Improving the performance of Surrogate-Based Optimization methods

Gary Goldblatt

February - July 2026

Contents

Abstract

Surrogate-based optimization is commonly used when the evaluation of objective functions and constraints is computationally expensive. In this work, several surrogate models are considered, including Tuned Radial Basis Function and Kriging models. Since their performance varies depending on the problem, different model-selection and aggregation strategies are investigated to improve robustness and optimization performance. The proposed approaches are evaluated through external validation and surrogate-based optimization experiments on a heterogeneous benchmark set. The results show that adaptive selection improves robustness compared with fixed surrogate models, while dynamic aggregation provides the best overall optimization performance. For high-dimensional problems, TRBF-based selection and aggregation strategies offer the most favourable compromise between performance and computational cost.

1 Introduction

The present work is conducted within the software environment introduced in [8]. In particular, the developments are implemented in Minamo, Cenaero's in-house platform dedicated to design space exploration and multidisciplinary optimization Many engineering optimization problems involve objective functions and constraints that are computationally expensive to evaluate, especially when high-fidelity numerical simulations are required. In such cases, directly using the true functions within an optimization process can become extremly costly. Surrogate-based optimization (SBO) addresses this issue by replacing expensive evaluations with surrogate models that provide fast approximations of the true functions. Among the most commonly used surrogate models are Radial Basis Function (RBF) models and Kriging models. However, the performance of a surrogate model strongly depends on the considered problem, the available data, and the stage of the optimization process. As a result, no single surrogate model consistently performs best in all situations. To overcome this issue, model selection and aggregation strategies can be used [6]. Selection approaches aim to identify the most suitable surrogate model, while aggregation approaches combine multiple surrogate predictions in order to improve robustness and predictive performance. The objective of this work is to investigate different model selection strategies, and aggregation approaches for surrogate-based optimization. Particular attention is given to weighting strategies based on surrogate predictive quality. Several quality metrics are studied, with a focus on the combined $r2\_rp$ criterion, which evaluates both prediction accuracy and preservation of the ranking of candidate solutions. External validation experiments are first performed to identify the most promising surrogate configurations. The selected approaches are then evaluated within surrogate-based optimization procedures in order to analyze their optimization performance and computational cost.

2 Quality Metrics for Surrogate Model Evaluation

In order to compare surrogate models and guide the selection or aggregation process, appropriate quality metrics are required to evaluate predictive performance. Different metrics can be used depending on the objective of the evaluation, such as prediction accuracy, error estimation, or preservation of the relative ordering of candidate solutions. Since optimization problems are particularly sensitive to the identification of promising candidates, ranking-based metrics are of particular interest in this work. The quality metrics considered in this study are presented in this section [4] [9]. In this work, surrogate quality is evaluated using a combined criterion based on both the R^2 coefficient and the Ranking Preservation (RP) metric. The objective of this combined criterion is to simultaneously account for global prediction accuracy and preservation of the ranking of candidate solutions. This is particularly relevant in surrogate-based optimization, where accurately identifying promising candidates is often as important as predicting exact objective values.

2.1 R^2 Coefficient

The coefficient of determination, R^2, measures how well the surrogate predictions $\hat{y}_i$ reproduce the variability of the true responses y_i. A value close to one indicates accurate predictions, while lower values correspond to poorer predictive performance. It is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{1}$$

where n denotes the number of samples, y_i the true responses, $\hat{y}_i$ the predicted responses, and $\bar{y}$ the mean value of the true responses. Although R^2 provides a global measure of prediction accuracy, it does not evaluate the preservation of the ranking of candidate solutions, which is particularly important in optimization problems.

2.2 Ranking Preservation Coefficient

The Ranking Preservation (RP) coefficient evaluates how well a surrogate model preserves the relative ordering of candidate solutions. Unlike error-based metrics, RP does not focus on the magnitude of prediction errors, but more on the consistency between the rankings of the predicted and true responses. It is based on pairwise comparisons between samples and measures the proportion of correctly preserved orderings:

$$\mathrm{RP} = \frac{1}{\frac{1}{2}n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} h(i,j), \tag{2}$$

where n denotes the number of samples and $h(i, j)$ indicates whether the relative ordering between samples i and j is correctly preserved. The RP coefficient takes values in the interval $[0, 1]$, where values close to one indicate a strong preservation of the ranking of candidate solutions. This metric is particularly relevant in optimization problems, where correctly identifying the best candidates is often more important than predicting exact objective values.

3 Model Selection and Aggregation Methods

3.1 Models used

3.1.1 Radial Basis Function Models

Radial Basis Function (RBF) models are interpolation-based surrogate models commonly used to approximate expensive objective functions [3]. Given a set of training samples, the approximation is constructed as a weighted sum of radial basis functions centered on the training points:

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i \, \phi(\|\mathbf{x} - \mathbf{x}_i\|), \tag{3}$$

where $\phi(\cdot)$ denotes a radial basis kernel function, λ_i are interpolation coefficients, and $\|\mathbf{x} - \mathbf{x}_i\|$ represents the Euclidean distance between the prediction point and the training samples. In this work, Tuned Radial Basis Function (TRBF) models are used. In contrast to standard RBF models relying on a fixed configuration, TRBF models adjust selected model parameters during the surrogate construction process in order to better adapt the approximation to the available data. The objective of this tuning procedure is to improve predictive quality while preserving the relatively low computational cost of RBF-based surrogates. Since TRBF is an internal Cenaero methodology, only its general principle is presented in this report. The behavior of the surrogate model depends on the selected kernel function. In this work, several kernels were investigated, including Gaussian, Exponential, Wave, Multiquadric, and Rational Quadratic kernels, in order to obtain surrogate models with different smoothness and generalization properties [3].

3.1.2 Kriging Models

Kriging models are probabilistic surrogate models based on Gaussian processes [10], [2]. Unlike RBF models, Kriging not only provides a prediction of the response, but also an estimation of the prediction uncertainty. The response is modeled as:

$$y(\mathbf{x}) = \mu + Z(\mathbf{x}), \tag{4}$$

where μ is a constant mean and $Z(\mathbf{x})$ is a Gaussian process defined through a correlation kernel. Several kernel formulations were investigated in this work, including Gaussian, Matern, Exponential, Fully Parameterized, and Gaussian

KPLS kernels. These kernels introduce different assumptions regarding the smoothness and regularity of the response surface. Different hyperparameter optimization strategies were also considered for the Kriging models; BFGS, Genetic Algorithm (GA), and Hybrid approaches. BFGS provides fast local optimization, GA allows broader global exploration at a higher computational cost, while the Hybrid approach combines both strategies.

3.2 Selection Strategies

Model selection aims to identify the surrogate model providing the best predictive performance among a set of candidate models [5]. Instead of relying on a single predefined surrogate, several surrogate models are evaluated and the one maximizing a chosen quality criterion is selected. Let $\{SM_1, SM_2, \ldots, SM_n\}$ denote a set of candidate surrogate models associated with quality coefficients q_i. The selected model is defined as:

$$SM_{selected} = \arg \max_{i \in \{1,\ldots,n\}} q_i \tag{5}$$

The selection is performed independently for each response of the optimization problem. This adaptive strategy allows the surrogate model to better match the characteristics of the considered response function.

3.3 Aggregation Strategies

While selection strategies choose a single surrogate model, aggregation strategies combine the predictions of several surrogate models [5]. The objective is to exploit the complementarity of different surrogate formulations in order to improve robustness and predictive stability. Let $\{SM_1, SM_2, \ldots, SM_n\}$ denote a set of surrogate models producing predictions $\hat{y}_i(\mathbf{x})$. The aggregated prediction is computed as:

$$\hat{y}_{agg}(\mathbf{x}) = \sum_{i=1}^{n} w_i \, \hat{y}_i(\mathbf{x}), \tag{6}$$

where w_i denotes the weight associated with surrogate model SM_i, with:

$$\sum_{i=1}^{n} w_i = 1 \quad \text{and} \quad w_i \geq 0. \tag{7}$$

The weights determine the contribution of each surrogate model in the final prediction and are typically based on surrogate predictive quality.

4 Proposed Weighting Strategies

Several weighting strategies were investigated in order to assign importance weights to surrogate models based on their predictive quality. The objective is to favor surrogate models providing reliable predictions while reducing the influence of poorly performing models. In this work, surrogate quality is evaluated using a combined criterion based on both prediction accuracy and ranking preservation. Models associated with higher quality values receive larger weights in the aggregation process.

4.1 Default Aggregation Strategy

A default aggregation strategy already implemented in the MINAMO framework was used as a reference approach for comparison with the newly investigated weighting strategies. This baseline strategy relies on quality-based weighting combined with a minimum quality threshold. Surrogate models whose predictive quality is considered insufficient are excluded from the aggregation process, while the remaining models receive normalized weights proportional to their estimated quality. A relaxation mechanism is also included in order to avoid situations where no surrogate model satisfies the required quality level. This approach provides a robust reference aggregation strategy and serves as the baseline for evaluating the new weighting methods proposed in this work.

4.2 Incremental Threshold Relaxation Weighting

An alternative aggregation strategy based on incremental threshold relaxation was investigated. Unlike the baseline strategy relying on multiplicative relaxation, this approach progressively decreases the minimum quality threshold using fixed steps. The objective is to avoid excluding surrogate models whose quality is only slightly below the initial threshold. At each iteration, surrogate models satisfying the current quality threshold are retained, while the others are discarded. The weights are then computed from the normalized quality coefficients of the remaining models. This strategy provides finer control over the selection process and aims to preserve potentially useful surrogate models that could be rejected too early by larger relaxation steps.

4.3 Dynamic Adaptive Weighting

A dynamic adaptive weighting strategy was investigated in order to continuously adjust the contribution of each surrogate model according to its predictive quality [6]. Unlike threshold-based approaches, this strategy does not explicitly discard models, but instead progressively reduces the influence of poorly performing surrogate models. The quality coefficient $q_i \in [0, 1]$ associated with surrogate model SM_i is first transformed into an error indicator:

$$e_i = 1 - q_i \tag{8}$$

where:

- q_i denotes the quality coefficient of surrogate model SM_i,
- e_i represents the associated prediction error indicator.

Aggregation weights are then computed using an inverse-error formulation:

$$w_i = \left(\frac{1}{e_i + \varepsilon} \right)^{\beta} \tag{9}$$

where:

- w_i is the weight assigned to surrogate model SM_i,
- $\varepsilon > 0$ is a small numerical tolerance preventing division by zero,
- $\beta > 0$ is a selectivity parameter controlling the aggressiveness of the weighting strategy.

Small values of β produce smoother and more balanced aggregations, while larger values progressively favor the best-performing surrogate model, approaching a winner-takes-all behavior. This strategy allows continuous and adaptive weighting of surrogate models while remaining robust to poorly performing configurations.

5 Experimental Setup

The experimental study compares the considered surrogate-model selection and aggregation strategies in the context of surrogate-based optimization. Different surrogate families, kernel formulations, and weighting approaches are evaluated on a heterogeneous benchmark set.

Each test case is repeated over 100 independent runs to obtain statistically reliable performance estimates. The size of the initial Design of Experiments (DOE) is defined proportionally to the problem dimension n. For external validation, the DOE contains $5n$ sample points, while it is set to $3n$ for optimization. The total evaluation budget is fixed to 200 function evaluations, resulting in $200 -$ DOE size sequential optimization iterations.

5.1 Benchmark Test Cases

The benchmark contains 24 test cases with different dimensions, variable types, numbers of constraints, and response-surface characteristics. It includes low- and high-dimensional problems, constrained and unconstrained cases, continuous and mixed-variable formulations, as well as functions exhibiting different levels of non-linearity, multimodality, and oscillatory behaviour.

This diversity is important for assessing whether the investigated methods remain robust across different problem structures rather than performing well only on a specific type of response surface. Table 1 summarizes the benchmark problems used in the study.

Test case	Dimension	# constraints
wing-grouping-frame-lightmaterial	9 (7 real, 2 categorical)	3
t10b-nogrouping-truss-lightmaterial	33 (13 real, 20 categorical)	3
t10b-grouping-truss-lightmaterial	6 (4 real, 2 categorical)	3
WB4	4	6
Spring-RI	3 (2 real, 1 integer)	4
Schwefel_20	20	0
Rosenbrock_uncomputable_10	10	0
Rosenbrock_30	30	0
ReinforcedConcreteBeam-RID	3 (1 real, 1 integer, 1 discrete)	2
Rastrigin_20	20	0
Hesse	6	6
Hart6	6	0
G8	2	2
G7	10	8
G6	2	2
G3MOD_plog_20	20	1
G2_plog_10	10	2
G19	15	5
G18	9	13
G16	5	38
G10	8	6
G1	13	9
CarSideImpact-RD	11 (9 real, 2 discrete)	10
Ackley_20	20	0

Table 1: Summary of the 24 benchmark test cases used in the study.

5.2 External Validation and Optimization Objectives

External validation is first used to assess the predictive reliability of the surrogate-model configurations before they are introduced into the optimization loop. This stage acts as a filtering procedure, allowing unstable or poorly performing models to be discarded and reducing the computational cost of the subsequent optimization experiments.

Predictive performance is evaluated using the quality criteria introduced previously, with particular attention given to the combined *r2_rp* criterion. This criterion accounts for both prediction accuracy and the preservation of the ranking of candidate solutions, which is particularly relevant in an optimization context.

The optimization phase then evaluates the practical impact of the retained surrogate models and adaptive strategies within the Surrogate-Based Optimization (SBO) process. A model with good predictive quality does not necessarily provide the best optimization performance, since the surrogate must also guide the search efficiently towards promising regions of the design space.

The optimization experiments therefore compare the methods in terms of convergence behaviour, final objective value, robustness across independent runs, and computational cost. All strategies are evaluated using the same optimization framework and evaluation budget to ensure a fair comparison.

5.3 Surrogate Model Configurations

Several surrogate-model configurations were evaluated during external validation.

For the Kriging family, the investigated correlation kernels include Gaussian, Matern, Exponential, Fully Parameterized, and Gaussian KPLS formulations. Different hyperparameter optimization strategies were also considered, namely BFGS, Genetic Algorithm (GA), and Hybrid approaches.

For the TRBF family, the considered kernels include Gaussian, Multiquadric, Rational Quadratic, Wave, and Exponential formulations. These configurations provide surrogate models with different smoothness assumptions and approximation behaviours.

The purpose of these experiments is to evaluate the influence of the kernel formulation and hyperparameter estimation strategy on predictive quality and optimization performance.

5.4 Selection Strategies

Several adaptive selection strategies were investigated to determine whether selecting a surrogate model according to its estimated quality improves performance compared with using a fixed configuration.

Three selection levels were considered:

- **Selection within the TRBF family:** the best TRBF configuration is selected among the candidate kernel formulations independently for each response.
- **Selection within the Kriging family:** the best Kriging configuration is selected among the considered correlation kernels and hyperparameter optimization strategies.
- **Global selection:** the best surrogate is selected from the complete set of Kriging and TRBF candidates.

The selection process relies on the surrogate quality criterion introduced previously and is performed independently for each response of each benchmark problem.

5.5 Aggregation Strategies

Aggregation strategies were also evaluated to determine whether combining surrogate models from different families improves predictive robustness and optimization performance.

For each response, one surrogate model is first selected within the TRBF family and one within the Kriging family. Their predictions are then combined as

$$\hat{y}_{\text{agg}}(\mathbf{x}) = w_{\text{TRBF}}\,\hat{y}_{\text{TRBF}}(\mathbf{x}) + w_{\text{OKG}}\,\hat{y}_{\text{OKG}}(\mathbf{x}), \tag{10}$$

where w_{TRBF} and w_{OKG} denote the weights assigned to the selected surrogate models.

Different weighting strategies are investigated to evaluate how the distribution of the weights affects predictive quality, optimization robustness, and computational cost.

6 Results

The external validation and optimization experiments are both presented in this section. External validation is first used as a preliminary filtering step to identify the most promising surrogate model configurations and define the baseline models retained for the optimization study.

The optimization results then compare fixed surrogate baselines, selection strategies, and aggregation strategies within the surrogate-based optimization framework. This makes it possible to assess whether the predictive trends observed during external validation are also reflected in the actual optimization performance.

6.1 Baseline Surrogate Models: External Validation

6.1.1 Kriging Baselines

Several hyperparameter optimization scenarios were first compared for the OKG surrogate models, including BFGS, Genetic Algorithm (GA), and Hybrid approaches. Their predictive performance was found to be globally similar, while their computational costs differed significantly. Since BFGS achieved comparable predictive quality at a substantially lower cost, it was retained for the subsequent kernel comparison.

The external validation analysis therefore focuses on the Gaussian, Matern, Exponential, and Fully Parameterized OKG kernels trained using BFGS scenario. Figure 1 presents the distribution of the combined quality metric $r2\_rp$ across all responses and benchmark test cases.

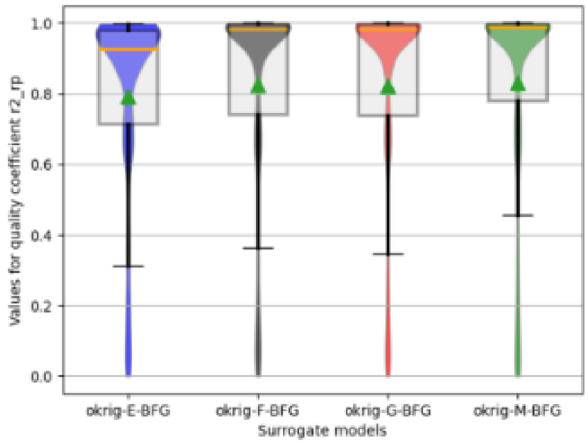

Figure 1: Distribution of the $r2\_rp$ quality metric for the different OKG kernels using the BFGS scenario across all responses and benchmark test cases.

The Matern and Gaussian kernels provide the strongest and most robust predictive performance. The Fully Parameterized kernel remains competitive but exhibits greater variability, whereas the Exponential kernel generally performs less well.

Figure 2 compares the normalized execution time of the same configurations. The Fully Parameterized kernel is clearly the most computationally expensive, while the Gaussian and Matern kernels provide a better compromise between predictive quality and computational cost.

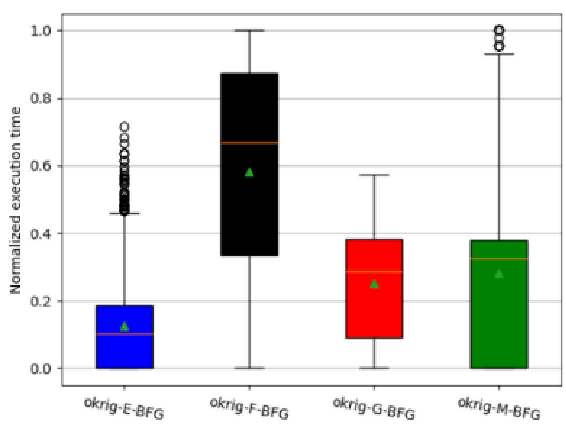

Figure 2: Normalized execution-time distribution for the different OKG kernels using the BFGS scenario across all responses and benchmark test cases.

Based on these results, the Gaussian and Matern kernels are retained for the subsequent selection and aggregation experiments.

6.1.2 TRBF Baselines

The external validation phase was used to compare the predictive quality and computational cost of the considered TRBF configurations. The investigated models include Gaussian, Multiquadric, Rational Quadratic, Wave, and the default TRBF configuration, which relies on a predefined sequence of kernels.

Figure 3 presents the distribution of the combined quality metric $r2\_rp$ across all responses and benchmark test cases.

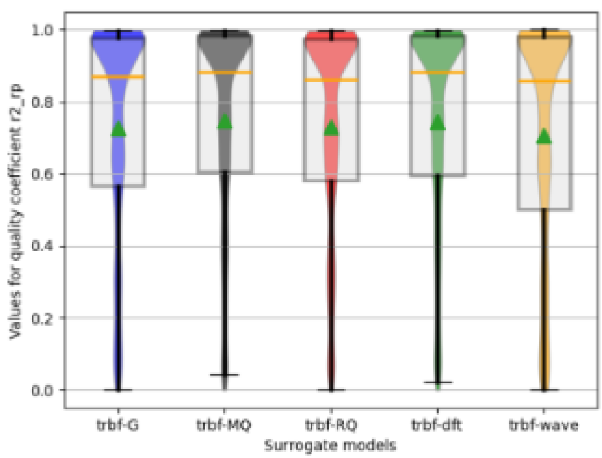

Figure 3: Distribution of the $r2\_rp$ quality metric for the different TRBF configurations across all responses and benchmark test cases.

The Multiquadric and default TRBF configurations provide the strongest overall predictive performance, with relatively high and stable values of the quality criterion. The Rational Quadratic configuration also shows competitive behaviour. In contrast, the Gaussian and Wave kernels exhibit greater variability, indicating that their predictive performance is more dependent on the characteristics of the considered response.

Figure 4 compares the normalized execution time of the same configurations.

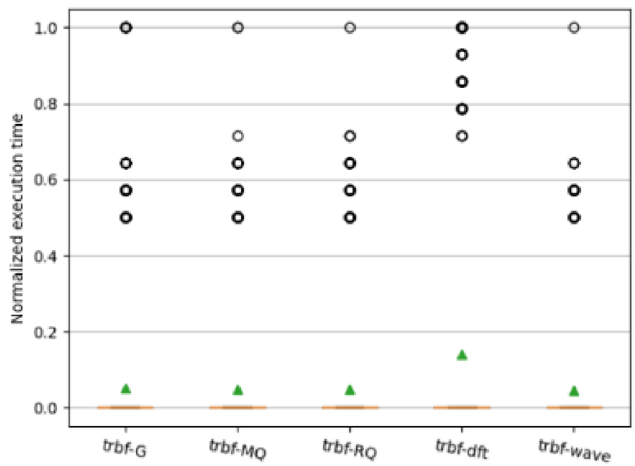

Figure 4: Normalized execution-time distribution for the different TRBF configurations across all responses and benchmark test cases.

The Gaussian, Multiquadric, Rational Quadratic, and Wave configurations have broadly similar computational costs, while the default TRBF configuration is slightly more expensive because it evaluates a predefined sequence of kernels.

Based on these results, the Multiquadric, Rational Quadratic, Wave, and default TRBF configurations are retained for the subsequent selection and aggregation experiments. Although the Wave kernel shows greater variability in external validation, it can still provide strong performance on specific responses and therefore remains relevant in an adaptive selection framework.

6.2 Baseline Surrogate Models: Optimization

6.2.1 Kriging Baselines

The Kriging baselines are evaluated in order to provide reference optimization performances before introducing the selection and aggregation strategies. The comparison focuses on the main kernel configurations retained after the preliminary validation phase. Figure 5 presents the data profile obtained for the lower quartile objective value. A data profile represents the proportion of optimization problems successfully solved as a function of the evaluation budget [7]. Curves located higher on the graph therefore indicate methods capable of solving a larger number of problems for a given number of objective function evaluations. The Matern and Gaussian kernels show very similar optimization behavior and provide the most robust performances among the tested Kriging configurations.

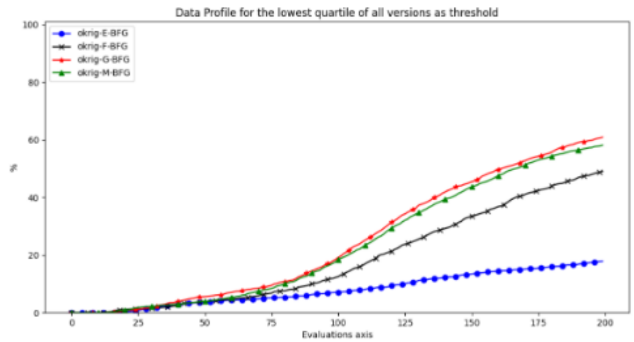

Figure 5: Data profile of percentage of solved problem at each evaluation for lowest quartile as threshold for the different OKG kernels.

The fully parameterized kernel also provides competitive optimization results in several cases. However, its significantly higher computational cost does not lead to sufficient performance improvements to justify its systematic use within the optimization framework. The exponential kernel generally exhibits lower overall performance, but remains competitive for a limited number of test cases, particularly for problems characterized by highly oscillatory response surfaces. Overall, the Matern and Gaussian Kriging models are retained as the main Kriging baselines for the comparison with the adaptive selection and aggregation strategies.

6.2.2 TRBF Baselines

The optimization results obtained with the different TRBF kernels show differences in robustness and convergence behavior across the benchmark problems. In the Minamo optimization platform, these models are referred to as Tuned Radial Basis Functions (TRBF). The term *tuned* refers to the automatic adaptation of the kernel hyperparameters during the surrogate model construction process. Figure 6 presents the data profile obtained for the lower quartile objective value. The Multiquadric (MQ), Rational Quadratic (RQ), and Wave kernels consistently provide the best optimization performances among the considered TRBF configurations.

The MQ kernel demonstrates stable and robust behavior across most test cases. The RQ kernel also provides competitive results, benefiting from its ability to adapt to multiple characteristic length scales of the response surface. The Wave kernel performs particularly well on several difficult and highly nonlinear problems due to its oscillatory formulation.

The Gaussian kernel shows competitive optimization behavior but does not provide significant advantages compared to MQ.

Based on these observations, the MQ, RQ, and Wave kernels are retained as the main TRBF baselines for the comparison with the adaptive selection and aggregation strategies.

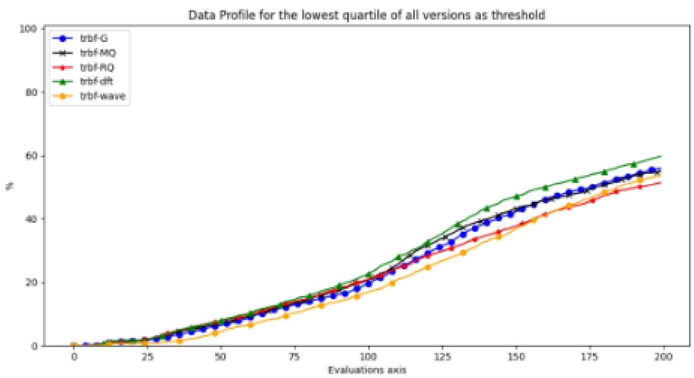

Figure 6: Data profile of percentage of solved problems at each evaluation for the lower quartile objective threshold for the different TRBF kernels.

6.3 Selection Strategies

6.3.1 External Validation Results

The external validation study evaluates whether adaptive surrogate selection improves predictive quality and robustness compared with fixed surrogate configurations. Three strategies are considered: selection within the O-Kriging family, selection within the TRBF family, and extra-familial selection combining models from both surrogate families.

The O-Kriging selection strategy considers the Gaussian, Matern, Exponential, and Fully Parameterized kernels, all trained using the BFGS hyperparameter optimization procedure. Figure 7 presents the distribution of the $r2\_rp$ quality metric obtained across all responses and benchmark test cases.

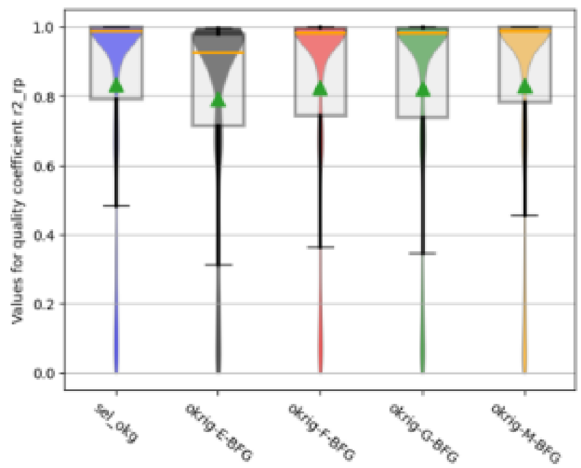

Figure 7: Distribution of the $r2\_rp$ quality metric obtained with the O-Kriging selection strategy during external validation.

The selection strategy achieves high predictive quality, with most values concentrated close to 1. Its selection frequencies are reported in Table 2.

Model	Selection frequency
O-Kriging Matern	**64.3%**
O-Kriging Gaussian	18.5%
O-Kriging Fully Parameterized	13.9%
O-Kriging Exponential	3.3%

Table 2: Selection frequency of the O-Kriging surrogate models during external validation.

The Matern kernel is selected most frequently, which is consistent with its strong individual predictive performance. However, the strategy introduces a significant computational overhead because all candidate models must be trained and evaluated before the final model is selected.

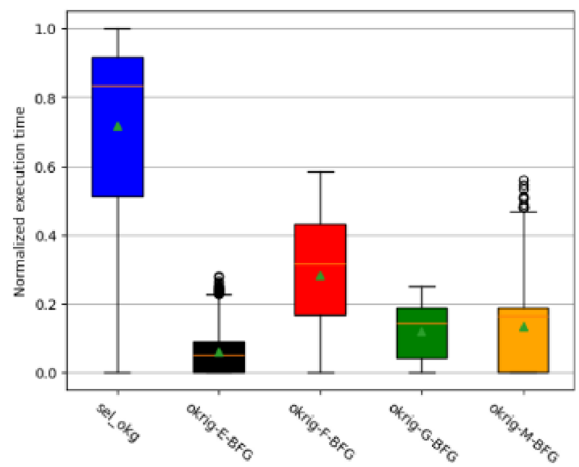

Figure 8: Normalized execution-time distribution for the O-Kriging selection strategy during external validation.

The TRBF selection strategy evaluates the default TRBF sequence together with the Gaussian, Multiquadric, Rational Quadratic, and Wave kernels. Figure 9 presents its predictive performance.

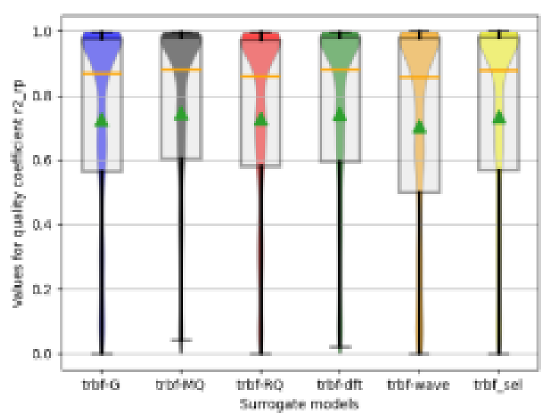

Figure 9: Distribution of the $r2\_rp$ quality metric obtained with the TRBF selection strategy during external validation.

The TRBF selection strategy achieves predictive quality comparable to the best fixed TRBF configurations, although no clear improvement over the strongest individual models is observed. Table 3 shows that the default sequence and the Wave kernel dominate the selection process.

Model	Selection frequency
TRBF default	**42.6%**
TRBF Wave	40.4%
TRBF Rational Quadratic	9.1%
TRBF Gaussian	7.0%
TRBF Multiquadric	0.8%

Table 3: Selection frequency of the TRBF surrogate models during external validation.

To better evaluate the individual kernels, the analysis was repeated after removing the default TRBF sequence. The resulting frequencies are presented in Table 4.

Model	Selection frequency
TRBF Wave	**40.4%**
TRBF Multiquadric	38.9%
TRBF Gaussian	13.7%
TRBF Rational Quadratic	7.0%

Table 4: Selection frequency of the TRBF kernels after removing the default sequence from the candidate set.

Without the default sequence, the Wave and Multiquadric kernels together account for nearly 80% of all selections, confirming their complementary predictive behaviour. As for O-Kriging selection, the TRBF selection strategy increases computational cost because several candidate models must be evaluated.

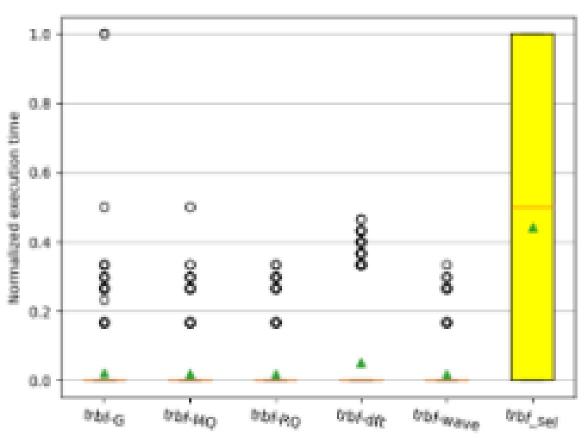

Figure 10: Normalized execution-time distribution for the TRBF selection strategy during external validation.

Finally, the extra-familial strategy selects between the Gaussian and Matern O-Kriging models and the Multiquadric and Wave TRBF models. Figure 11 presents the resulting predictive quality.

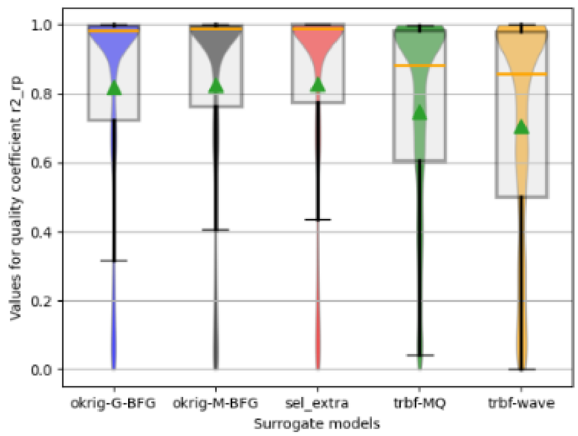

Figure 11: Distribution of the $r2\_rp$ quality metric obtained with the extra-familial selection strategy during external validation.

The distribution is strongly concentrated near 1, indicating high predictive accuracy and robust ranking preservation. The corresponding model-selection frequencies are given in Table 5.

Model	Selection frequency
O-Kriging Matern	**69.5%**
O-Kriging Gaussian	23.5%
TRBF Wave	4.2%
TRBF Multiquadric	2.8%

Table 5: Selection frequency of the surrogate models in the extra-familial strategy during external validation.

The Matern model clearly dominates the selection process, showing that the O-Kriging family provides the most consistent predictive quality on this benchmark. Nevertheless, the TRBF models remain useful for responses where different approximation properties are beneficial.

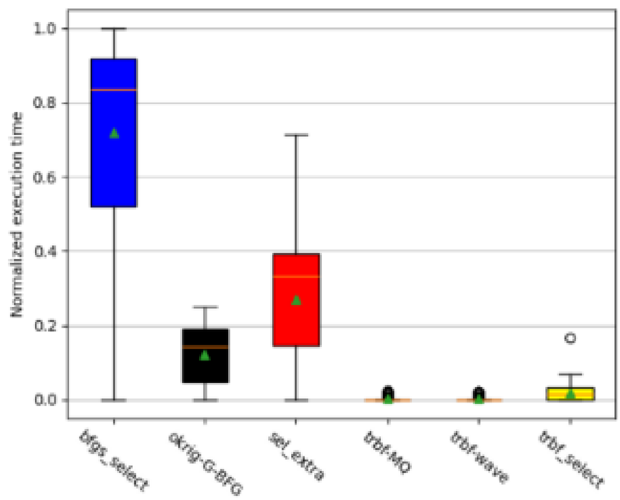

Figure 12: Normalized execution-time distribution for the extra-familial selection strategy during external validation.

Overall, the external validation results show that adaptive selection can improve predictive robustness by choosing the most suitable surrogate for each response. The O-Kriging and extra-familial strategies provide the strongest predictive quality, while the TRBF strategy remains competitive but does not clearly outperform the best fixed TRBF configurations. In every case, this increased adaptability introduces additional computational cost.

6.3.2 Optimization Results

The objective of the selection strategies is not necessarily to outperform every individual surrogate model on each optimization problem, but rather to improve robustness across a wide range of problem characteristics by dynamically selecting the most suitable surrogate configuration.

Three selection strategies were investigated. The `sel-okg` strategy performs a selection between the Gaussian and Matern O-Kriging models, both optimized using the BFGS hyperparameter estimation procedure. The `trbf-sel` strategy performs a selection between the three retained TRBF kernels: Multiquadric (MQ), Rational Quadratic (RQ) and Wave. Finally, the `sel-extra` strategy extends the selection process across surrogate families by considering both the TRBF kernels (MQ, RQ and Wave) and the O-Kriging models (Gaussian and Matern), thereby increasing the diversity of candidate surrogate models. Figure 13 presents the data profile obtained for the lower quartile objective value. The results compare the fixed surrogate baselines with the different adaptive selection strategies.

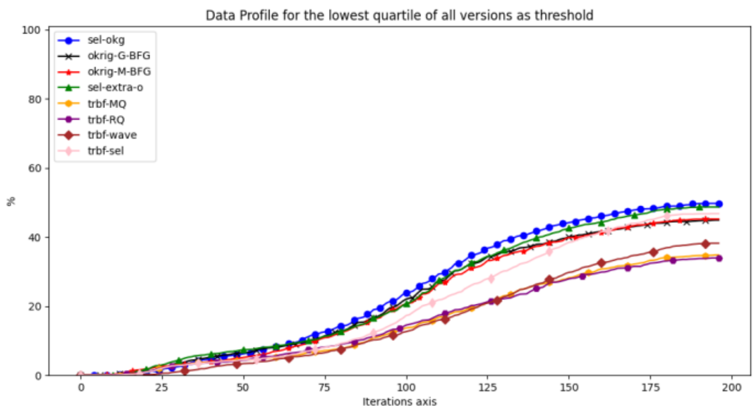

Figure 13: Data profile comparison of the different selection strategies using the lower quartile objective threshold.

Kernel	Selection frequency
Matern	**68.2%**
Gaussian	31.8%

Table 6: Selection frequency of each kernel in the OKG selection strategy.

TRBF kernel	Selection frequency
Multi-Quadric	**61.6%**
Wave	28.1%
Rational-Quadratic	10.3%

Table 7: Selection frequency of each kernel in the TRBF selection strategy.

Surrogate model	Selection frequency
Kriging Matern	**43.1%**
TRBF Multi-Quadric	24.4%
Kriging Gaussian	21.7%
TRBF Wave	7.5%
TRBF Rational-Quadratic	3.4%

Table 8: Selection frequency of each surrogate model in the extra-familial selection strategy.

Several observations can be drawn from these selection frequencies. For the OKG-based selection strategy, the Matern kernel is selected more than twice as often as the Gaussian kernel, suggesting that its increased flexibility is generally beneficial across the considered benchmark problems.

For the TRBF selection strategy, the Multi-Quadratic kernel clearly dominates the selection process, accounting for more than 60% of all decisions. The Wave kernel is selected less frequently but still contributes significantly, while the Rational-Quadratic kernel is only occasionally retained. These results are consistent with the optimization performances observed previously, where the Multi-Quadratic kernel emerged as one of the most robust TRBF configurations.

The extra-familial selection strategy provides additional insight into the relative competitiveness of the different surrogate families. The Matern Kriging model remains the most frequently selected surrogate, but the Multi-Quadratic TRBF kernel is selected nearly one quarter of the time. Together, the Matern and Gaussian Kriging models account for approximately 65% of all selections, while the TRBF models account for the remaining 35%. This distribution indicates that both surrogate families contribute to the overall performance of the selection strategy and confirms the complementarity between Kriging and TRBF models. Such diversity explains the strong robustness observed for the extra-familial selection strategy in Figure 13.

The results show that the OKG-based selection and the extra-familial selection provide the most robust optimization performances across the benchmark set. Both strategies consistently outperform the individual TRBF kernels and remain competitive with the best fixed Kriging configurations. The TRBF selection also improves robustness compared to individual TRBF kernels by combining complementary modeling assumptions. In particular, the MQ, RQ, and Wave kernels provide different approximation behaviors, allowing the selection strategy to adapt to both smooth and highly non-linear optimization problems. The extra-familial selection further extends this idea by allowing the optimization process to select surrogate models across both Kriging and TRBF families. This approach increases modeling diversity and reduces the risk of poor surrogate assumptions when the characteristics of the optimization problem are not known in advance. Although the OKG selection sometimes achieves slightly better optimization performance, the extra-familial strategy provides a better compromise between robustness and computational cost since it avoids the full computational overhead associated with the BFGS-based Kriging selection. Overall, the results confirm that adaptive selection strategies improve the robustness of surrogate-based optimization compared to relying on a single fixed surrogate model.

6.4 Aggregation Strategies

6.4.1 External Validation Results

The external validation study compares three aggregation approaches: the default, relaxed, and dynamic aggregation strategies. Each strategy combines two independently selected surrogate models: one selected from the O-Kriging family and one selected from the TRBF family. For the O-Kriging family, the considered kernels are the Gaussian, Matern, fully-parameterized and exponential one (again both in the BGFS scenario), and for the TRBF family, the candidate set contains the default TRBF sequence, the Gaussian sequence, the Multiquadric sequence, the Rational-Quadratic sequence and the Wave sequence. Their predictions are then combined using the weighting mechanism associated with the considered aggregation strategy.

Default aggregation Figure 14 presents the predictive quality obtained with the default aggregation strategy.

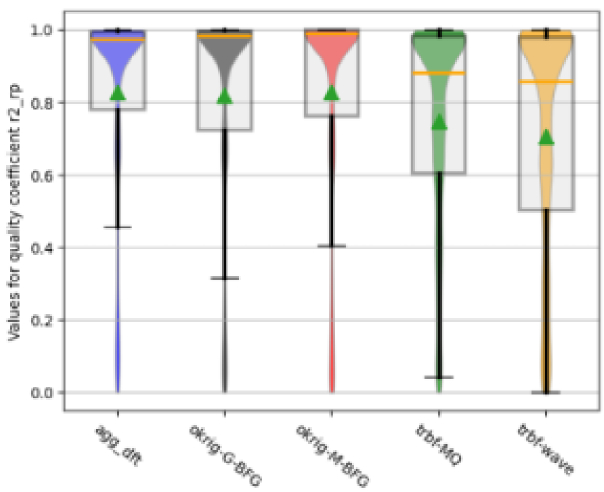

Figure 14: Distribution of the $r2\_rp$ quality metric obtained with the default aggregation strategy compared with individual surrogate models.

The default aggregation strategy achieves predictive quality comparable to the best individual surrogate models. Its $r2\_rp$ distribution is strongly concentrated near 1 and exhibits limited variability, indicating that combining models from different families improves predictive stability.

This increased robustness comes with a higher computational cost because several candidate models must be trained and evaluated before constructing the aggregated prediction.

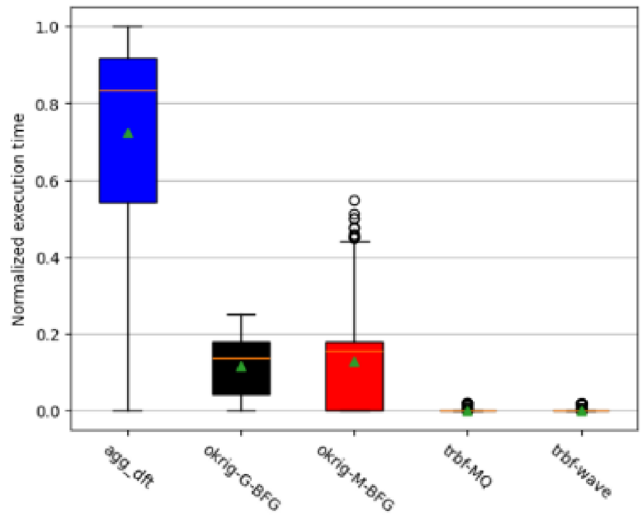

Figure 15: Normalized execution-time distribution for the default aggregation strategy.

Relaxed aggregation Figure 16 presents the predictive quality obtained with the relaxed aggregation strategy.

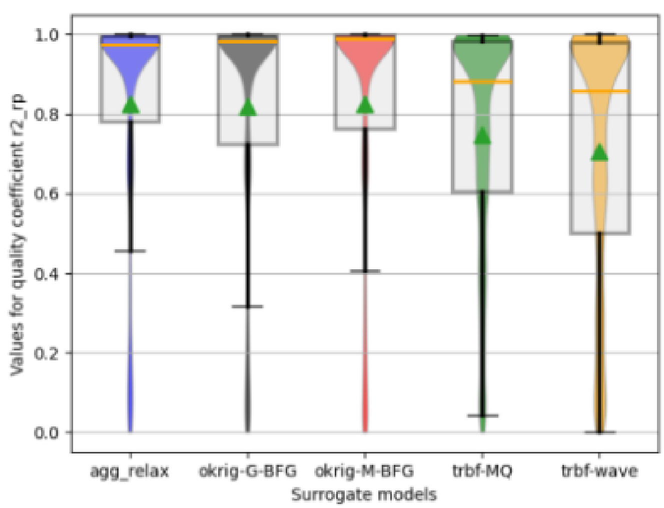

Figure 16: Distribution of the $r2\_rp$ quality metric obtained with the relaxed aggregation strategy compared with individual surrogate models.

The relaxed strategy also provides predictive performance comparable to the best individual models, with a high concentration of quality values near 1. By distributing the weights more evenly, it preserves a greater contribution from both selected surrogate models, while still reducing the variability observed for some individual configurations.

Its execution time remains higher than that of fixed surrogate models because both selection mechanisms must be applied before aggregation.

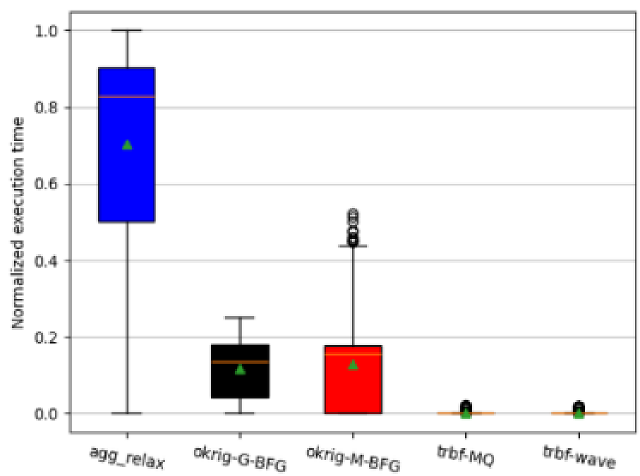

Figure 17: Normalized execution-time distribution for the relaxed aggregation strategy.

Dynamic aggregation The dynamic aggregation strategy uses a selectivity parameter β to control how strongly the model with the highest estimated quality is favoured. Preliminary experiments with different values of β led to the selection of $\beta = 5$.

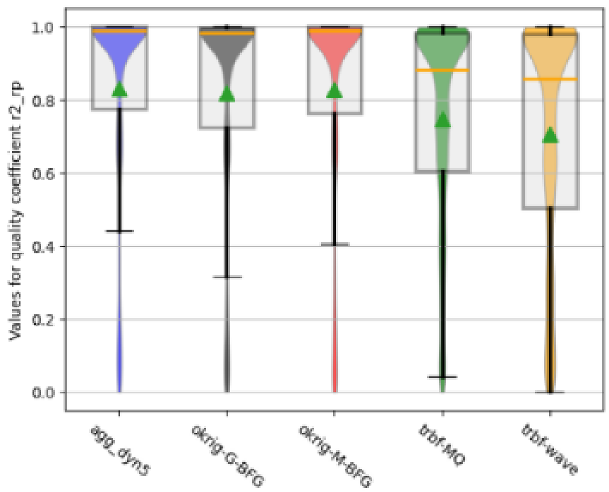

Figure 18: Distribution of the $r2\_rp$ quality metric obtained with the dynamic aggregation strategy using $\beta = 5$.

The dynamic approach achieves predictive performance comparable to the best individual surrogates and presents a slightly stronger concentration of high $r2\_rp$ values. When important quality differences exist between the two selected models, the weighting mechanism increasingly favours the most reliable surrogate while retaining a limited contribution from the second model.

As for the other aggregation strategies, this additional robustness is obtained at a higher computational cost than for individual models.

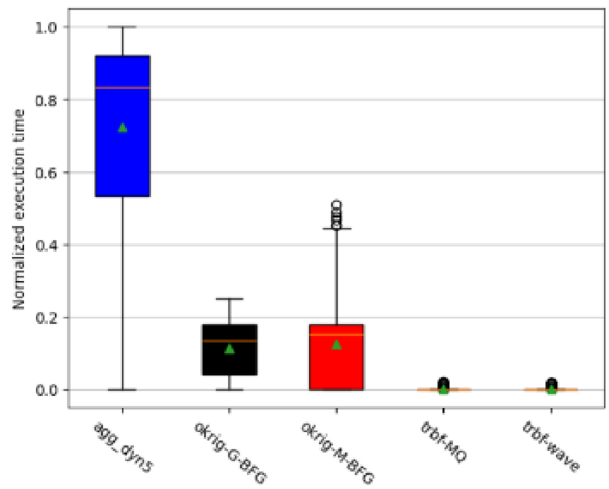

Figure 19: Normalized execution-time distribution for the dynamic aggregation strategy using $\beta = 5$.

Comparison of aggregation strategies Figure 20 directly compares the predictive quality of the three aggregation strategies.

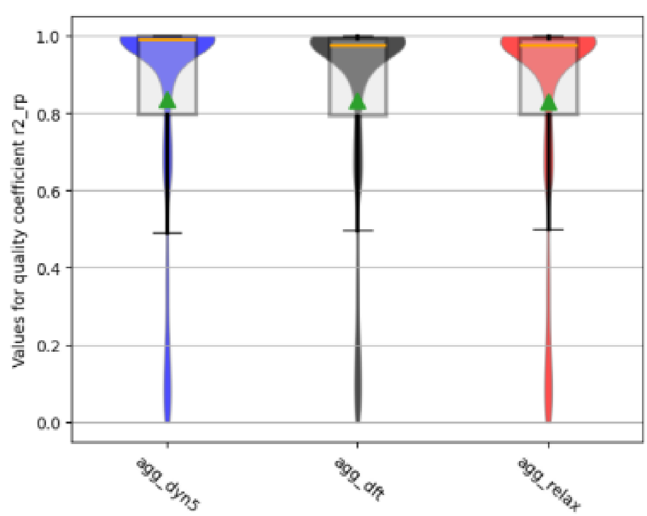

Figure 20: Distribution of the $r2\_rp$ quality metric for the three aggregation strategies.

The three approaches provide very similar predictive performance, with distributions strongly concentrated near 1. The dynamic strategy with $\beta = 5$ shows a slightly higher concentration of high-quality predictions, while the default and relaxed approaches exhibit marginally greater dispersion.

To further analyse the behaviour of the aggregation mechanisms, Table 9 reports the selection frequencies and median aggregation weights obtained with the default `agg-cpp` strategy during external validation. Only this strategy is shown, as the relaxed and dynamic variants exhibit very similar selection and weighting patterns.

Table 9: Selection frequency and median aggregation weight for the `agg-cpp` strategy in external validation.

Family	Kernel / seq.	It. 1		Mid it.		Last it.	
		Sel.	Med. w.	Sel.	Med. w.	Sel.	Med. w.
TRBF	Default	36.11	48.01	36.11	48.01	36.11	48.00
	Gaussian	9.15	**49.76**	9.15	**49.76**	9.15	**49.76**
	MQ	0.66	7.45	0.66	7.45	0.66	7.85
	RQ	4.90	0.00	4.90	0.00	4.90	0.00
	Wave	**49.18**	49.15	**49.18**	49.15	**49.18**	49.15
OKG	Exp.	4.31	50.64	4.31	50.64	4.35	50.64
	Full param.	14.01	**70.15**	14.01	**70.15**	13.91	**77.24**
	Gaussian	17.09	51.48	17.09	51.48	17.06	51.50
	Matern	**64.60**	51.19	**64.60**	51.19	**64.68**	51.19

The selection patterns remain remarkably stable throughout the validation process. Wave is the most frequently selected TRBF kernel, whereas Gaussian receives the highest median weight. Within the OKG family, Matern clearly dominates the selections, while the Full parameter kernel receives the largest median weight. This confirms that selection frequency and aggregation contribution capture complementary aspects of surrogate relevance.

Their computational costs are also comparable, as shown in Figure 21. Overall, the three aggregation strategies exhibit similar predictive quality, model contribution patterns and computational cost, indicating that the alternative weighting formulations only marginally affect the overall behaviour.

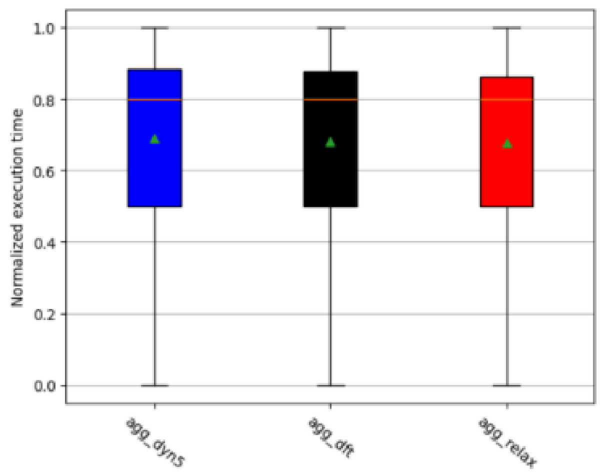

Figure 21: Normalized execution-time distribution for the three aggregation strategies.

Overall, the external validation results show that aggregation improves predictive stability compared with relying on a single surrogate model. The three weighting strategies provide similar levels of predictive quality and computational cost, although the dynamic strategy with $\beta = 5$ appears slightly more selective and robust.

6.4.2 Optimization Results

This section evaluates the optimization performance, robustness, and computational cost of the considered aggregation strategies. The selection strategies are also included in the comparison as adaptive reference methods, allowing the additional benefit of combining surrogate models to be assessed.

Both the selection and aggregation strategies rely on two surrogate-model families. The O-Kriging candidate set contains the Gaussian, Matern, Fully Parameterized, and Exponential O-Kriging models, all trained using the BFGS hyperparameter estimation procedure. The TRBF candidate set contains the default TRBF sequence together with the Gaussian, Multiquadric (MQ), Rational Quadratic (RQ), and Wave configurations.

For each response of each benchmark test case, the aggregation strategies combine two adaptively selected surrogate models. One O-Kriging model is selected from the O-Kriging candidate set, while one TRBF model is independently selected from the TRBF candidate set. The aggregated prediction therefore combines the currently selected representative of each surrogate family rather than all available models simultaneously.

The standalone selection strategies use the same candidate pools but retain only the surrogate selected by their respective selection mechanisms. They are included as reference strategies to determine whether combining the selected

O-Kriging and TRBF models provides additional robustness or optimization performance.

Three aggregation strategies are considered. The `agg-dft` strategy corresponds to the default aggregation method implemented in the Minamo platform. The `agg-rel` strategy uses a more balanced distribution of the weights between the selected models. Finally, `agg-dyn5` denotes the dynamic aggregation strategy with a selectivity parameter $\beta = 5$, which increasingly favours the surrogate model with the highest estimated predictive quality.

Figure 22 presents the data profile obtained using the lower-quartile objective value as threshold.

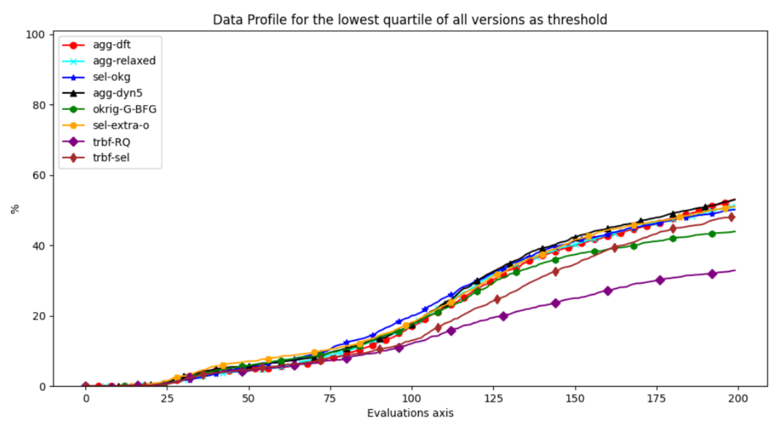

Figure 22: Data profile per evaluation using the lower-quartile threshold for the selection and aggregation strategies.

The aggregation strategies improve optimization robustness by combining surrogate models with complementary approximation properties. In particular, the dynamic aggregation strategy, `agg-dyn5`, achieves the best overall optimization performance among the investigated approaches. By adapting the weights according to the estimated surrogate quality, it frequently reaches better final objective values than both the fixed models and the standalone selection strategies.

The default and relaxed aggregation strategies also provide robust results, although their performance remains generally below that of the dynamic aggregation strategy. The more balanced weights used by the relaxed approach reduce its ability to strongly favour the most accurate surrogate model.

To further analyse the behaviour of the aggregation mechanism, Table 10 reports the selection frequencies and median aggregation weights obtained with the dynamic aggregation strategy using $\beta = 5$. The values are presented for the first, middle, and final stages of the optimization process. Similar overall trends were observed for the default and relaxed aggregation strategies; therefore, only the detailed results of the dynamic strategy are reported.

Family	Kernel / seq.	It. 1		Mid it.		Last it.	
		Sel.	Med. w.	Sel.	Med. w.	Sel.	Med. w.
TRBF	Default	18.32	**0.00**	**40.03**	0.07	**49.21**	0.52
	Gaussian	7.44	**0.00**	10.60	0.03	10.80	0.09
	MQ	0.15	**0.00**	0.93	**3.18**	1.01	**1.13**
	RQ	1.18	**0.00**	5.44	**0.26**	6.77	**0.98**
	Wave	**72.91**	**0.00**	**43.01**	0.00	**32.22**	0.22
OKG	Exp.	1.38	**100.00**	3.24	77.24	4.18	79.72
	Full param.	10.77	**100.00**	17.81	**83.13**	22.01	74.38
	Gaussian	**42.69**	**100.00**	**29.28**	**99.96**	**28.22**	**99.91**
	Matern	**45.17**	**100.00**	**49.67**	**100.00**	**45.59**	**99.99**

Table 10: Selection frequency and median aggregation weight for the dynamic aggregation strategy across all benchmark test cases. Sel. denotes the selection frequency in percent, and Med. w. denotes the median aggregation weight in percent. Bold values indicate the two largest values within each family, metric, and iteration stage.

The results show that the models selected within each surrogate family evolve during the optimization process. Within the TRBF family, the Wave kernel is strongly favoured during the first iterations, while the default sequence becomes increasingly selected during the middle and final stages. Within the O-Kriging family, the Gaussian and Matern kernels are the most frequently selected configurations throughout the optimization.

The median weights also show that the dynamic aggregation strongly favours the selected O-Kriging model in most cases. The TRBF contribution remains limited, although it slightly increases during the later stages of the optimization. This behaviour indicates that the dynamic strategy becomes highly selective when an important quality difference exists between the two surrogate families, while still preserving a small contribution from the alternative model. The strong influence of the Gaussian and Matern O-Kriging models is consistent with their robust predictive performance observed during external validation.

Aggregation strategies are, however, more computationally expensive than selection strategies because two surrogate models must be retained and evaluated simultaneously. Overall, the dynamic aggregation strategy provides the strongest optimization quality, while the extra-familial selection strategy remains more attractive when computational cost is an important consideration.

7 KPLS for High-Dimensional Problems

Surrogate modeling becomes increasingly challenging as the dimensionality of the optimization problem grows. In particular, Kriging models require the estimation of several kernel hyperparameters, which can become computationally expensive and difficult to optimize in high-dimensional spaces. This issue is closely related to the curse of dimensionality, which often leads to reduced predictive accuracy and increased model construction costs.

To alleviate these limitations, Kriging with Partial Least Squares (KPLS) was investigated [2]. KPLS combines Gaussian process modeling with a dimension-reduction step based on Partial Least Squares (PLS), allowing the surrogate model to be constructed in a reduced latent space.

7.1 High-Dimensional Test Cases

A dedicated benchmark composed exclusively of high-dimensional problems was considered to evaluate the investigated surrogate-modeling strategies under more demanding conditions. The benchmark contains 15 test cases with dimensions ranging from 20 to 162 design variables, as summarized in Table 11.

The selected problems exhibit different characteristics, including constrained and unconstrained formulations, analytical benchmark functions, engineering applications, multimodal response surfaces, and problems with different levels of non-linearity and variable interactions. This diversity makes it possible to assess whether the considered methods remain robust across different high-dimensional problem structures rather than being adapted to a single type of response surface.

Test case	**Dimension**	**# constraints**
G3MOD_plog_20	20	1
G3MOD_plog_50	50	1
G3MOD_plog_100	100	1
G2_plog_50	50	2
G2_plog_100	100	2
Schwefel_20	20	0
Schwefel_50	50	0
Schwefel_100	100	0
Ackley_20	20	0
Rastrigin_20	20	0
Rosenbrock_30	30	0
bldg-grouping	32	3
bldg-extgrouping	96	3
dome-extgrouping	29	3
wing-nogrouping	162	3

Table 11: High-dimensional benchmark problems used in this study.

The following subsections evaluate the predictive and optimization behaviour of the considered surrogate strategies on this benchmark, with particular attention given to robustness, convergence performance, and computational cost.

7.2 Optimization Results

Since no external validation experiment was performed for the KPLS configuration, the comparison is conducted directly within the optimization framework. The Gaussian KPLS model is compared against the standard Gaussian and Matern O-Kriging models, as well as the default TRBF sequence.

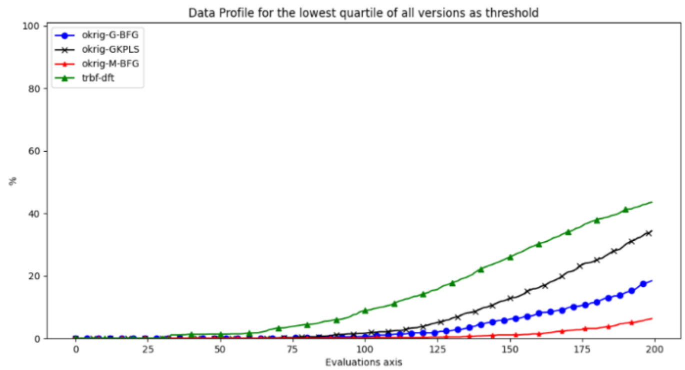

Figure 23: Data profile obtained for the considered surrogate models on the high-dimensional benchmark using the lower-quartile threshold.

Figure 23 shows that the Gaussian KPLS model clearly outperforms the standard Gaussian and Matern O-Kriging models. The introduction of Partial Least Squares significantly improves the robustness of the OKG family on high-dimensional problems, confirming the benefits of dimensionality reduction in this context. The default TRBF sequence nevertheless remains the best performing approach overall. However, Gaussian KPLS emerges as the strongest OKG variant and the closest competitor to the TRBF sequence. These results suggest that future adaptive strategies dedicated to high-dimensional problems could focus on a reduced set of candidate models composed of Gaussian KPLS and TRBF-based approaches.

7.3 Selection and Aggregation Strategies for High-Dimensional Problems

To evaluate whether adaptive surrogate management strategies could further improve KPLS-based optimization, a selection strategy (`sel-high-di`) and an aggregation strategy (`agg-high-di`) which is the dynamic aggregation strategy mentioned previously, were investigated. Both approaches combine the Gaussian KPLS model with the default TRBF sequence.

Surrogate	Selection frequency
`trbf-dft`	**51.1%**
`okg-GKPLS`	48.9%

Table 12: Global selection frequency of each surrogate family in the high-dimensional selection strategy across all benchmark test cases.

The selection frequencies are almost evenly distributed between the two surrogate models. The default TRBF sequence is selected slightly more often, with 51.1% of the selections, while Gaussian KPLS accounts for 48.9%. This balanced distribution confirms that both models provide complementary behaviour across the high-dimensional benchmark and that neither surrogate is consistently optimal for all considered problems.

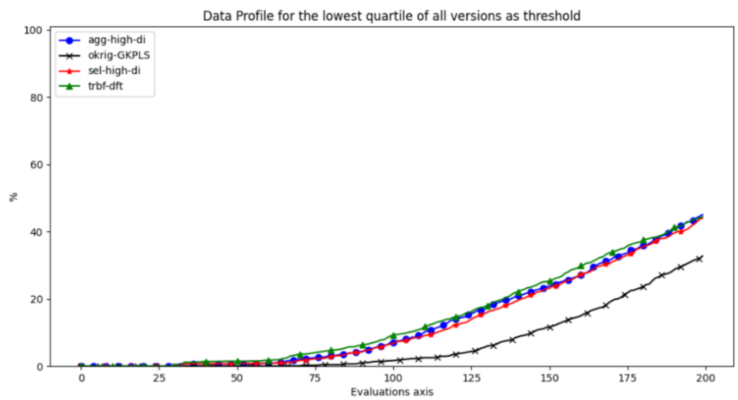

Figure 24: Data profile obtained on the high-dimensional benchmark using the lower-quartile threshold.

Figure 24 shows that both adaptive strategies consistently outperform the standalone `okrig-GKPLS` model. The selection and aggregation mechanisms solve a larger proportion of problems for a given evaluation budget, demonstrating that the complementary strengths of KPLS and TRBF can be successfully exploited within a single optimization framework. The aggregation and selection approaches also achieve performances close to those of `trbf-dft`, which remains the best-performing surrogate on the considered benchmark. These results indicate that adaptive surrogate management improves the robustness of KPLS-based optimization but does not fully surpass the performance of the TRBF sequence.

Surrogate	**First**	**Middle**	**Last**
`okg-GKPLS`	**84.42%**	49.69%	49.83%
`trbf-dft`	15.58%	**50.31%**	**50.17%**

Table 13: Average median weights assigned by the high-dimensional aggregation strategy across all benchmark test cases.

The aggregation weights evolve during the optimization process. During the first iterations, the Gaussian KPLS model receives most of the weight, with an average median value of 84.42%, while the default TRBF sequence contributes only 15.58%. During the middle and final stages, the weights become almost perfectly balanced between both surrogate models. This evolution suggests that KPLS is particularly important during the early exploration phase, whereas the contribution of the TRBF sequence increases as more evaluations become available. The balanced weights observed later in the optimization confirm the complementarity of the two surrogate families.

Overall, the experiments confirm that combining multiple surrogate models is beneficial in high-dimensional settings. However, the strong performance of `trbf-dft` suggests that TRBF-based approaches constitute the most promising direction for high-dimensional surrogate-assisted optimization.

7.4 TRBF-Based Strategies for High-Dimensional Problems

Since the previous experiments identified TRBF as the most competitive surrogate family in high-dimensional settings, several TRBF-based strategies were investigated. Three individual kernels (MQ, RQ, and Wave), the default TRBF sequence, a TRBF selection strategy, and a TRBF aggregation strategy (dynamic based aggregation strategy) were compared.

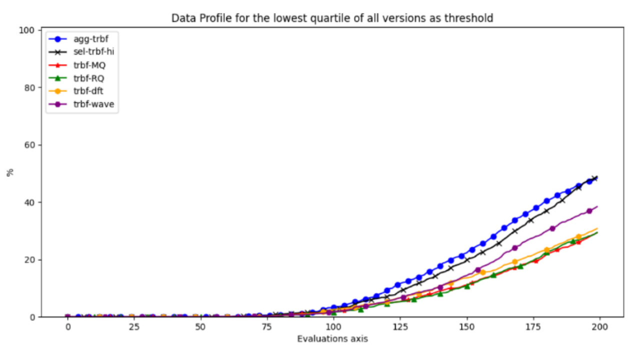

Figure 25: Data profile obtained on the high-dimensional benchmark using the lower-quartile threshold.

Figure 25 shows a clear advantage of the adaptive approaches over the individual TRBF kernels. Both the selection and aggregation strategies consistently outperform MQ, RQ, and Wave, confirming the benefit of exploiting the complementarity of several TRBF models. The aggregation strategy achieves the best overall performance, while the selection strategy remains highly competitive and also improves upon the default TRBF sequence.

Table 14 reports the kernel-selection frequencies obtained with the TRBF selection strategy.

Table 14: Kernel-selection frequencies for the TRBF selection strategy.

Kernel	Selection count	Frequency (%)
MQ	**1409**	**46.97**
Wave	844	28.13
RQ	747	24.90

MQ is the most frequently selected kernel, with approximately 46.97% of the selections. However, Wave and RQ together account for more than half of the selections, showing that no single kernel is consistently optimal across all high-dimensional problems. This supports the use of an adaptive selection mechanism rather than relying on one fixed TRBF kernel.

Table 15 presents the average median weights assigned by the TRBF aggregation strategy at different stages of the optimization.

Table 15: Average median kernel weights for the TRBF aggregation strategy.

Kernel	First iterations	Middle iterations	Last iterations
MQ	9.32	33.05	34.27
RQ	3.50	20.58	24.00
Wave	**70.59**	**39.88**	**39.21**

Wave receives the largest weight throughout the optimization, particularly during the first iterations, where its average median weight reaches 70.59. As more evaluations become available, the weights become more balanced, with MQ and RQ gaining importance during the middle and final stages. This evolution indicates that the aggregation strategy adapts the contribution of each kernel as the amount of available information increases.

These improvements are obtained with limited computational overhead, since the adaptive strategies involve only three TRBF kernels. Overall, the results confirm that adaptive TRBF management provides a favourable compromise between robustness, optimization performance, and computational cost for high-dimensional problems. Among the investigated methods, the TRBF aggregation strategy provides the strongest overall performance.

7.5 Computational Cost on High-Dimensional Problems

The computational cost of the investigated methods was evaluated by averaging the total execution time over all available runs and high-dimensional benchmark problems. The resulting mean execution times are presented in Figure 26.

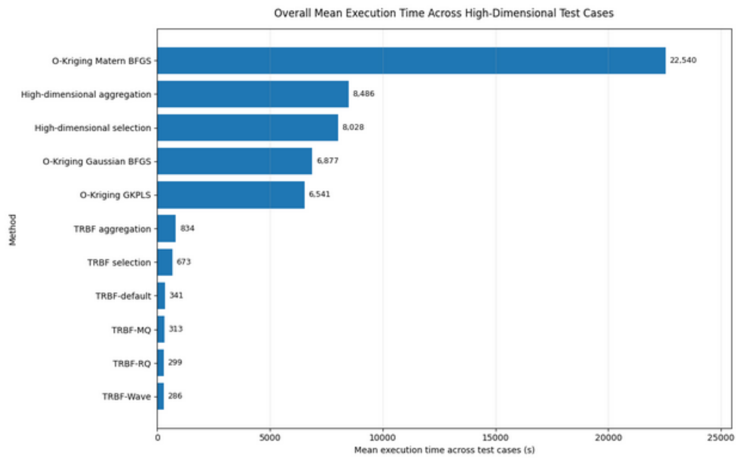

Figure 26: Overall mean execution time of the investigated methods across the high-dimensional benchmark problems.

The TRBF models are the least computationally expensive methods, with only limited differences between the considered kernels. The TRBF selection and aggregation strategies introduce additional overhead, but remain substantially faster than the Kriging-based approaches.

Gaussian KPLS is more expensive than the TRBF methods, while remaining less costly than standard high-dimensional O-Kriging. Among the investigated methods, the Matern O-Kriging model optimized with BFGS exhibits the highest execution time. The high-dimensional selection and aggregation strategies also incur a significant computational cost due to the repeated construction and management of both KPLS and TRBF models.

Overall, the results highlight a trade-off between computational cost and robustness. TRBF-based adaptive strategies provide a favourable compromise, whereas KPLS-based strategies require more computational resources.

8 Conclusion and Perspectives

This work investigated adaptive surrogate-model selection and aggregation strategies for improving the robustness of surrogate-based optimization. Several Kriging and Tuned Radial Basis Function (TRBF) models were evaluated using the combined *r2_rp* quality criterion, considering both optimization performance and computational cost.

The experiments confirmed that no single surrogate model performs best on all problems. Among the Kriging models, the Gaussian and Matern kernels provided the most robust results, while the fully parameterized configuration introduced a high computational cost without sufficient performance improvement. Within the TRBF family, the Multiquadric, Rational Quadratic, and Wave kernels achieved the strongest overall results.

Adaptive selection improved robustness compared with fixed surrogate models. In particular, the extra-familial selection strategy provided a favourable compromise between optimization performance and computational cost by exploiting the complementary properties of Kriging and TRBF models. Aggregation achieved the best overall optimization performance, with the dynamic aggregation strategy using $\beta = 5$ providing the most robust results, although at a higher computational cost.

The high-dimensional experiments showed that standard Kriging models become less competitive as the number of variables increases. Gaussian KPLS reduced the difficulties associated with high-dimensional hyperparameter estimation, but the TRBF-based strategies remained more effective on the considered benchmark. In particular, TRBF aggregation based on the MQ, RQ, and Wave kernels achieved the best overall performance, while TRBF selection provided a competitive alternative with a lower computational cost.

The execution-time analysis confirmed that TRBF models are substantially less expensive than Kriging-based approaches. Their adaptive selection and aggregation variants introduced only a moderate overhead, whereas strategies involving Gaussian KPLS or standard O-Kriging required significantly more computational resources.

Overall, dynamic aggregation is the preferred strategy when optimization quality is the main objective, while extra-familial selection offers a more balanced solution for general problems. For high-dimensional optimization, TRBF aggregation provides the best performance, whereas TRBF selection is more attractive when computational efficiency is a priority.

Several directions could be considered for future work. A particularly promising one would be the development of a new TRBF sequence combining the Wave, Rational Quadratic, and Multiquadric kernels. Since these kernels showed strong and complementary behaviour, particularly on high-dimensional problems, such a sequence could adapt the active kernel through the skip mechanism according to the characteristics of the problem and the current stage of the optimization process. This could provide a more flexible alternative to the current fixed TRBF sequence while preserving the low computational cost of TRBF models.

Another perspective would be to investigate aggregation methods based on optimized weight factors. In the approach proposed by Acar and Rais-Rohani [1], several surrogate models are combined through a weighted sum, while the weights are obtained by solving an optimization problem that minimizes a prediction-error criterion. Unlike the aggregation methods considered in this work, where the weights are directly derived from the estimated surrogate quality, this approach explicitly optimizes the contribution of each model to the ensemble. It could therefore provide an interesting alternative for improving prediction accuracy and robustness, although it would introduce an additional computational cost during the construction of the aggregated surrogate.

Further work could also investigate cost-aware quality criteria, dynamically adapted candidate sets, and additional constrained or engineering optimization problems to assess the generality of the proposed strategies.

Overall, the results demonstrate that adaptive surrogate strategies can significantly improve optimization robustness, while TRBF-based approaches appear particularly promising for achieving a favourable balance between performance and computational efficiency, especially in high-dimensional settings.

References

[1] Erdem Acar and Masoud Rais-Rohani. "Ensemble of Metamodels with Optimized Weight Factors". In: *Structural and Multidisciplinary Optimization* 37.3 (2009), pp. 279–294. DOI: 10.1007/s00158-008-0230-y.

[2] Mohamed-Amine Bouhlel et al. *Improving Kriging Surrogates of High-Dimensional Design Models by Partial Least Squares Dimension Reduction*. HAL preprint. HAL Id: hal-01232938. 2015. URL: https://hal.science/hal-01232938v1.

[3] Martin D. Buhmann. *Radial Basis Functions: Theory and Implementations*. Cambridge University Press, 2003.

[4] Cenaero. *Minamo 4.4.0 Theoretical Manual*. Cenaero. Apr. 2026.

[5] A. E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing*. Natural Computing Series. Berlin, Heidelberg: Springer, 2015.

[6] Tushar Goel et al. "Ensemble of Surrogates". In: *Structural and Multidisciplinary Optimization* (2007).

[7] Jorge J. Moré and Stefan M. Wild. "Benchmarking Derivative-Free Optimization Algorithms". In: *SIAM Journal on Optimization* 20.1 (2009), pp. 172–191. DOI: 10.1137/080724083.

[8] C. Sainvitu, V. Iliopoulou, and I. Lepot. "Global Optimization with Expensive Functions – Sample Turbomachinery Design Application". In: *Recent Advances in Optimization and its Applications in Engineering*. Springer Berlin Heidelberg, 2010, pp. 499–509. DOI: 10.1007/978-3-642-12598-0_44.

[9] Steve Smeets. *Implémentation de nouveaux coefficients de qualité dans la plateforme d'optimisation multidisciplinaire, Minamo, et tentatives d'amélioration par le biais de méthodes d'agrégation*. Internship report. Cenaero, May 9, 2020.

[10] Wikipedia contributors. *Kriging — Wikipedia, The Free Encyclopedia*. 2025. URL: https://en.wikipedia.org/wiki/Kriging.