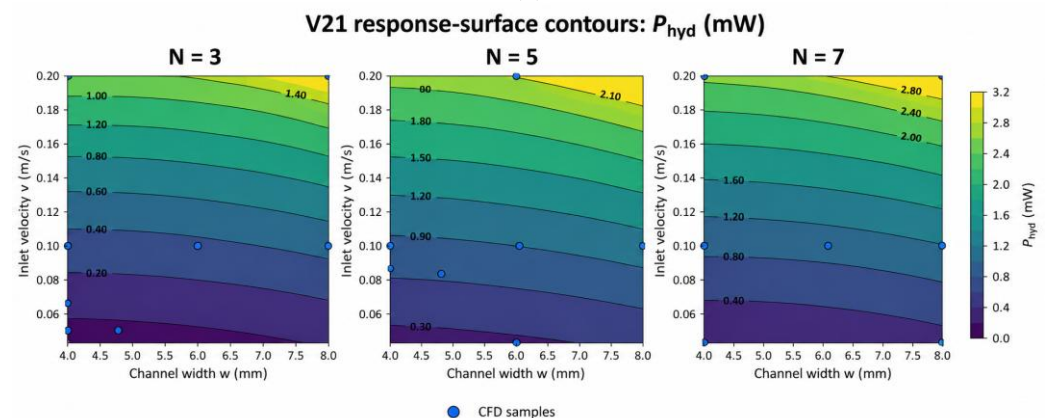
For the $N = 3$ low-flow section in Figure 5a, the predicted non-monotonic variation in with channel width should therefore be interpreted as a case-specific local response arising from these competing transport effects. Because the local CFD support remains limited, this behavior is not interpreted as a universal monotonic or non-monotonic law outside the investigated design range.

From an engineering perspective, the $N = 3$ structure shows strong competitiveness in the low-energy region. This does not imply that fewer channels are always superior. Rather, under the single total-inlet and total-flow-rate definition used in this study, the $N = 3$ branch can maintain acceptable peak temperature and temperature uniformity at lower hydraulic power.

In contrast, the $N = 5$ structure has certain advantages in reducing ΔT , but some of its candidates are limited by $T_{\max}$ or P_{hyd} . Although $N = 7$ increases the number of cooling channels, this structural change does not yield an overall Pareto advantage under the present design range and flow-rate definition.



(a)



(b)

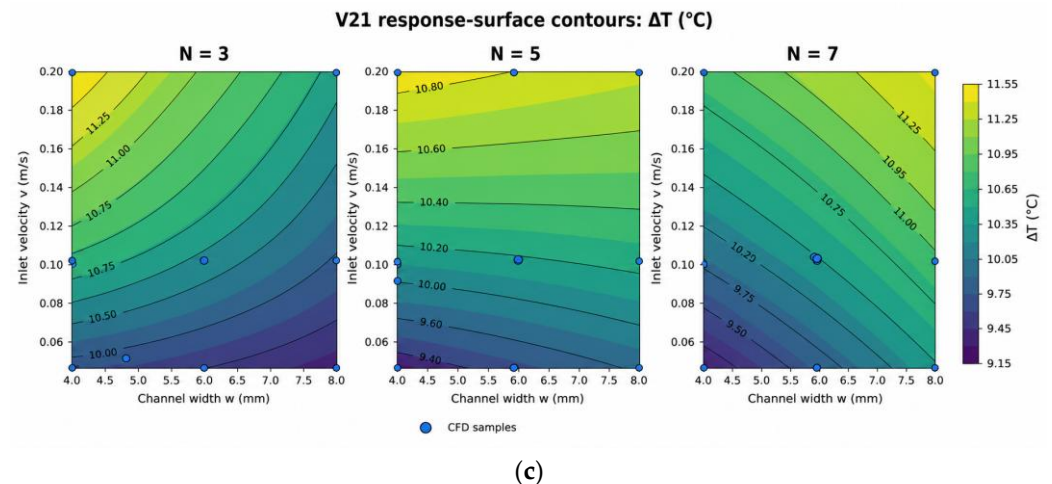


Figure 6. Response-surface contours of the final V21 surrogate model under different channel-number levels: (a) maximum temperature ($T_{\max}$); (b) hydraulic power (P_{hyd}); (c) maximum temperature difference (ΔT).

It should be noted that the V21 results support only the conclusions within the present variable ranges. For the $N = 3$ -dominated branch, it can be stated that this branch tends toward the lower boundary of channel width, i.e., $w \approx 4$ mm, within the investigated width range. This reflects a local optimum trend under the current cold-plate structure and flow-rate definition, and it should not be extrapolated as a globally optimal width for all cold-plate sizes or all AUV battery modules.

3.3. Nominally Constrained Pareto Front and Identification of Engineering Branches

In the V21 surrogate-model space, a high-density scan was conducted over the three discrete structural levels $N = 3, 5$, and 7 , producing a total of 271,803 candidate points. After applying the nominal engineering constraints of $T_{\max} \leq 50$ °C and $\Delta T \leq 10.5$ °C, 37,348 thermally feasible points were obtained, accounting for 13.74% of all candidates. Further nondominated filtering resulted in 138 engineering-feasible Pareto points, including 120 points for $N = 3$, 18 points for $N = 5$, and no points for $N = 7$. The distribution of thermally feasible candidates and Pareto points for each channel-number level is listed in Table 5.

Table 5. Distribution of thermally feasible candidates and Pareto points for each channel-number level.

N	Candidate Points	Thermally Feasible Points	Thermally Feasible Ratio	Constrained Pareto Points
3	90,601	5692	6.28%	120
5	90,601	27,693	30.57%	18
7	90,601	3963	4.37%	0

This result has a clear engineering implication. Although the $N = 5$ layer contains the largest number of thermally feasible points, indicating a wider feasible region under the individual thermal constraints, only a small number of its candidates enter the Pareto front after $T_{\max}$, ΔT , and P_{hyd} are considered simultaneously. Conversely, the $N = 3$ layer does not have the highest thermally feasible ratio, but it occupies most Pareto points, indicating stronger global competitiveness in the three-objective trade-off. This phenomenon reveals an important fact: the size of the thermally feasible region is not equivalent to the strength of engineering optimality. The actual design recommendation is governed by nondominance relationships among objectives rather than by the number of feasible points under a single constraint. The corresponding Pareto-front projections in the objective space are shown in Figure 7.

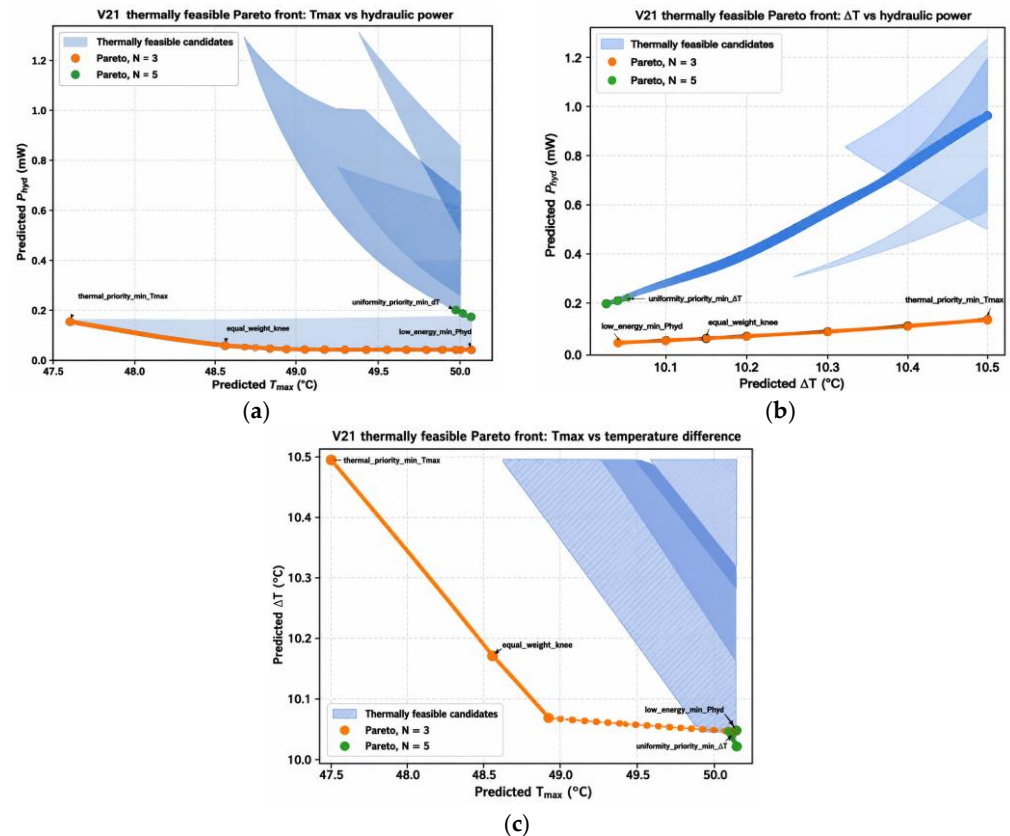


Figure 7. Nominally constrained thermally feasible V21 Pareto front in the objective space: (a) $T_{\max}$ versus P_{hyd} ; (b) ΔT versus P_{hyd} ; (c) ΔT versus $T_{\max}$. The blue shaded regions indicate the distribution of thermally feasible candidate solutions, while darker shades result from overlapping transparent regions and do not represent additional solution categories.

From the Pareto branches, $N = 3$ forms the low-energy and thermal-safety-dominated branch. Its main advantage is that it maintains the peak temperature and temperature difference within the nominal thermal constraints at relatively low P_{hyd} , making it suitable for AUV applications that are sensitive to auxiliary energy consumption. $N = 5$ forms a supplementary low-temperature-difference branch; its advantage lies mainly in temperature uniformity, but the peak-temperature boundary and increased hydraulic power must be considered. $N = 7$ does not enter the engineering-feasible Pareto front, indicating that increasing the channel number does not transform into a three-objective comprehensive advantage under the present design-variable range and constraints. Representative points on the nominally constrained V21 Pareto front are summarized in Table 6.

Table 6. Representative points on the nominally constrained V21 Pareto front.

Representative Point	<i>N</i>	<i>w</i> (mm)	<i>v</i> (m s ^{−1})	Predicted <i>T</i> _{max} (°C)	Predicted ΔT (°C)	Predicted <i>P</i> _{hyd} (mW)	<i>T</i> _{max} Margin (°C)	ΔT Margin (°C)
Thermal-priority endpoint	3	4.000	0.079	47.522	10.499	0.247	2.478	0.001
Uniformity-priority endpoint	5	7.987	0.066	50.000	9.933	0.324	0.000	0.567
Low-energy end-point	3	4.280	0.050	49.997	9.960	0.092	0.003	0.540
Equal-weight knee point	3	4.000	0.055	48.719	10.089	0.114	1.281	0.411

Note: Margins were calculated relative to $T_{\max} \leq 50$ °C and $\Delta T \leq 10.5$ °C.

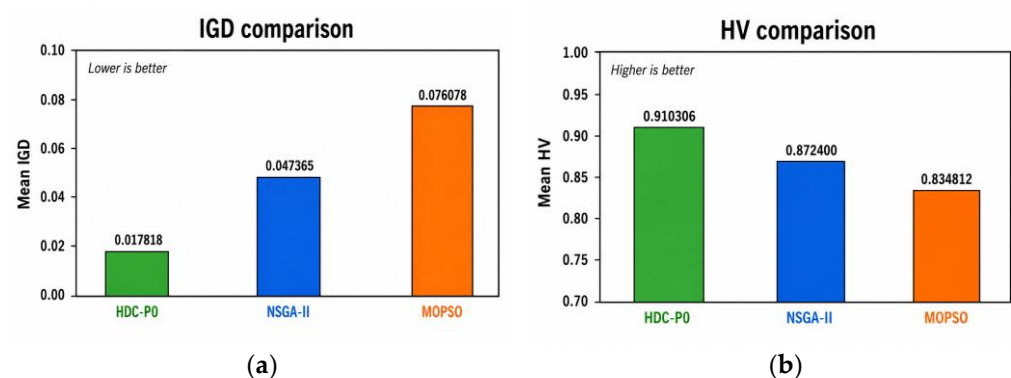
The representative points in Table 6 are mainly used to explain the boundary and trade-off characteristics of the Pareto front; they should not be directly treated as final engineering solutions without Fluent recalculation. In particular, the thermal-priority endpoint, uniformity-priority endpoint, and low-energy endpoint are all close to at least one constraint boundary. Three of the four representative points therefore lie essentially on an active thermal constraint, highlighting the need for independent CFD confirmation before engineering recommendation. This indicates that the surrogate-model Pareto front has the nature of a set of boundary candidates and requires subsequent independent CFD validation before final judgment.

3.4. Algorithm Comparison Under a Fixed Evaluation Budget

After the V21 reference Pareto front was obtained, NSGA-II, MOPSO, and HDC-PO were further compared under an identical surrogate-model evaluation budget. It should be emphasized that the purpose of the algorithms is not to find physical solutions that are better than the reference Pareto front, but to approximate that reference front within a limited number of evaluations. Because the reference front was obtained from a high-density scan, it can be used as a benchmark for algorithm evaluation in the present low-dimensional surrogate-model space. The graphical comparison of the algorithmic results is shown in Figure 8, and the quantitative statistics are listed in Table 7.

Table 7. Results of 30 repeated runs for the three algorithms under the same evaluation budget.

Algorithm	IGD Mean	IGD Median	HV Mean	HV Median	Final Feasible	Archive <i>N</i> = 3	Archive <i>N</i> = 5
					Ratio	Mean	Mean
HDC-PO	0.017818	0.003871	0.910306	0.928891	1.000	106.23	13.77
NSGA-II	0.047365	0.003972	0.872400	0.928622	1.000	99.43	20.57
MOPSO	0.076078	0.018352	0.834812	0.925006	0.465	93.67	26.33



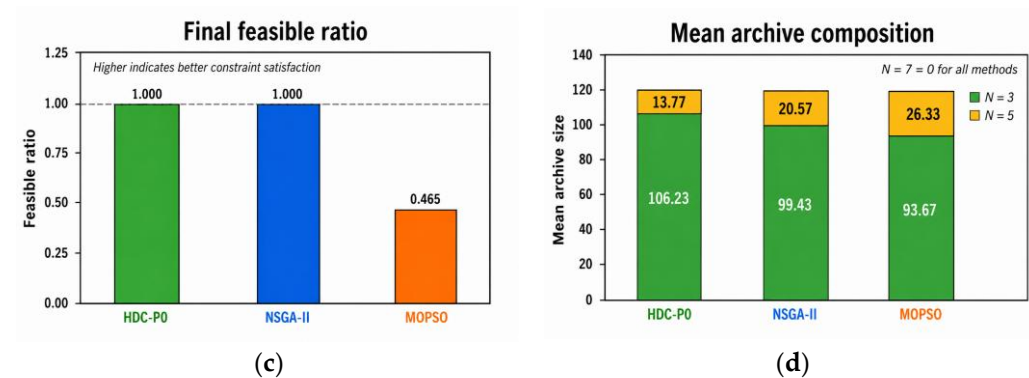


Figure 8. Performance comparison of NSGA-II, MOPSO, and HDC-PO under the same surrogate-model evaluation budget: (a) mean inverted generational distance (IGD); (b) mean hypervolume (HV); (c) final feasible solution ratio; (d) mean archive composition for different channel-number levels ($N = 3$ and $N = 5$).

The results show that HDC-PO produced the lowest mean IGD and the highest mean HV, indicating more stable overall Pareto-front approximation quality over repeated runs. Relative to NSGA-II and MOPSO, HDC-PO reduced the mean IGD by approximately 62.4% and 76.6%, respectively, while increasing the mean HV by approximately 4.35% and 9.04%. Both HDC-PO and NSGA-II achieved a final feasible ratio of 100%, whereas the corresponding value for MOPSO was 46.5%. MOPSO had a higher mean IGD and a final feasible ratio of only 0.465, suggesting that particle-swarm search was more strongly affected by boundary constraints and competition among discrete structural levels in this problem.

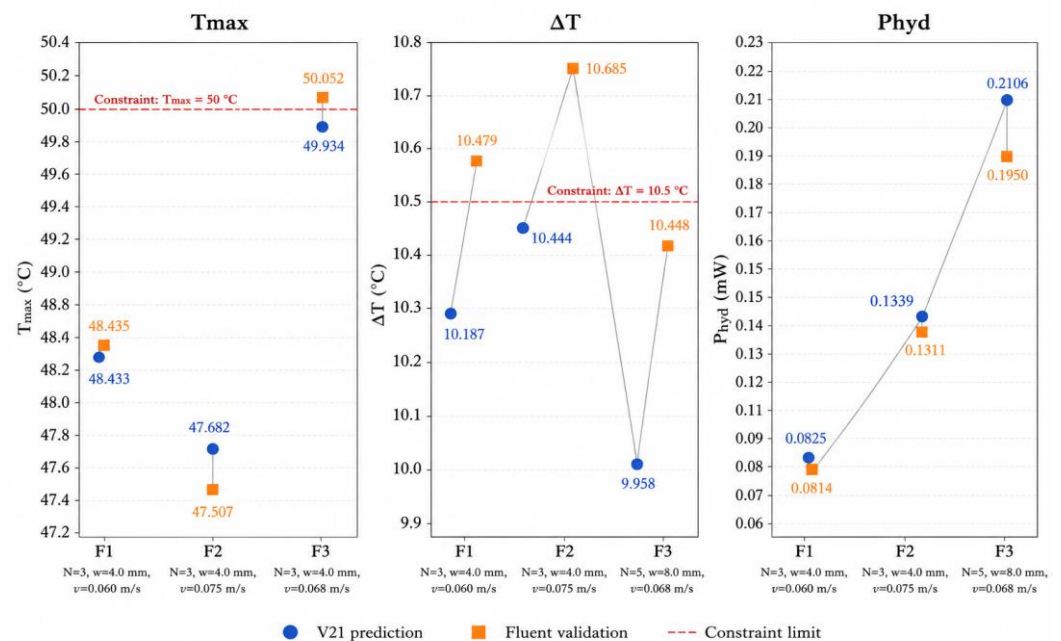
However, the advantage of HDC-PO should not be overstated. The median IGD and median HV of HDC-PO and NSGA-II were very close, implying that the two methods could obtain similar front quality when they converged successfully. The main value of HDC-PO lies in reducing the influence of low-quality repeated runs on the average performance and in producing an archive more strongly concentrated on the dominant $N = 3$ branch while retaining the supplementary $N = 5$ branch. In other words, HDC-PO should not be claimed as a universally superior new algorithm; rather, in this mixed discrete–continuous low-dimensional optimization problem, explicit stratification of the discrete structural levels improved the search stability under a finite budget.

3.5. Independent Fluent Validation and Engineering Implications

After the V21 surrogate model was frozen, F1–F3 were selected for independent Fluent validation. Unlike E0 and A1–A5, F1–F3 were not involved in any model training, and their calculation results were not used to update the surrogate model. Therefore, they can be used to evaluate the predictive capability of V21 near engineering decision branches and to make final judgments on representative candidates screened by the surrogate model. The V21 predictions and independent Fluent validation results for F1–F3 are listed in Table 8, and the comparisons of the three indicators are shown in Figure 9.

Table 8. V21 predictions and independent Fluent validation results for F1–F3.

ID	Functional Branch	N	w (mm)	v (m s ⁻¹)	Predicted $T_{\max}$ (°C)	CFD $T_{\max}$ (°C)	Predicted ΔT (°C)	CFD ΔT (°C)	CFD P_{hyd} (mW)	Final Decision
F1	Balanced trade-off branch	3	4.0	0.060	48.433	48.435	10.187	10.479	0.136	Satisfies nominal constraints; ΔT margin ≈ 0.021 °C
F2	Thermal-priority branch	3	4.0	0.075	47.682	47.507	10.444	10.685	0.218	ΔT exceeds 10.5 °C
F3	N = 5 wide-channel low- ΔT branch	5	8.0	0.068	49.934	50.052	9.958	10.448	0.325	$T_{\max}$ slightly exceeds 50 °C

**Figure 9.** Comparison between V21 predictions and independent Fluent validation results for F1–F3.

Across the three independent cases, the mean absolute deviation between V21 and Fluent was approximately 0.098 °C for $T_{\max}$ and 0.341 °C for ΔT . However, the independent CFD results also show that small prediction errors near active constraints can change the final feasibility classification.

F1 represents the balanced trade-off branch of $N = 3$. The Fluent results show that $T_{\max} = 48.435$ °C and $\Delta T = 10.479$ °C, which still satisfy the nominal constraints; however, the margin to the ΔT upper limit is only about 0.021 °C. Therefore, F1 can be regarded as a boundary candidate for balanced operation, but its temperature-uniformity safety margin is very small and should be treated cautiously in practical engineering applications.

F2 is located in the same $N = 3$ structure, but its inlet velocity is increased to 0.075 m s⁻¹. From the perspective of peak temperature, the CFD value of F2 decreases to 47.507 °C, indicating that increasing the flow velocity indeed enhances the cooling capability. However, its ΔT increases to 10.685 °C, exceeding the nominal constraint. This result illustrates an inherent thermal trade-off in the present liquid-cooling configuration: enhanced peak-temperature suppression does not necessarily improve temperature uniformity. When the coolant removes more heat near the inlet region, the internal temperature

gradient of the module may be amplified, causing a conflict between temperature reduction and temperature uniformity.

F3 represents the $N = 5$ wide-channel low- ΔT competing branch. Its CFD value of $\Delta T = 10.448$ °C indicates that this branch has an advantage in temperature uniformity. However, its $T_{\max} = 50.052$ °C is slightly higher than the 50 °C constraint. Therefore, the engineering value of the $N = 5$ branch is not that it provides the final recommended design, but that it demonstrates how low- ΔT design can still be limited by the peak-temperature constraint. This observation further supports the need for the three-objective optimization framework, because optimizing $T_{\max}$ or ΔT in isolation may lead to incomplete engineering conclusions.

The constraint violations observed for F2 and F3 also reveal the remaining boundary risk of PAR-SSM. Both candidates are located close to an active constraint and/or a geometric boundary, where the available training information is effectively one-sided and small surrogate errors can cause feasibility misclassification. Therefore, Pareto-active-region refinement reduces local prediction risk but does not eliminate the need for high-fidelity confirmation of boundary candidates.

Considering the V21 reference Pareto front and the F1–F3 independent Fluent validation results together, F2 and F3 should not be presented as final feasible recommended solutions. The main engineering recommendation should be positioned as the low-energy operating anchor of the $N = 3$ branch. This point corresponds to A5, which has already been included in the V21 training set; therefore, it should be referred to as a CFD anchor point of the low-energy-dominated branch rather than an independent validation point. F1 can be used as a balanced boundary candidate at a slightly higher flow velocity, illustrating that the temperature-uniformity margin narrows rapidly while the peak temperature is reduced.

3.6. Limitations and Robustness Considerations

The CFD data used in the present study are deterministic numerical outputs rather than experimental measurements; therefore, random measurement noise was not explicitly modeled. Nevertheless, uncertainty remains from numerical discretization, solver convergence, equivalent heat-source approximation, and surrogate-model error. PAR-SSM reduces decision-region surrogate error through targeted high-fidelity enrichment, but it does not constitute a formal uncertainty-quantification or robust-optimization framework, as also illustrated by the boundary-crossing behavior of F2 and F3.

From the perspective of marine engineering application, the present fixed-boundary results should be regarded as a nominal design basis rather than a complete mission-cycle solution. In practical AUV operation, seawater temperature, depth-related structural requirements, pump efficiency, and duty-cycle variation may shift the feasible margin of each Pareto branch. Nevertheless, the identified low-energy and boundary-sensitive branches provide a useful starting point for later robust design, because they reveal which structural and operating combinations are most sensitive to changes in the thermo-fluid boundary.

4. Conclusions

This study developed a Pareto-active-region-guided sequential surrogate modeling framework (PAR-SSM) for CFD-based multi-objective optimization of an AUV liquid-cooled battery module under a limited computational budget. Rather than uniformly refining the surrogate across the entire design space, PAR-SSM concentrates high-fidelity CFD evaluations in Pareto-active and constraint-sensitive regions where local prediction errors are most likely to alter feasibility, dominance relationships, and engineering selection. The resulting workflow integrates initial FCCD sampling, decision-sensitive region

identification, targeted CFD refinement, surrogate updating, multi-objective optimization, and independent Fluent validation after model freezing.

The results show that the initial V15 surrogate captured the overall thermal–hydraulic trends but retained local deficiencies in decision-sensitive regions. Targeted sequential enrichment enabled the final V21 model to better represent these locally sensitive responses without uniformly increasing the CFD sample density over the entire design space. Under the nominal thermal constraints, the $N = 3$ configuration dominated the engineering-feasible Pareto set and provided the most favorable compromise between thermal performance and hydraulic energy consumption, whereas $N = 5$ mainly contributed to improved temperature uniformity. No $N = 7$ solution entered the constrained Pareto front within the investigated design range.

Under the same surrogate-model evaluation budget, HDC-PO achieved the lowest mean IGD and the highest mean HV among the three tested optimizers, while NSGA-II exhibited comparable median performance. The benefit of HDC-PO should therefore be interpreted as improved search stability for the present mixed discrete–continuous problem rather than universal algorithmic superiority. More importantly, the independent Fluent validation revealed that small surrogate errors near active constraints can change feasibility classification. F1 remained nominally feasible but possessed only a small temperature-uniformity margin, whereas F2 and F3 crossed the corresponding thermal constraints. These results demonstrate that PAR-SSM improves decision-oriented screening efficiency but does not eliminate the need for high-fidelity verification of boundary candidates.

Overall, this study establishes a Pareto-active-region-guided surrogate-refinement strategy that reallocates limited high-fidelity CFD resources toward regions that directly govern engineering feasibility and trade-off selection. By combining targeted sequential enrichment with post-freeze independent CFD validation, the framework provides a practical means of balancing computational cost and decision credibility in constrained thermal-management optimization. The resulting solutions should be interpreted as nominal design benchmarks rather than uncertainty-inclusive robust designs. Future work will extend the framework through uncertainty quantification, environmental-temperature and pump-efficiency variations, battery aging, and additional high-fidelity validation under representative AUV operating conditions.

Author Contributions: Conceptualization, Z.L., L.W. and D.S.; Methodology, Z.L., L.W. and D.S.; Software, Z.L.; Validation, Z.L., L.W. and D.S.; Formal analysis, Z.L.; Investigation, Z.L.; Writing—original draft, Z.L.; Writing—review and editing, L.W. and D.S.; Visualization, Z.L.; Supervision, L.W. and D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Appendix A.1. Equivalent Battery Model and Heat-Generation Treatment

The complete AUV battery system consists of 450 cylindrical 21,700 cells in a 30-series and 15-parallel configuration. Considering the computational cost of transient thermo-fluid simulations, one representative 5S15P module containing 75 cells was selected as the research object. To reduce geometric complexity, the fifteen cylindrical cells in each column were replaced by one equivalent heat-generating region, resulting in five equivalent

heat sources. The equivalence was established by preserving the total heat generation, thermal capacity, and principal spatial distribution of the heat sources.

Battery heat generation under the 2C discharge condition was calculated using the Bernardi heat-generation model and converted into a volumetric heat source for CFD calculation:

$$\mathbf{q}''' = \frac{\mathbf{Q}}{V} \quad (11)$$

Where $\mathbf{Q}$ is the total battery heat-generation rate, V is the corresponding heat-source volume, and $\mathbf{q}'''$ is the volumetric heat-generation rate. Thermal-capacity equivalence was further maintained during model simplification, while the anisotropic thermal-conductivity characteristics of the cylindrical cells were retained.

The equivalent model and the detailed 75-cell model were simulated under identical 2C discharge and thermal boundary conditions. As shown in Figure A1, the two models produced consistent temperature-field characteristics. Their maximum temperatures were 39.01 °C and 39.15 °C, respectively, corresponding to an absolute deviation of only 0.14 °C. The equivalent model was therefore considered adequate for the subsequent comparative CFD optimization.

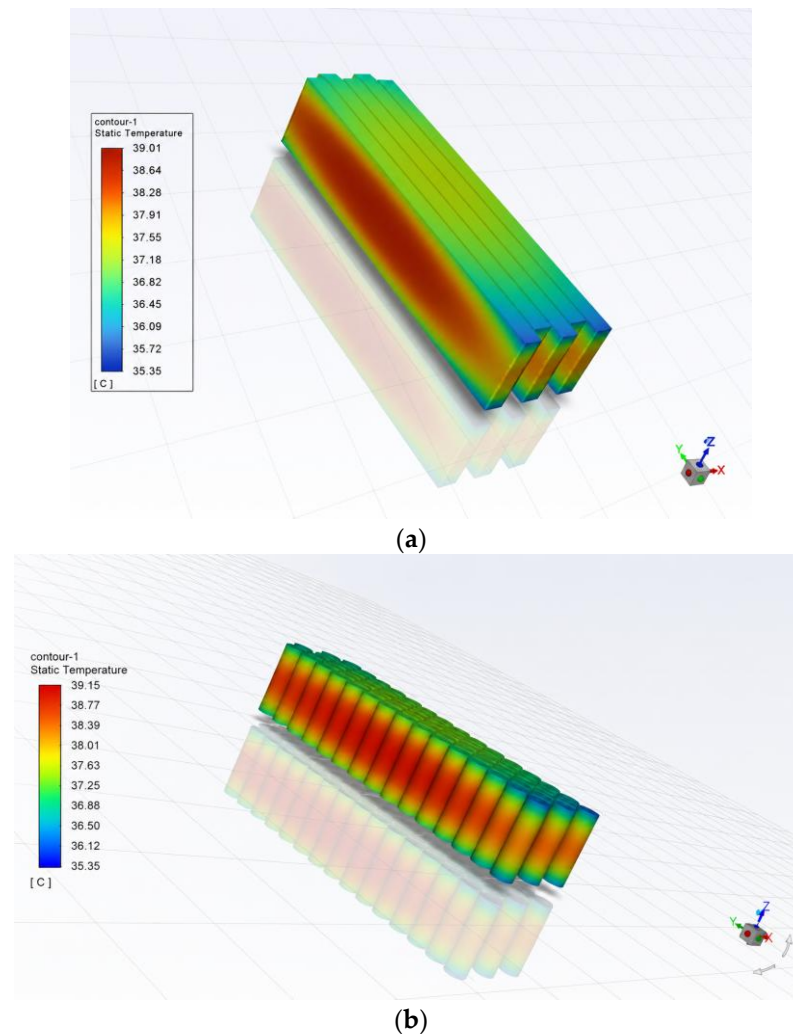


Figure A1. Comparison of temperature fields between (a) the five-region equivalent heat-source model and (b) the detailed 75-cell model.

Appendix A.2. Mesh-Independence Verification

Mesh independence was evaluated at the representative condition, mm, and m/s. Three mesh resolutions were generated for comparison. The coarse mesh contained 602,313 cells, with a heat-source mesh size of 4 mm, a maximum surface mesh size of 9.35 mm, and a maximum volume mesh size of 9 mm. The medium mesh contained 1,081,389 cells, with the corresponding mesh sizes reduced to 3, 7.5, and 7 mm, respectively. The fine mesh contained 12,549,705 cells, with mesh sizes of 2, 5.5, and 5 mm.

The three meshes produced consistent temperature-field distributions and similar transient trends in maximum temperature and pressure drop. As shown in Figure A2, the maximum-temperature and pressure-drop responses obtained using the three mesh resolutions exhibit consistent transient trends. Further refinement from the medium to the fine mesh did not materially alter the principal thermo-hydraulic responses, while the computational cost increased substantially. Therefore, the medium mesh with approximately 1.08 million cells was adopted for all subsequent CFD sampling, surrogate-model construction, and independent Fluent validation.

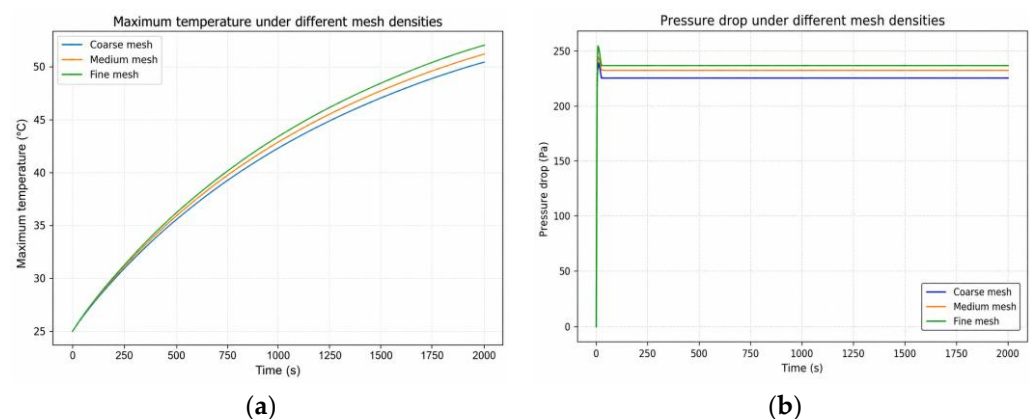


Figure A2. Mesh-independence comparison of transient thermo-hydraulic responses under different mesh resolutions: (a) maximum temperature; (b) pressure drop.

Appendix A.3. CFD Dataset for Surrogate Modeling and Sequential Refinement

The complete CFD dataset used for surrogate construction and sequential refinement consists of 21 samples. The first 15 cases were generated using the face-centered central composite design (FCCD) and were used to construct the initial V15 surrogate model. The additional E0 and A1–A5 cases were introduced sequentially for external checking and Pareto-active-region-guided refinement, resulting in the final V21 model. The dataset contains the design variables channel number, channel width, and inlet velocity, together with the corresponding thermo-hydraulic responses $T_{\max}$, $T_{\min}$, ΔT and Δp . The independent validation cases F1–F3 are not included in this dataset and were evaluated only after V21 had been frozen. The complete dataset is summarized in Table A1.

Table A1. Complete CFD dataset used for surrogate modeling and sequential refinement.

Case ID	N	w/mm	$v/(\text{m}\cdot\text{s}^{-1})$	$T_{\max}/^{\circ}\text{C}$	$T_{\min}/^{\circ}\text{C}$	$\Delta T/^{\circ}\text{C}$	$\Delta p/\text{Pa}$
C0	5	6	0.100	49.83	39.58	10.25	232.36
N3_W6_V01	3	6	0.100	50.35	40.12	10.23	148.22
N7_W6_V01	7	6	0.100	49.68	38.90	10.78	313.03
N5_W4_V01	5	4	0.100	50.78	40.82	9.96	366.69
N5_W8_V01	5	8	0.100	49.29	38.80	10.49	193.33
N5_W6_V005	5	6	0.050	51.52	42.61	8.91	103.76
N5_W6_V02	5	6	0.200	48.69	38.00	10.69	567.53
N3_W4_V005	3	4	0.050	49.17	38.86	10.31	92.22

N3_W4_V02	3	4	0.200	43.85	32.53	11.32	464.17
N3_W8_V005	3	8	0.050	51.09	41.15	9.94	53.36
N3_W8_V02	3	8	0.200	48.88	38.09	10.79	311.33
N7_W4_V005	7	4	0.050	52.30	43.23	9.07	199.52
N7_W4_V02	7	4	0.200	49.10	38.20	10.90	1011.11
N7_W8_V005	7	8	0.050	50.51	39.98	10.53	121.10
N7_W8_V02	7	8	0.200	48.01	36.77	11.24	700.44
V16_N3_W4	3	4	0.100	46.35	35.41	10.94	202.80
A1_N3_W4	3	4	0.068	47.91455	37.31805	10.5965	130.2844
A2_N5_W4	5	4	0.090	51.07815	41.03967	10.03848	320.2462
A3_N3_W4.8	3	4.8	0.055	52.81527	43.61587	9.199402	85.51375
A4_N5_W4.8	5	4.8	0.090	50.62466	40.38033	10.24432	236.1099
A5_N3_W4	3	4	0.055	48.79553	38.39984	10.39569	103.7359

These CFD samples cover the investigated design space and provide the physical basis for surrogate construction, Pareto-front identification, and decision-sensitive sequential refinement.

References

- Ho, K.-C. Advanced capacitor-based battery equalizer for underwater vehicles. *J. Mar. Sci. Eng.* **2024**, *12*, 1357. <https://doi.org/10.3390/jmse12081357>.
- Xu, H.; Zheng, R.; Yang, B.; Ning, W. Inductive wireless power transfer for autonomous underwater vehicles: A comprehensive review of technological advances and challenges. *J. Mar. Sci. Eng.* **2025**, *13*, 1855. <https://doi.org/10.3390/jmse13101855>.
- Qi, Y.; Wu, X.; Zhang, G.; Sun, Y. Energy-saving depth control of an autonomous underwater vehicle using an event-triggered sliding mode controller. *J. Mar. Sci. Eng.* **2022**, *10*, 1888. <https://doi.org/10.3390/jmse10121888>.
- Zhang, X.; Li, Z.; Luo, L.; Fan, Y.; Du, Z. A review on thermal management of lithium-ion batteries for electric vehicles. *Energy* **2022**, *238*, 121652. <https://doi.org/10.1016/j.energy.2021.121652>.
- Fu, P.; Zhao, L.; Wang, X.; Sun, J.; Xin, Z. A review of cooling technologies in lithium-ion power battery thermal management systems for new energy vehicles. *Processes* **2023**, *11*, 3450. <https://doi.org/10.3390/pr11123450>.
- Ghalkhani, M.; Habibi, S. Review of the Li-ion battery, thermal management, and AI-based battery management system for EV application. *Energies* **2023**, *16*, 185. <https://doi.org/10.3390/en16010185>.
- Tai, L.D.; Garud, K.S.; Hwang, S.G.; Lee, M.Y. A review on advanced battery thermal management systems for fast charging in electric vehicles. *Batteries* **2024**, *10*, 372. <https://doi.org/10.3390/batteries10100372>.
- Saber, N.; Richter, C.P.; Unnthorsson, R. Review of thermal management techniques for prismatic Li-ion batteries. *Energies* **2025**, *18*, 492. <https://doi.org/10.3390/en18030492>.
- Li, B.; Mao, Z.; Song, B.; Wang, X.; Tian, W.; Sun, Q.; Wang, Y.F.; Jin, Z. Experimental investigation on efficient thermal management of autonomous underwater vehicle battery packs using anisotropic expanded graphite/paraffin composite materials. *Appl. Therm. Eng.* **2024**, *242*, 122477. <https://doi.org/10.1016/j.applthermaleng.2024.122477>.
- Li, B.; Mao, Z.; Song, B.; Tian, W.; He, S.; Wang, H.; Jin, Z. A comprehensive investigation of autonomous underwater vehicle battery thermal management system using metal foam/paraffin composite. *Int. Commun. Heat Mass Transf.* **2024**, *157*, 107786. <https://doi.org/10.1016/j.icheatmasstransfer.2024.107786>.
- Wang, Z.; Liu, D.; Li, Z.; Qi, X.; Wan, C. Simulation and experimental study on heat transfer performance of bionic structure-based battery liquid cooling plate. *World Electr. Veh. J.* **2024**, *15*, 464. <https://doi.org/10.3390/wevj15100464>.
- Zhao, Y.; Zhang, T.; Wang, C.; Tang, Y.B.; Xi, H. Analysis and optimization of thermal management system for cylindrical power battery based on distributed liquid cooling plates. *Sci. China Technol. Sci.* **2024**, *67*, 3695–3706. <https://doi.org/10.1007/s11431-024-2770-1>.
- Chen, Y.; Hu, Z. Computational fluid dynamic modeling of pack-level battery thermal management systems in electric vehicles. *Energies* **2025**, *18*, 484. <https://doi.org/10.3390/en18030484>.
- Ding, Y.; Wei, M.; Liu, R. Channel parameters for the temperature distribution of a battery thermal management system with liquid cooling. *Appl. Therm. Eng.* **2021**, *186*, 116494. <https://doi.org/10.1016/j.applthermaleng.2020.116494>.
- Xu, X.; Tong, G.; Li, R. Numerical study and optimizing on cold plate splitter for lithium battery thermal management system. *Appl. Therm. Eng.* **2020**, *167*, 114787. <https://doi.org/10.1016/j.applthermaleng.2019.114787>.

16. Liu, J.; Yuan, W.; Zhou, Y.; Zhang, H. Design optimization of bionic liquid cooling plate based on PSO-BP neural network surrogate model and multi-objective genetic algorithm. *Batteries* **2025**, *11*, 141. <https://doi.org/10.3390/batteries11040141>.
17. Aydin, N.; Gurses, D.; Beyazoglu, E. Efficient immersion cooling of lithium-ion batteries: A CFD and MOGA-based optimization study. *Appl. Sci.* **2025**, *15*, 11564. <https://doi.org/10.3390/app152111564>.
18. Li, L.; Mao, Z.; Cai, T.; Hu, X.; Xu, J. Thermal-hydraulic performance optimization of Tesla-type minichannel structures for AUV battery thermal management using genetic algorithms. *Therm. Sci. Eng. Prog.* **2025**, *67*, 104130. <https://doi.org/10.1016/j.tsep.2025.104130>.
19. Li, B.; Wu, M.; Hong, L.; Li, J.; Cheng, B.; Tian, W.; Mao, Z.; Song, B. Investigation on thermal management of AUVs battery packs via an innovative bionic ram liquid cooling design. *Appl. Therm. Eng.* **2026**, *302*, 131890. <https://doi.org/10.1016/j.applthermaleng.2026.131890>.
20. Qian, W.; Fang, W.; Tian, Y.; Dai, G.; Yan, T.; Yang, S.; Wang, P. Data-driven prediction of Li-ion battery thermal behavior: Advances and applications in thermal management. *Processes* **2025**, *13*, 2769. <https://doi.org/10.3390/pr13092769>.
21. Ye, L.; Li, C.; Wang, C.; Zheng, J.; Zhong, K.; Wu, T. A multi-objective optimization approach for battery thermal management system based on the combination of BP neural network prediction and NSGA-II algorithm. *J. Energy Storage* **2024**, *99*, 113212. <https://doi.org/10.1016/j.est.2024.113212>.
22. Cho, H.; Lee, H.; Ryu, S. Sequential multi-objective Bayesian optimization of air-cooled battery thermal management system with spoiler integration. *J. Energy Storage* **2025**, *114*, 115586. <https://doi.org/10.1016/j.est.2025.115586>.
23. Tan, H.; Wu, H.; Zhang, Q.; Lei, G.; Chen, Q. Surrogate-assisted multi-objective optimization of a liquid oxygen vacuum sub-cooling system based on ejector and liquid ring pump. *Processes* **2022**, *10*, 1188. <https://doi.org/10.3390/pr10061188>.
24. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. <https://doi.org/10.1109/4235.996017>.
25. Wang, G.G.; Shan, S. Review of metamodeling techniques in support of engineering design optimization. *J. Mech. Des.* **2007**, *129*, 370–380. <https://doi.org/10.1115/1.2429697>.
26. Jones, D.R.; Schonlau, M.; Welch, W.J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **1998**, *13*, 455–492. <https://doi.org/10.1023/A:1008306431147>.
27. Knowles, J. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Trans. Evol. Comput.* **2006**, *10*, 50–66. <https://doi.org/10.1109/TEVC.2005.851274>.
28. Picheny, V. Multiobjective optimization using Gaussian process emulators via stepwise uncertainty reduction. *Stat. Comput.* **2015**, *25*, 1265–1280. <https://doi.org/10.1007/s11222-014-9477-x>.
29. Couckuyt, I.; Deschrijver, D.; Dhaene, T. Fast calculation of multiobjective probability of improvement and expected improvement criteria for Pareto optimization. *J. Glob. Optim.* **2014**, *60*, 575–594. <https://doi.org/10.1007/s10898-013-0118-2>.
30. Coello Coello, C.A.; Toscano Pulido, G.; Salazar Lechuga, M. Handling multiple objectives with particle swarm optimization. *IEEE Trans. Evol. Comput.* **2004**, *8*, 256–279. <https://doi.org/10.1109/TEVC.2004.826067>.
31. Zitzler, E.; Thiele, L.; Laumanns, M.; Fonseca, C.M.; da Fonseca, V.G. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Trans. Evol. Comput.* **2003**, *7*, 117–132. <https://doi.org/10.1109/TEVC.2003.810758>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.