Gaussian Processes for AI-based Antenna Data Analysis

Executive Summary Report

ESA Contract No. ESA 4000141763/23/NL/GLC/cb

***Abstract*—This document provides a high-level overview of the work carried out under the ESA Contract No. 4000141763/23/NL/GLC/cb. It reports on a new software prototype for designing and optimizing antenna systems using a Bayesian Optimization (BO) framework integrated into TICRA Tools. The prototype efficiently handles complex electromagnetic (EM) design problems by leveraging meta-modeling and advanced acquisition functions tailored for multi-output Gaussian Processes (MOGPs). The framework enables the automation of data collection, meta-model construction, and optimization, significantly reducing the computational burden of evaluating large-scale antenna designs. Results demonstrate that the prototype can offer substantial improvements in robustness, efficiency and accuracy over traditional methods, providing antenna engineers with a powerful tool for optimizing designs with thousands of residuals, paving the way for advanced design tasks such as uncertainty quantification and multi-objective optimization.**

***Index Terms*—Machine learning, meta-modelling, space applications, antenna design software**

I. Introduction

Modern antenna design for space applications, such as telecommunications or Earth observation, requires high-precision fine-tuning of both geometric and material parameters to meet stringent performance specifications. This necessity becomes even more pronounced as the design process advances, making it mission-critical to achieve absolutely optimal performance. Sub-optimal or non-robust designs can be extremely costly in space applications, where resources are scarce and failures can have significant consequences. As such, ensuring robust and efficient antenna designs is essential to avoid mission failures and maximize the return on investment. In many cases, the fine-tuning task can effectively be addressed using advanced simulation-based optimisation software. However, general-purpose optimization algorithms, including gradient-based methods [1] and meta-heuristics [2], [3], typically lack robustness and rely on a large number of computationally expensive electromagnetic (EM) simulations to identify the optimal design parameters. These methods often become computationally prohibitive for many practical applications, particularly in the final design stages, when full-wave simulations are required to capture the complex interactions of the antenna with itself and its environment. In such cases, the evaluation of a single design may take hours, making design optimization intractable within project deadlines. To address the challenges of conventional optimization methods, this work investigates data-driven meta-modeling for black-box, compute-intensive antenna design tasks. In this context, meta-modeling refers to machine-learning (ML)-based approximations to full-scale EM simulators, built from carefully selected simulation data. The key benefit is that meta-models can in many cases effectively replace true EM simulations for practical purposes, offering significantly faster evaluations while retaining sufficient accuracy [4]. The primary target of the proposed framework is computationally intensive antenna design tasks, where conventional methods typically fall short. The targeted tasks are characterized by one or more of the following criteria:

- Expensive, black-box objective functions that, e.g., rely on repeated full-wave simulations.
- First – and Second-Order Derivatives are unavailable and expensive to approximate.
- Complex optimisation landscapes with multiple local optima.
- Good starting guesses for the design parameters are uncertain or not known.
- Design tasks involving many frequencies and design criteria.

To allow antenna engineers to address such design tasks more efficiently, the proposed framework uses Bayesian optimisation (BO), which is particularly well-suited for optimising expensive black-box functions using low volume data [4]–[6]. BO leverages a probabilistic meta-model, in this case a Gaussian Process (GP) [7], to guide the search for the global optimum. This approach is advantageous because it is sample-efficient, gradient-free and yet efficient in navigating the search space to find high-performing solutions with fewer evaluations.

Overall, the proposed meta-modelling framework offers a robust alternative to conventional optimization methods. It eliminates the need for gradient information, making it well-suited for problems with expensive objective functions and uncertain initial conditions. Additionally, by employing a BO framework, the prototype provides a more efficient approach than global search methods like genetic algorithms or particle swarm optimization, offering faster convergence and reduced computational costs.

To showcase the framework's potential, this paper considers a multi-criteria design problem, where the goal is to optimally balance the conflicting objectives of gain and return loss for a dual reflector system consisting of two rotationally symmetric reflectors. This serves as a representative example of an antenna design task, that would be both time-consuming and computational expensive using traditional methods.

The report is structured as follows. Section II provides background on meta-modelling and Bayesian Optimisation. Section III gives a high-level introduction to the Meta-modelling framework and its key components. The multi-criteria case study is presented in Section IV, while Section V draws overall conclusions and discusses perspectives.

II. Background

The primary objective of the meta-modelling framework is to address large-scale, black-box optimization problems of the form:

$$\text{minimize} \quad h(\overline{r}(\overline{x})) \tag{1a}$$

$$\text{subject to} \quad \overline{l} \preceq \overline{x} \preceq \overline{u}, \tag{1b}$$

where $\overline{x} \in \mathbb{R}^n$ are the design variables, $\overline{r} : \mathbb{R}^n \to \mathbb{R}^m$ is a vector function that produces m residuals for a given design, and $h : \mathbb{R}^m \to \mathbb{R}$ is a function that maps the vector of residuals to a scalar objective function value.

For each design variable, lower and upper bounds are specified in (1b). The vector residual function $\overline{r}(\overline{x})$ is defined as:

$$r_i(\overline{x}) = w_i(\gamma_i - f_i(\overline{x})), \tag{2}$$

where w_i is the weight of the i'th residual, γ_i is the goal for the i'th residual, and $f_i : \mathbb{R}^n \to \mathbb{R}$ computes the i'th performance of interest.

While many optimization problems (1) can be solved with conventional gradient-based or global search algorithms, challenges arise when dealing with expensive, black-box objective functions, as these rely on repeated full-wave simulations to compute the residuals. In such cases, first- and second-order derivatives may be unavailable or costly to approximate, and the optimization landscape might have multiple local optima, with uncertain or unknown starting points. Moreover, if changes in objective function h, weights ($\overline{w}$), or goals ($\overline{\gamma}$) are needed, the optimization must be restarted, wasting expensive computations of the residuals $\overline{r}(\overline{x})$. This is often the case, when multiple conflicting objectives must be balanced. Here different weights ($\overline{w}$), or goals ($\overline{\gamma}$) must be adjusted to represent different trade-offs.

A. Meta-Modelling

Given an expensive black-box function $F : \mathcal{X} \subset \mathbb{R}^i \to \mathbb{R}^o$ (e.g., an EM simulator), the goal of meta-modelling is to construct a fast, cheap-to-evaluate approximation $G(x) \approx F(x)$, while ensuring sufficient accuracy within a subset of the design space $\mathcal{X}$:

$$\max_{x \in \mathcal{X}} \|F(x) - G(x)\| \leq \epsilon, \quad \text{given the tolerance} \quad \epsilon > 0. \tag{3}$$

Typically, G is a data-driven model built from an initial training set $\mathcal{D} = \{(x_i, y_i) \mid i = 1, ..., N\}$, where $F(x_i) = y_i$ are the observations for design sites $X = \{x_i\}_{i=1}^N$, and $x_i \in \mathbb{R}^i$. Once trained and validated, the meta-model G can be used in place of F, allowing faster evaluations, especially useful in optimization - and risk/uncertainty quantification tasks that would otherwise be computationally prohibitive or practically infeasible, due to the need for a large number of full simulations.

B. Gaussian Processes

There are numerous methods for constructing data-driven meta-models, such as neural networks, reduced-order models, polynomial chaos expansions, and support vector regressors [8]. This activity specifically uses Gaussian Processes (GPs) [7]. This choice is driven by the well-established use of GPs in practical applications, including antenna design, and its ability to build approximations for highly nonlinear engineering simulations, especially in sparse data regimes [4]. A Gaussian Process (GP) defines a distribution over functions, where the distribution of function values at a finite set of points $\{x_1, x_2, ..., x_N\}$ follows a multivariate Gaussian distribution on the corresponding function values $\{y_1, y_2, ..., y_N\}$. The GP model is given by:

$$G(x) = \mu(x; \beta) + Z(x; \theta), \quad Z \sim \mathcal{N}(0, \mathbf{K}), \tag{4}$$

where $\mu(x; \beta)$ is the mean function, often assumed to be linear, and $Z(x; \theta)$ is a zero-mean GP with covariance matrix $\mathbf{K}$ parameterized by hyperparameters θ. The covariance between two points x_i and x_j is:

$$\mathbf{K}(x_i, x_j \mid \theta) = \text{Cov}(Z(x_i), Z(x_j)). \tag{5}$$

Given training data $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$, the GP provides a predictive mean $\mu_{\mathcal{D}}(x^*)$ and variance $\sigma^2_{\mathcal{D}}(x^*)$ for any new input x^*, derived as:

$$\mu_{\mathcal{D}}(x^*) = \mathbf{k}(x^*)^\top \mathbf{K}^{-1} \mathbf{y}, \tag{6}$$

$$\sigma^2_{\mathcal{D}}(x^*) = K(x^*, x^*) - \mathbf{k}(x^*)^\top \mathbf{K}^{-1} \mathbf{k}(x^*), \tag{7}$$

where $\mathbf{k}(x^*)$ is the covariance vector between x^* and the training points, and $\mathbf{y}$ is the vector of observed outputs. The variance $\sigma^2_{\mathcal{D}}(x^*)$ is an uncertainty estimate, which can be used to express confidence in predictions and as a means to guide further sampling.

C. Bayesian Optimization

BO collects the properties of GPs into a framework for robust, sample-efficient, black-box optimisation, specifically targeting problems of the general type (1). The BO process iteratively refines the GP meta-model, balancing exploration of high-uncertainty regions and exploitation of promising designs. The framework is illustrated in Figure 1, where the key steps include:

1) **Initial Sampling:** An initial set of sample points is generated, often using low-discrepancy sampling methods, to cover the design space.
2) **Meta-Model Construction:** A GP is trained on the initial data, modeling both the objective function and the uncertainty in its predictions.
3) **Active Learning:** Active learning uses the meta-model to identify the most informative points in the design space to sample next. An acquisition function, which

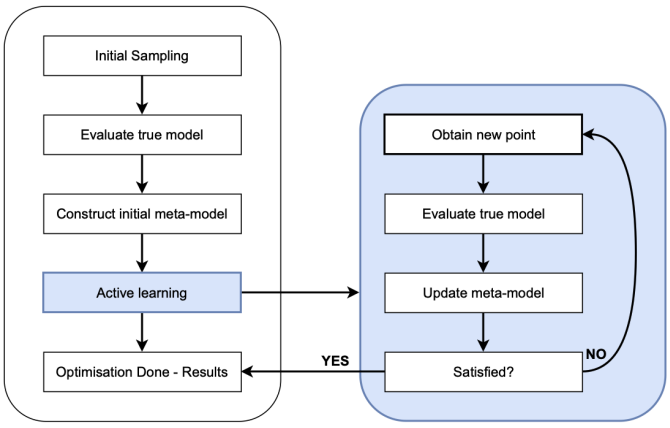

Fig. 1. Bayesian Optimisation workflow

balances exploration and exploitation, guides this selection using the predictions and uncertainty estimates of the meta-model.

4) **Iterative Refinement:** The selected design is evaluated using the true function, and the GP is updated with the new data. This process repeats until convergence or a stopping criterion (e.g., computational budget) is met.

Overall, BO allows for efficient exploration and exploitation of complex design spaces with minimal computational resources, making it a valuable tool for addressing computational expensive antenna designs tasks.

III. The Meta-Modelling Framework

The following chapter provides a high-level overview of the essential components required to develop an end-to-end BO framework integrated into TICRA Tools. The framework is driven by the need to adapt the concept of meta-modelling to the specific challenges of electromagnetic (EM) design while making it accessible to antenna engineers without requiring expertise in machine learning. This chapter outlines the key methods and design decisions that streamline the processes of automated data collection, meta-model construction, and optimization, particularly in complex antenna design tasks where the number of residuals can reach into the thousands.

A. Multi-Residual Modelling

A key objective of the meta-modelling framework is the efficient handling of multiple outputs, particularly in large-scale antenna design tasks where the number of residuals can grow rapidly. For example, optimizing the cross-polar far-field magnitude across angular regions and multiple frequencies can result in hundreds or even thousands of residuals. Addressing this complexity requires a framework capable of modeling many outputs simultaneously, while ensuring flexibility and scalability.

In conventional BO, the focus is on constructing a meta-model for a scalar objective function $h(\overline{r}(\overline{x}))$, which limits the reuse of simulation data for different objectives. To overcome this limitation, the proposed framework models the underlying output functions $f_i(\overline{x})$, where:

$$r_i(\overline{x}) = w_i(\gamma_i - f_i(\overline{x})),$$

for $i = 1, \ldots, m$, allowing the reuse of data across different optimization tasks. This approach enables modifications to the goals γ_i, weights w_i, and other parameters without needing additional simulations, as these changes do not depend on $\overline{x}$.

1) Single-Output Gaussian Processes: One approach for modeling each residual function $f_i(\overline{x})$ is to use individual Gaussian Processes (GPs) for each output. For each f_i, a GP is defined as:

$$f_i(\overline{x}) \sim \mathcal{GP}(\mu_i(\overline{x}), K_i(\overline{x}, \overline{x}')),$$

where $\mu_i(\overline{x})$ is the mean function and $K_i(\overline{x}, \overline{x}')$ is the covariance (kernel) function. Given a set of training data $\mathcal{D} = \{(\overline{x}_i, y_i)\}_{i=1}^N$, the posterior predictive mean and variance for a new point $\overline{x}_*$ are given by:

$$\mu_{i,\mathcal{D}}(\overline{x}_*) = \mathbf{k}_i(\overline{x}_*)^\top \mathbf{K}_i^{-1} \mathbf{y}_i, \tag{8}$$

$$\sigma^2_{i,\mathcal{D}}(\overline{x}_*) = K_i(\overline{x}_*, \overline{x}_*) - \mathbf{k}_i(\overline{x}_*)^\top \mathbf{K}_i^{-1} \mathbf{k}_i(\overline{x}_*), \tag{9}$$

where $\mathbf{K}_i$ is the covariance matrix of the training data and $\mathbf{k}_i(\overline{x}_*)$ is the covariance vector between $\overline{x}_*$ and the training points.

While this approach works well for a small number of outputs, it quickly becomes computationally expensive as the number of outputs m increases, due to the need to train and update m separate GP models. The computational complexity of training each GP model is $\mathcal{O}(N^3)$, and with m outputs, the total cost scales as $\mathcal{O}(N^3 m)$.

2) Multi-Output Gaussian Processes and High-Order Gaussian Processes: To address the limitations of single-output GPs, the proposed framework uses Multi-Output Gaussian Processes (MOGPs) [9]. Instead of modeling each output $f_i(\overline{x})$ independently, MOGPs capture correlations between the outputs, improving efficiency and accuracy when outputs are not independent. In this case, the vector of outputs $\overline{f}(\overline{x}) = [f_1(\overline{x}), \ldots, f_m(\overline{x})]^\top$ is modeled as:

$$\overline{f}(\overline{x}) \sim \mathcal{GP}(\overline{\mu}(\overline{x}), \mathbf{K}(\overline{x}, \overline{x}')),$$

where $\mathbf{K}(\overline{x}, \overline{x}')$ is now a block covariance matrix that captures both input correlations and correlations between outputs:

$$\mathbf{K}(\overline{x}, \overline{x}') = \begin{bmatrix} K_{11}(\overline{x}, \overline{x}') & \cdots & K_{1m}(\overline{x}, \overline{x}') \\ \vdots & \ddots & \vdots \\ K_{m1}(\overline{x}, \overline{x}') & \cdots & K_{mm}(\overline{x}, \overline{x}') \end{bmatrix}.$$

Each block $K_{ij}(\overline{x}, \overline{x}')$ represents the covariance between outputs f_i and f_j at inputs $\overline{x}$ and $\overline{x}'$. This approach is more efficient when the outputs are correlated, reducing the number of independent models that need to be trained. However, standard MOGPs face significant computational challenges, especially for large-scale problems, as the complexity grows cubically with the number of outputs m and data points N. To overcome this bottleneck, the framework employs

High-Order Gaussian Processes (HOGPs) [10], which extend MOGPs by organizing the outputs into multi-dimensional tensors. In HOGPs, the outputs are structured as a tensor $\mathcal{F} \in \mathbb{R}^{m_1 \times m_2 \times \cdots \times m_Q \times N}$, where Q represents the number of modes (e.g., angles, frequencies), and N is the number of input points. The covariance structure is modeled using Kronecker products, significantly reducing the computational complexity:

$$\mathbf{K} = \mathbf{K}_X \otimes \mathbf{K}_Q \otimes \cdots \otimes \mathbf{K}_1,$$

where $\mathbf{K}_X$ captures the input correlations and $\mathbf{K}_1, \ldots, \mathbf{K}_Q$ capture the correlations along the tensor modes. This structure allows the complexity to scale more efficiently as $\mathcal{O}(N^3 + \sum_{i=1}^{Q} m_i^3)$, making HOGPs suitable for large-scale multi-output problems.

By employing HOGPs, it is ensured that the framework can handle complex antenna design tasks with thousands of outputs while maintaining computational efficiency. Moreover, the tensor structure of HOGPs enables better reuse of simulation data across different tasks, ensuring flexibility and scalability in optimization, uncertainty quantification, and multi-objective design.

B. Model Construction

To build an efficient and accurate meta-model for BO, several crucial decisions need to be made regarding data collection, the number of initial samples, and model selection. These aspects are important for minimizing computational costs while ensuring sufficient coverage of the input space. This section highlights the selected strategies used in TICRA Tools to automate these decisions, ensuring that antenna enginners can benefit form meta-models without requiring ML expertise.

1) Sampling Plan: The sampling plan determines how the input space is explored with an initial set of true model evaluations. Since each evaluation is computationally expensive, the choice of a sampling plan is critical to capture the model's behavior efficiently. A variety of strategies were investigated, including Full Factorial (FF), Latin Hypercube (LHC), Space-filling LHC (S-LHC), and Quasi Monte Carlo (QMC) [6]. After comparative studies, the Latin Hypercube Sampling with Gradient Correction (LHC-GC) was selected as the default due to its balance between computational efficiency and space-filling properties. LHC-GC starts with a standard LHC design and improves space-fillingness through gradient-based corrections. The space-fillingness is measured using the Morris-Mitchell criterion [11], which aims to maximize the minimum distance between points:

$$\Phi_q(X) = \left(\sum_{j=1}^{m} J_j d_j^{-1} \right)^{1/q},$$

where d_j are the unique distances between all point pairs in the set X, and J_j counts how often each distance appears. Minimizing $\Phi_q(X)$ ensures better space-fillingness, which reduces redundancy and improves model accuracy. The LHC-GC method strikes a pragmatic balance by improving the uniform coverage of the space without the high computational demands associated with space-filling LHC.

2) Number of Initial Sample Points: Determining the number of initial sample points is another critical factor, as too few samples can result in overfitting, while too many can lead to unnecessary computational cost. The minimum number of samples N_{min} depends on the complexity of the regression model used in the GP. This minimum can be calculated as follows:

$$N_{\text{min}} = \begin{cases} 1+1 & \text{for constant regression,} \\ (D+1)+1 & \text{for linear regression,} \\ \frac{1}{2}(D+1)(D+2)+1 & \text{for quadratic regression,} \end{cases}$$

where D is the number of design variables. While this gives the minimum required samples, additional samples are often needed to ensure the GP can model deviations from the regression model effectively. The selected approach follows a heuristic that adds a buffer to the minimum sample count, ensuring that the initial meta-model has enough data to capture the true behavior of the objective function. A linear regression model is typically paired with $N = 2(D+1)$ samples, while a quadratic model may use $N = 1.25 \times N_{\text{min}}$, ensuring a balance between computational cost and accuracy.

3) Model Selection Strategy: Model selection is an automated process where the goal is to identify the best GP meta-model based on the initial dataset [12]. Each candidate model $\mathcal{M}_k$ is defined by a combination of mean and kernel functions, such as:

$$\mathcal{M}_k = \mathcal{GP}(\mu_k(\theta), R_k(\theta)),$$

where $\mu_k(\theta)$ is the mean function (e.g., constant, linear, quadratic), and $R_k(\theta)$ is the kernel function (e.g., RBF, Matern32, Matern52). The challenge is to automatically select the best model that fits the dataset.

The process involves a grid search over combinations of mean and kernel functions and uses hyperparameter optimization to fit each model. The selected model is evaluated using the Akaike Information Criterion (AIC). AIC was found to be the most robust criterion for smaller initial sample budgets, as it balances model fit and complexity, penalizing overly complex models:

$$\text{AIC}(\theta) = 2K + \text{NLL}(\theta),$$

where K is the number of model parameters. AIC helps avoid overfitting, making it a reliable choice in TICRA Tools for the smaller to medium budgets used with BO.

Summary of Design Decisions

The selected strategies ensure that TICRA Tools can build efficient meta-models for BO by automating the key decisions involved in sampling and model construction:

- **Sampling Plan**: LHC-GC was chosen for its superior space-filling properties and computational efficiency.
- **Number of Initial Sample Points**: A heuristic-based approach, with additional buffer points, ensures that the meta-model has sufficient data for accurate predictions.

- **Model Selection**: The AIC criterion provides a robust, automated approach to selecting the best GP model for small to medium budgets, preventing overfitting and ensuring reliable performance.

C. Active Learning - Design of Acquisition Functions

In BO, acquisition functions guide the search for optimal points by balancing exploration and exploitation, with the goal of minimizing the number of true function evaluations. The acquisition function, $\alpha(x; \mathcal{D})$, is maximized to select the most informative point for sampling:

$$x^* = \underset{x \in \mathcal{X}}{\operatorname{argmax}}\, \alpha(x; \mathcal{D}),$$

where $\mathcal{D}$ represents the data collected so far. For antenna design, where optimization involves many correlated residuals, the choice of acquisition function is critical for efficiency.

1) Default Acquisition Function: The Log-Expected Improvement (LogEI) is chosen as the default acquisition function in TICRA Tools due to its numerical stability and effectiveness in high-dimensional optimization problems [13]. It improves upon the traditional Expected Improvement (EI) function by reformulating it in log-space, which prevents vanishing gradients and enhances the stability of gradient-based optimization.

The traditional EI measures the expected improvement over the best-known objective value, f^*:

$$\mathrm{EI}(x) = (\mu(x) - f^*)\Phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right),$$

where $\mu(x)$ and $\sigma(x)$ are the mean and standard deviation of the GP at x, and $\Phi(\cdot)$, $\phi(\cdot)$ are the cumulative and probability density functions of the standard normal distribution.

LogEI addresses numerical stability by applying a log transformation:

$$\mathrm{LogEI}(x) = \log_h\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) + \log(\sigma(x)),$$

where $\log_h(\cdot)$ is a numerically stable implementation that prevents underflow and ensures non-zero gradients, even when the improvement is small. This allows for reliable optimization, especially in complex, multi-modal landscapes.

2) Scaling and Stabilizing Acquisition Functions: To extend acquisition functions for multi-output models, particularly when residuals are correlated, Monte Carlo-based Expected Improvement for Composite Functions (EI-CF) is used [14]. EI-CF accounts for the vector-valued nature of residuals and computes the expected improvement over composite functions:

$$\text{EI-CF}_n(x) = \mathbb{E}_n\left[\max\left(h(\overline{r}(x)) - f_n^*, 0\right)\right],$$

where $h(\overline{r}(x))$ is a composite objective function of the residuals $\overline{r}(x)$, and f_n^* is the best observed value. EI-CF requires Monte Carlo (MC) simulations for estimation, which increases computational complexity. To address this increase in compute, the activity employs the following techniques:

a) Matheron's Rule for Efficient Sampling: Matheron's rule [15] is employed to efficiently sample from the posterior distribution of MOGPs, reducing the complexity of generating samples from high-dimensional distributions.

b) Combining LogEI with EI-CF: To further stabilize the MC-based acquisition functions, LogEI is combined with EI-CF. This combination ensures that even small improvements are captured, preventing vanishing gradients, and allowing robust optimization in high-dimensional landscapes.

IV. Case Study

To showcase the potential of the meta-modelling prototype as a new means for solving computationally expensive antenna design tasks, this case study addresses the common challenge in antenna design of balancing conflicting performance criteria. Specifically, the study focuses on optimizing a dual-reflector antenna system consisting of two rotationally symmetric reflectors, with a main reflector diameter of 1.5 m and a subreflector diameter of 0.225 m (See Figure 2). The system operates at 8 GHz and is modeled in CHAMP 3D [16], providing high accuracy and relatively fast simulation times, with each full-system simulation taking approximately 1 second on a standard laptop.

The optimization task involves shaping the geometry of subreflector using eight design variables, where the primary trade-off to balance is between maximizing gain and minimizing return loss. Traditionally, balancing these conflicting objectives requires single-objective optimization, where the engineer manually adjusts goal weights to reflect the relative importance of each performance criterion. This process is time-consuming and inefficient, as the engineer must perform multiple optimization runs to explore the various trade-offs. Additionally, choosing appropriate weights is non-trivial, as the performance criteria often differ in scale and units, making it challenging to accurately reflect the engineer's preferences. As a result, this approach becomes computationally expensive, time-consuming and the manual approach limits the ability to dynamically explore multiple trade-offs during the optimization process.

In contrast, the meta-modelling prototype overcomes these limitations by integrating GP meta-models with multi-objective optimization. In this case study, the NSGA-II multi-objective algorithm [17] is combined with a meta-model to find a set of optimal solutions. The idea is to only use true function evaluations whenever the meta-model is uncertain in its predictions, thereby significantly reducing computational complexity. Of the 4000 evaluations required by NSGA-II, only 134 needed full simulations, while the remaining 3867 evaluations were handled by the GP meta-model. This approach reduced computational costs by 97%, making the exploration of complex design spaces both feasible and efficient.

The results, as shown in Figure 2, demonstrate the wide range of potential trade-offs between gain and return loss, known as the Pareto front. In a conventional setup, generation of the frontier will often be practically intractable as each data point represents a full single-objective optimisation. In

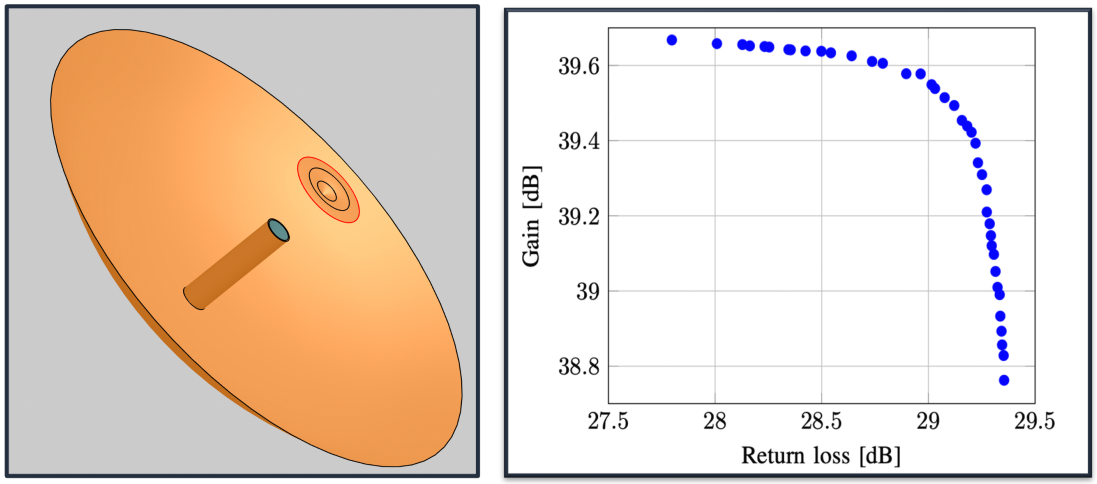

Fig. 2. Right) Axially displaced ring focus dual-reflector system 8 variables for shaping of sub reflector. Left) The Pareto front of optimal trade-offs between gain and return loss.

contrast, the combination of NSGA-II and meta-modelling allows the engineer to perform visual inspection of the frontier, providing a near complete picture of the possible trade-offs. For instance, improving the gain by approximately 0.8 dB can be achieved by accepting a reduction in return loss of around 1 dB.

Overall, this case study highlights the significant potential of the meta-modelling framework to expand the scope of antenna designs that can be explored within practical time frames. By enabling engineers to efficiently navigate high-dimensional design spaces and balance multiple performance criteria with minimal computational overhead, the prototype offers a scalable solution for addressing complex optimization problems in antenna design.

V. Conclusions and Perspectives

In conclusion, the developed meta-modeling prototype represents a robust and efficient alternative to traditional gradient-based and global search optimization methods. By eliminating the need for gradient information, the prototype is well-suited for antenna design problems with expensive-to-evaluate objective functions, where good starting guesses are not known. The Bayesian Optimization (BO) framework employed provides a sample-efficient strategy that outperforms conventional global methods, such as genetic algorithms and particle swarm optimization, by exploring the design space more efficiently and reducing computational costs.

While BO offers significant advantages for computationally expensive, black-box problems, it is important to recognize its situational limitations. In cases where good starting points and derivative information are readily available, a local optimization approach may still be preferable. Similarly, for inexpensive objective functions with well-constrained search spaces, global search methods can be more practical. BO, however, fills a critical gap by addressing antenna design tasks where conventional methods struggle, providing a complementary tool for complex and challenging scenarios.

Moreover, the development of persistent meta-models extends the utility of this framework beyond mere optimization. These models can be reused across different design tasks, allowing for efficient data reuse and the flexibility to modify objectives and constraints without repeating costly simulations. As such, this activity has laid the technical foundation for meta-models that adapt as new data becomes available, setting the stage for broader applications in antenna design within TICRA Tools and paving the way for future advancements of antenna design software.

References

[1] J. Nocedal and S. Wright, *Numerical Optimization*, ser. Springer Series in Operations Research and Financial Engineering. Springer New York, 2006.

[2] J. Kennedy, "Swarm intelligence," in *Handbook of nature-inspired and innovative computing: integrating classical models with emerging technologies*. Springer, 2006, pp. 187–219.

[3] E. Cantu-Paz and D. E. Goldberg, "Efficient parallel genetic algorithms: theory and practice," *Computer methods in applied mechanics and engineering*, vol. 186, no. 2-4, pp. 221–238, 2000.

[4] S. Koziel and S. Ogurtsov, *Antenna Design by Simulation-Driven Optimization*. Springer Publishing Company, Incorporated, 2014.

[5] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global optimization*, vol. 13, pp. 455–492, 1998.

[6] A. Forrester, A. Sobester, and A. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*. Wiley, 2008.

[7] C. E. Rasmussen, C. K. Williams *et al.*, *Gaussian processes for machine learning*. Springer, 2006, vol. 1.

[8] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006. [Online]. Available: /bib/bishop/Bishop2006/Pattern-Recognition-and-Machine-Learning-Christophe-M-Bishop.pdf,/bib/bishop/Bishop2006/978-0-387-31073-2_sm.pdf,https://www.microsoft.com/en-us/research/people/cmbishop/#!prml-book

[9] M. A. Alvarez, L. Rosasco, N. D. Lawrence *et al.*, "Kernels for vector-valued functions: A review," *Foundations and Trends® in Machine Learning*, vol. 4, no. 3, pp. 195–266, 2012.

[10] S. Zhe, W. Xing, and R. M. Kirby, "Scalable high-order gaussian process regression," in *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR, 2019, pp. 2611–2620.

[11] M. D. Morris and T. J. Mitchell, "Exploratory designs for computational experiments," *Journal of Statistical Planning and Inference*, vol. 43, no. 3, pp. 381–402, 1995.
[12] J. Ding, V. Tarokh, and Y. Yang, "Model selection techniques: An overview," *IEEE Signal Processing Magazine*, vol. 35, no. 6, pp. 16–34, 2018.
[13] S. Ament, S. Daulton, D. Eriksson, M. Balandat, and E. Bakshy, "Unexpected improvements to expected improvement for bayesian optimization," *Advances in Neural Information Processing Systems*, vol. 36, pp. 20 577–20 612, 2023.
[14] R. Astudillo and P. I. Frazier, "Bayesian optimization of composite functions," 2019.
[15] G. Matheron, "Principles of Geostatistics," *Economic Geology*, pp. 1246–1266, 1963.
[16] "TICRA Tools CHAMP 3D," https://www.ticra.com/software/champ3d/.
[17] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.