

FREQUENCY-AWARE SURROGATE MODELING WITH SMT KERNELS FOR ADVANCED DATA FORECASTING

N. Gonel^{1,2*}, P. Saves³, J. Morlier⁴

1: Fédération ENAC ISAE-SUPAERO ONERA,
Université de Toulouse,
31000, Toulouse, France.

2: IFP Energies nouvelles,
Institut Carnot IFPEN Transports Energie,
92852 Rueil-Malmaison, France.
nicolas.gonel@ifpen.fr

3: IRIT,
Université de Toulouse, CNRS, Toulouse INP, UT3, UT Capitole,
Toulouse, France.
paul.saves@irit.fr

4: ICA,
ISAE-SUPAERO, INSA, CNRS, MINES ALBI, UPS,
Université de Toulouse,
31400 Toulouse, France.
joseph.morlier@isae-supero.fr

Abstract. *This paper introduces a comprehensive open-source framework for developing correlation kernels, with a particular focus on user-defined and composition of kernels for surrogate modeling. By advancing kernel-based modeling techniques, we incorporate frequency-aware elements that effectively capture complex mechanical behaviors and time-frequency dynamics intrinsic to aircraft systems. Traditional kernel functions, often limited to exponential-based methods, are extended to include a wider range of kernels such as exponential squared sine and rational quadratic kernels, along with their respective first- and second-order derivatives. The proposed methodologies are first validated on a sinus cardinal test case and then applied to forecasting Mauna-Loa Carbon Dioxide (CO₂) concentrations and airline passenger traffic. All these advancements are integrated into the open-source Surrogate Modeling Toolbox (SMT 2.0), providing a versatile platform for both standard and customizable kernel configurations. Furthermore, the framework enables the combination of various kernels to leverage their unique strengths into composite models tailored to specific problems. The resulting framework offers a flexible toolset for engineers and researchers, paving the way for numerous future applications in metamodeling for complex, frequency-sensitive domains.*

Keywords: Frequency correlation kernels, Gaussian process, Aircraft data prediction, Surrogate Modeling Toolbox, Open-Source Software

1 INTRODUCTION

Over the past few decades, the aeronautical industry has witnessed significant advancements driven primarily by incremental improvements in aircraft design optimization [1]. The design process for aeronautic systems is inherently complex and multidisciplinary, requiring a careful balance between various domains such as aerodynamics, structural integrity, geometry, performance, weight management, and dynamic stability [2]. This integrated process, referred to as Multidisciplinary Design Analysis (MDA), lays the foundation for the iterative improvements carried out in Multidisciplinary Design Optimization (MDO) [3, 4]. In our approach to MDO, we adopt the multidisciplinary design feasible strategy, which ensures that the equilibrium between disciplines is re-evaluated and maintained, through a fixed point algorithm, at every modification of the design variables during optimization. However, this approach significantly increases computational demands, posing a major challenge.

To address these challenges, research has increasingly turned to surrogate modeling techniques, which offer an efficient alternative to high-fidelity simulations [5–7]. Software such as the Surrogate Modeling Toolbox (SMT) have been developed to reduce computational costs by leveraging Gaussian Process (GP) interpolation and Bayesian Optimization (BO) methods [8, 9]. However, traditional kernel functions used in GP models—primarily based on exponential formulations—often struggle to capture the complex cyclic and high-dimensional phenomena encountered in advanced aerospace applications [10].

This paper introduces a comprehensive open-source framework for the development of advanced correlation kernels, specifically designed for surrogate modeling in frequency-sensitive environments. Our framework extends conventional methods by incorporating frequency-aware elements into the kernel design. In doing so, we broaden the list of available kernels to include not only standard forms but also alternatives such as exponential squared sine and rational quadratic kernels, along with their first and second-order derivatives. These enhancements enable the efficient optimization of surrogate models that address intricate mechanical behaviors and time-frequency dynamics, which are crucial for analyzing vibrational and aerodynamic characteristics of aircraft structures.

The proposed methodologies are validated through a series of test cases, beginning with a sinus cardinal example and extending to real-world applications such as forecasting Mauna-Loa CO₂ concentrations and airline passenger traffic. All these advancements are encapsulated within SMT 2.0, offering a versatile platform that supports both standard and user-defined kernel configurations. By allowing the combination of multiple kernel types, our approach enables the construction of composite models that leverage the unique strengths of each kernel to address specific modeling challenges.

This extended framework not only enhances predictive capabilities in the preliminary design phase of aerospace engineering but also paves the way for future applications in diverse fields where complex, frequency-sensitive phenomena are of interest. The remainder of the paper is organized as follows: Section 2 reviews the background of surrogate modeling and kernel methods in aerospace applications; Section 3 details the theoretical underpinnings and implementation of our advanced kernel framework; Section 4 presents validation results on benchmark cases and real-world applications; and Section 5 concludes with a discussion of future directions and potential extensions.

2 GAUSSIAN PROCESSES

A Gaussian Process (GP) is a stochastic process, *i.e.* a collection of random variables, where any finite collection of those variables has a joint Gaussian distribution. Hereinafter, we detail the mathematical aspects of GP interpolation, also known as Kriging [11]. First, the GP distribution is completely determined by its mean and covariance functions [12]. Formally, let f denote the (scalar) target function we wish to estimate:

$$\begin{aligned} f &: \mathcal{D} \subseteq \mathbb{R}^d \rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto f(\mathbf{x}) \end{aligned} \quad (1)$$

where \mathbf{x} is a vector of input variables from the input set \mathcal{D} called the *design space*. In the function-space view, we assume that $f(\mathbf{x})$ is a realization of a GP Y with mean function m and covariance function k :

$$Y \sim \mathcal{GP}(m, k) \quad (2)$$

Indeed, the GP defines a probability distribution (called the *prior*) over possible functions that fit the set of observations. The main advantage of the GP modeling is that any finite set of n input points $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \subset \mathcal{D}$, the corresponding observations follow a multivariate Gaussian distribution:

$$\underbrace{\begin{bmatrix} f(\mathbf{x}_1) \\ f(\mathbf{x}_2) \\ \vdots \\ f(\mathbf{x}_n) \end{bmatrix}}_{\text{outputs vector } \mathbf{f}} \sim \mathcal{N} \left(\underbrace{\begin{bmatrix} m(\mathbf{x}_1) \\ m(\mathbf{x}_2) \\ \vdots \\ m(\mathbf{x}_n) \end{bmatrix}}_{\text{mean vector } \mathbf{m}}, \underbrace{\begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & k(\mathbf{x}_1, \mathbf{x}_2) & \cdots & k(\mathbf{x}_1, \mathbf{x}_n) \\ k(\mathbf{x}_1, \mathbf{x}_2)^T & k(\mathbf{x}_2, \mathbf{x}_2) & \cdots & k(\mathbf{x}_2, \mathbf{x}_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{x}_1, \mathbf{x}_n)^T & k(\mathbf{x}_2, \mathbf{x}_n)^T & \cdots & k(\mathbf{x}_n, \mathbf{x}_n) \end{bmatrix}}_{\text{covariance matrix } \mathbf{K}} \right). \quad (3)$$

We can also write it more compact using the shorthands $\mathbf{f} = f(\mathbf{X})$, $\mathbf{m} = m(\mathbf{X})$ and $\mathbf{K} = k(\mathbf{X}, \mathbf{X})$:

$$p(f(\mathbf{x}_1), \dots, f(\mathbf{x}_n)) = p(\mathbf{f} \mid \mathbf{X}) = \mathcal{N}(\mathbf{m}, \mathbf{K}). \quad (4)$$

We highlight the fact that the covariance matrix \mathbf{K} (also called *Gram matrix*) is a Positive Semi-Definite (PSD) and symmetric $n \times n$ matrix [13] whose components are given by the covariance function k : $K_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)$. An example of a valid (stationary) covariance function is the anisotropic¹ Squared Exponential (SE) kernel²:

$$K^{\theta_i}(\mathbf{x}_i, \mathbf{x}_j) = \sigma^2 \prod_{i=1}^n \exp(-\theta_i (\mathbf{x}_i - \mathbf{x}_j)^2). \quad (5)$$

where σ^2 is the process variance parameter and $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_n\}$ is the vector of inverse length-scale parameters. We call these quantities the hyperparameters of the GP that are needed to be estimated in order to best fit the model to our target function.

To make predictions \mathbf{f}_* for a new set of n_* inputs \mathbf{X}_* (independent of \mathbf{X}), we first write the joint distribution of \mathbf{f} and \mathbf{f}_* which also follows a multivariate Gaussian distribution:

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{m} \\ \mathbf{m}_* \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{K}_* \\ \mathbf{K}_*^\top & \mathbf{K}_{**} \end{bmatrix} \right). \quad (6)$$

¹In the isotropic version, θ is scalar so the length-scale is the same for each dimension.

²This kernel is also denoted as the Radial Basis Function (RBF) kernel

where $\mathbf{K}_* = k(\mathbf{X}, \mathbf{X}_*)$ denotes the $n \times n_*$ matrix of covariance between all pairs of training and test points, and similarly $\mathbf{K}_{**} = k(\mathbf{X}_*, \mathbf{X}_*)$ is the covariance matrix of the test set.

In practice, the true function value $f(\mathbf{x})$ is often not directly observable. Instead, we observe a noisy output $y(\mathbf{x}) = f(\mathbf{x}) + \varepsilon$. We assume that this additive noise is an independent and identically distributed (i.i.d.) Gaussian noise [14] and we denote its variance η^2 . Using the previous shorthands and letting \mathbf{y} denotes the vector of noisy observations at inputs \mathbf{X} , the joint distribution (6) then becomes:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{m} \\ \mathbf{m}_* \end{bmatrix}, \begin{bmatrix} \mathbf{K} + \eta^2 \mathbf{I} & \mathbf{K}_* \\ \mathbf{K}_*^\top & \mathbf{K}_{**} \end{bmatrix} \right). \quad (7)$$

Thus, we compute the *posterior* distribution over functions by conditioning the joint prior (7) with the observed data. Using the properties of a multivariate normal distribution, we obtain the predicted distribution:

$$\mathbf{f}_* | \mathbf{y}, \mathbf{X}, \mathbf{X}_* \sim \mathcal{N} \left(\mathbf{m}_* + \mathbf{K}_*^\top [\mathbf{K} + \eta^2 \mathbf{I}]^{-1} (\mathbf{y} - \mathbf{m}), \mathbf{K}_{**} - \mathbf{K}_*^\top [\mathbf{K} + \eta^2 \mathbf{I}]^{-1} \mathbf{K}_* \right). \quad (8)$$

For the ease of notations, we will then only consider centered GPs, *i.e.* zero mean functions: $\mathbf{m} = \mathbf{m}_* = \mathbf{0}$, but equations can be generalized for more general cases with constant (simple Kriging), linear (ordinary Kriging) or polynomial (universal Kriging) priors, adding hyperparameters to the model [12].

Using the usual formalism of GP, we can rewrite the conditional distribution $\mathbf{f}_* | \mathbf{y}$ as:

$$Z_n \triangleq \tilde{Y} \left| \left\{ \tilde{Y}(\mathbf{x}_i) = y(\mathbf{x}_i), 1 \leq i \leq n \right\} \right. \sim \mathcal{GP}(\mu_n, c_n). \quad (9)$$

where the posterior mean and covariance functions are defined as follows:

$$\mu_n(\mathbf{x}) \triangleq k(\mathbf{x}, \mathbf{X}) [\mathbf{K}^{-1} + \eta^2 \mathbf{I}] \mathbf{y}. \quad (10)$$

$$c_n(\mathbf{x}, \mathbf{x}') \triangleq k(\mathbf{x}', \mathbf{x}') - k(\mathbf{x}', \mathbf{X}) [\mathbf{K} + \eta^2 \mathbf{I}]^{-1} k(\mathbf{X}, \mathbf{x}'). \quad (11)$$

An alternative way of formulating the above functions, implemented in Surrogate Modeling Toolbox (SMT) [8, 15], is to introduce the correlation function R and to consider a unique GP variance σ^2 that is an homoscedastic process:

$$\text{Cov}(Y(\mathbf{x}), Y(\mathbf{x}')) = k(\mathbf{x}, \mathbf{x}') = \sigma^2 R(\mathbf{x}, \mathbf{x}'). \quad (12)$$

With this new notation, the equations (10) and (11) can be rewritten as follows [16]:

$$\mu_n(\mathbf{x}) = \mathbf{r}_n^\top(\mathbf{x}) \mathbf{R}_n^{-1} \mathbf{y}. \quad (13)$$

$$c_n(\mathbf{x}, \mathbf{x}') = \sigma^2 [r(\mathbf{x}, \mathbf{x}') - \mathbf{r}_n^\top(\mathbf{x}) \mathbf{R}_n^{-1} \mathbf{r}_n(\mathbf{x}')]. \quad (14)$$

where $\mathbf{y} = [y(\mathbf{x}_1), \dots, y(\mathbf{x}_n)]^\top$ is the observations vector, $\mathbf{r}_n(\mathbf{x}) = k(\mathbf{X}, \mathbf{x}) = (k(\mathbf{x}_i, \mathbf{x}))_{1 \leq i \leq n}$ is the cross-covariance vector and $\mathbf{R}_n = \mathbf{K} + \eta^2 \mathbf{I} = (k(\mathbf{x}_i, \mathbf{x}_j))_{1 \leq i, j \leq n}$ is the covariance matrix. Note that if we consider noise-free observations (*i.e.* $\eta^2 = 0$), we simply have $\mathbf{R}_n = \mathbf{K}$. The conditional process defined by Eq.(12) is then used as a surrogate model to approximate our target function f .

To conclude this short presentation of the GP regression model, we address the optimization of the hyperparameters mentioned above. Indeed, it is necessary to tune the

model parameters $\sigma^2, \boldsymbol{\theta}$ and η^2 on the training set \mathbf{X} before making predictions at new inputs \mathbf{X}_* . Here, everything being Gaussian and our prior being an uniform distribution between given bounds, we only optimize the *marginal likelihood* $p(\mathbf{y})$ to obtain our optimal hyperparameters [17]. This likelihood is given by:

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{0}, \mathbf{K} + \eta^2 \mathbf{I}) = \frac{1}{(2\pi)^{N/2} \det(\mathbf{K} + \eta^2 \mathbf{I})^{1/2}} \exp \left(-\frac{1}{2} \mathbf{y}^\top (\mathbf{K} + \eta^2 \mathbf{I}) \mathbf{y} \right). \quad (15)$$

The optimal set of hyperparameters is then obtained by maximizing the log marginal likelihood:

$$\log p(\mathbf{y}) = -\frac{1}{2} \log (\det (\mathbf{K} + \eta^2 \mathbf{I})) - \frac{1}{2} \mathbf{y}^\top (\mathbf{K} + \eta^2 \mathbf{I})^{-1} \mathbf{y} - \frac{N}{2} \log(2\pi). \quad (16)$$

We recall that the marginal likelihood $p(\mathbf{y})$ is the integral of the *likelihood* $p(\mathbf{y}|\mathbf{f}, \mathbf{X})$ times the prior $p(\mathbf{f}|\mathbf{X})$. Thus, given the prior $\mathbf{f}|\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$ and the likelihood $\mathbf{y}|\mathbf{f} \sim \mathcal{N}(\mathbf{f}, \eta^2 \mathbf{I})$, the computation leads to (15). We saw that to define the GP, we need to define its kernel first. This kernel will give the structure of the GP. We already showed the squared exponential kernel but other kernels exists in SMT such as the absolute exponential, the Matern kernel or the power exponential kernel.

3 A NOVEL USER-DEFINED KERNEL LIBRARY FOR GAUSSIAN PROCESSES

The SMT 2.0 interface is more user-friendly and offers improved and detailed documentation for both users and developers. Hosted publicly, SMT 2.0 can be directly imported within Python scripts, is released under the New BSD license, and runs on Linux, macOS, and Windows operating systems. Furthermore, regression tests are automatically executed on each supported platform whenever changes are committed to the repository, ensuring robustness and compatibility. To enhance the modeling capabilities of Gaussian Processes (GP), we have expanded SMT by implementing two new kernels:

- **Rational Quadratic Kernel:** This kernel is particularly effective for modeling functions that exhibit irregularities. It can be interpreted as a scale mixture of squared exponential kernels, thereby allowing for varying length scales. Its formula is given by:

$$K^{\theta_l, \theta_k}(x, x') = \left(1 + \frac{(x - x')^2}{\theta_l} \right)^{-\theta_k} \quad (17)$$

where θ_l controls the length scale and θ_k adjusts the weighting across scales.

- **Periodic Kernel:** Useful for modeling periodic data, this kernel captures cyclic patterns inherent in many physical and time-series processes. Its formula is defined as:

$$K^{\theta_l, \theta_k}(x, x') = \exp \left(-\frac{\sin^2((x_l - x'_l)\theta_l)}{\theta_k} \right) \quad (18)$$

where θ_l controls the periodicity and θ_k determines the smoothness of the periodic oscillations.

In order to further expand the range of functions that can be modeled with GP, it is essential to leverage the property of kernel composition. This property states that the sum and product of PSD kernels are also valid kernels. Adding kernels models the superposition of independent functions, while multiplying them accounts for interactions between different phenomena [18]. However, SMT previously did not offer a native mechanism for combining kernels, so we introduce a novel architecture that supports flexible kernel composition.

The optimal structure for managing a wide variety of kernel functions is a class-based architecture. To address the need for flexibility and ease of use, we have defined a new class structure within SMT that can accommodate any kernel form. The architecture, illustrated in Fig. 1, is composed of three main elements:

1. **Base Kernel Class:** An abstract class that provides a template for all kernel implementations. Any new kernel can inherit from this class to ensure consistency and compatibility.
2. **Standard Kernel Implementations:** Existing kernels in SMT have been refactored into class-based structures. This includes not only the traditional kernels (such as the squared exponential or Matern kernels) but also the newly added rational quadratic and periodic kernels.
3. **Kernel Composition Utilities:** To simplify the process of creating complex kernels, we provide built-in tools that allow users to combine kernels using basic Python operators (e.g., addition and multiplication) as the product and sum of SDP kernels still results in a kernel [19]. These utilities enable the effortless formation of composite kernels, which can model both the independent superposition and interactions between different data features.

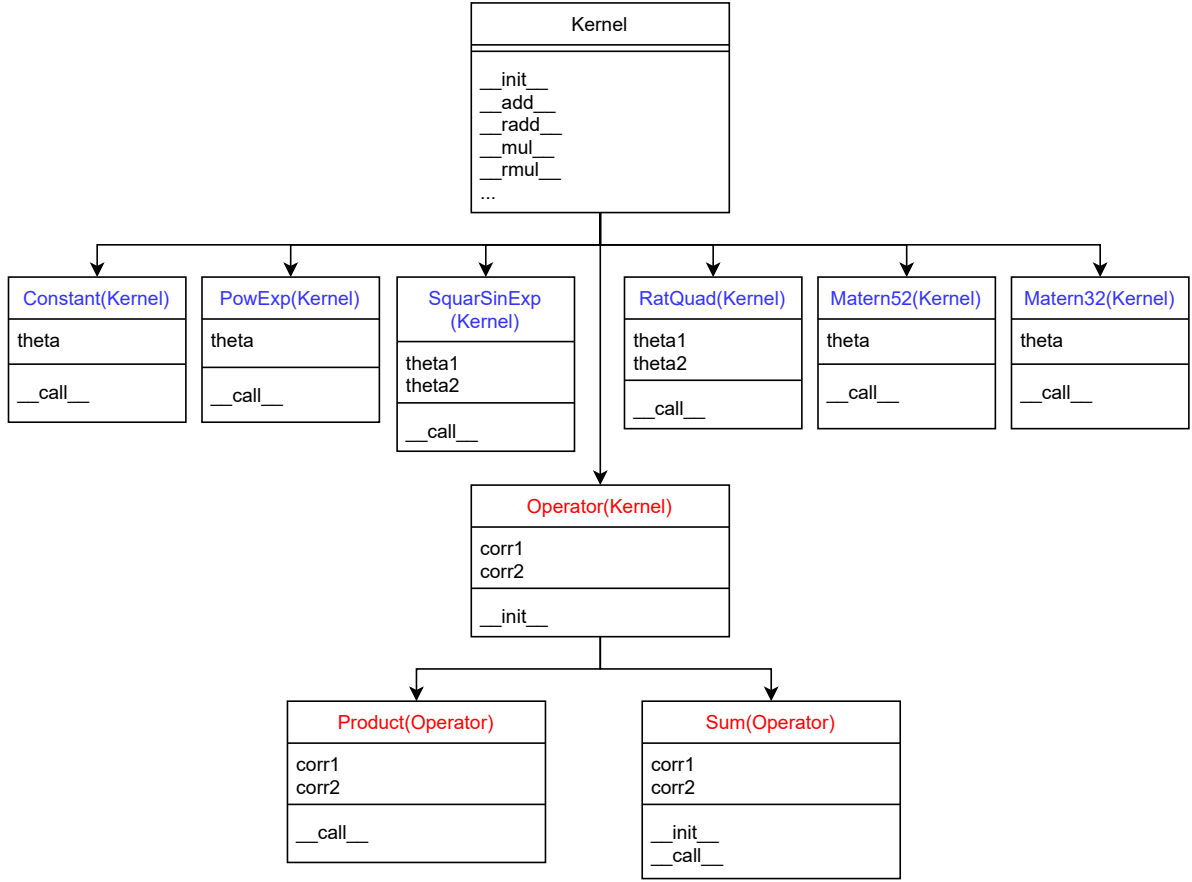


Figure 1: Class diagram for kernels in SMT. In black, the main class, in blue the basic kernels already implemented, and in red the operator.

The new architecture empowers users to construct expressive, hybrid kernels tailored to specific applications. For example, in modeling the oscillatory behavior of an aircraft wing subject to aerodynamic forces, one might combine a periodic kernel to capture cyclic oscillations with a squared exponential kernel to model broader trends. By integrating these advancements, SMT 2.0 significantly extends the flexibility and power of Gaussian Process modeling. This novel kernel library not only supports standard kernel configurations but also facilitates the development of custom, composite models tailored to the unique challenges of complex, frequency-sensitive applications as shown in the following section.

4 RESULTS

To test our framework, we evaluated its ability to model a range of functions and datasets, and assess its capacity to uncover underlying structures and extrapolate beyond the given data. Below, we provide results from different case studies, each with its own characteristics and challenges.

Cardinal Sine function We begin by testing our framework with the well-known cardinal sine (sinc) function. This function, although not strictly periodic, exhibits a periodic component with a varying amplitude. The sinc function is a good candidate for testing, as it has a periodicity that is modified by amplitude fluctuations, which we aim to capture with our kernel model. We employed a kernel formed by the multiplication of a squared exponential (SE) kernel and a periodic kernel:

$$K(x, x') = K_{SE}^{\theta_1}(x, x') \cdot K_{Periodic}^{\theta_2, \theta_3}(x, x') \quad (19)$$

This kernel combination allows us to model both the periodic nature of the function and the amplitude variations. As shown in Fig. 2, the model perfectly captures the periodic structure and the largest amplitude, providing a good representation of the sinc function over the given input domain.

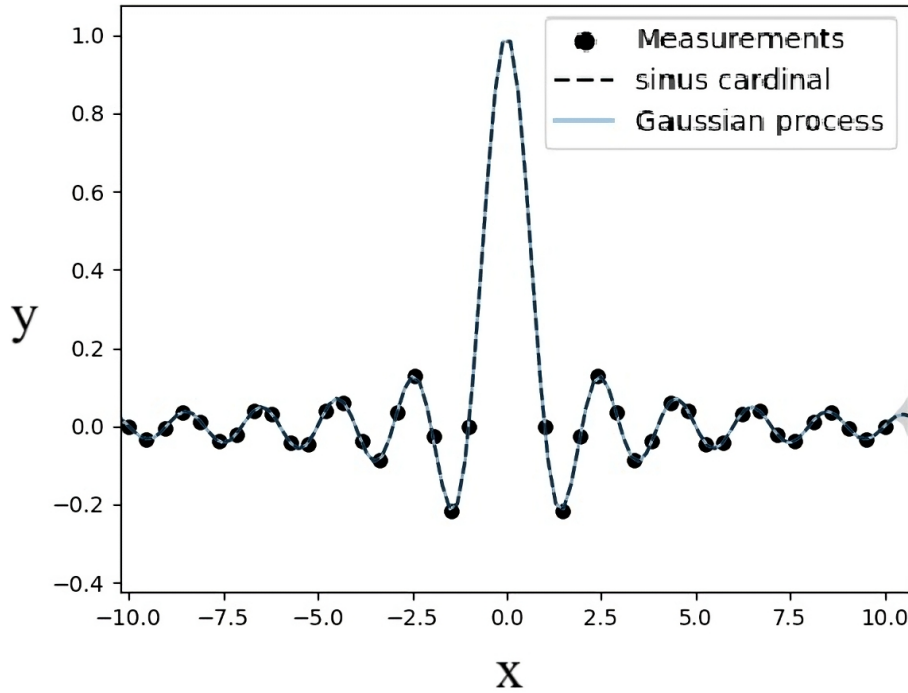


Figure 2: Posterior GP with kernel of Eq.(19) on the sinc function

Mauna-Loa CO₂ observations Next, we apply our framework to the Mauna-Loa CO₂ dataset [20], which contains historical data on the concentration of atmospheric CO₂ since 1958. The Mauna-Loa dataset is often used to test models for environmental data, as it contains a combination of long-term trends, seasonal oscillations, and noise. The primary goal of this case study is to test our model's ability to capture these components, particularly the periodic oscillations and the irregularities introduced by noise.

To model this dataset, we employed a kernel that combines multiple kernel functions to address the different behaviors in the data. Specifically, we use a periodic kernel to model the oscillations, a squared exponential kernel to model the small variations around the periodicity, another squared exponential kernel to account for the linear growth trend,

and a rational quadratic kernel to handle irregularities:

$$K(x, x') = K_{SE}^{\theta_1}(x, x') \cdot K_{Periodic}^{\theta_2, \theta_3}(x, x') + K_{SE}^{\theta_4}(x, x') + K_{rational_quadratic}^{\theta_5, \theta_6}(x, x') \quad (20)$$

The results in Fig. 3 show that the model effectively captures the periodicity and growth, with the interpolation fitting the observed data closely. Additionally, the extrapolation provides reasonable forecasts, with the identified components appearing to repeat in the future, as expected from the underlying structure of the data.

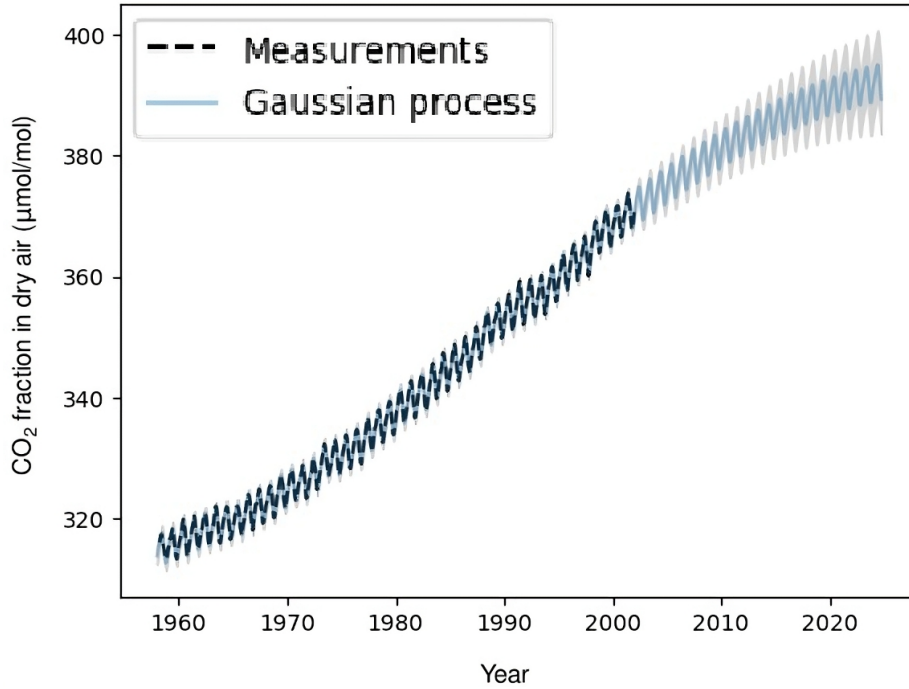


Figure 3: CO₂ forecasting with GP kernel of Eq.(20).

Airlines Annual Passengers Traffic Another case we examine is a dataset containing monthly international airline passenger counts from January 1949 to December 1960 [21]. This dataset is commonly used for testing forecasting models due to its long-term trend, seasonal fluctuations, and sudden changes in direction. Similar to the Mauna-Loa dataset, this data exhibits three primary components: a long-term trend, periodic oscillations, and sharp shifts in direction, such as a sudden increase in the number of passengers during specific periods.

To model this dataset, we apply the same kernel we used for the Mauna-Loa data, given its similar structure:

$$K(x, x') = K_{SE}^{\theta_1}(x, x') \cdot K_{Periodic}^{\theta_2, \theta_3}(x, x') + K_{SE}^{\theta_4}(x, x') + K_{rational_quadratic}^{\theta_5, \theta_6}(x, x') \quad (21)$$

The results shown in Fig. 4 indicate that the kernel successfully captures the underlying structure of the data. The interpolation is smooth and accurate, and the extrapolated

part is plausible, although the variance is higher, possibly reflecting the uncertainty in predicting sudden changes in the future. This suggests that while the model captures the general trends, future data points may have a high degree of uncertainty.

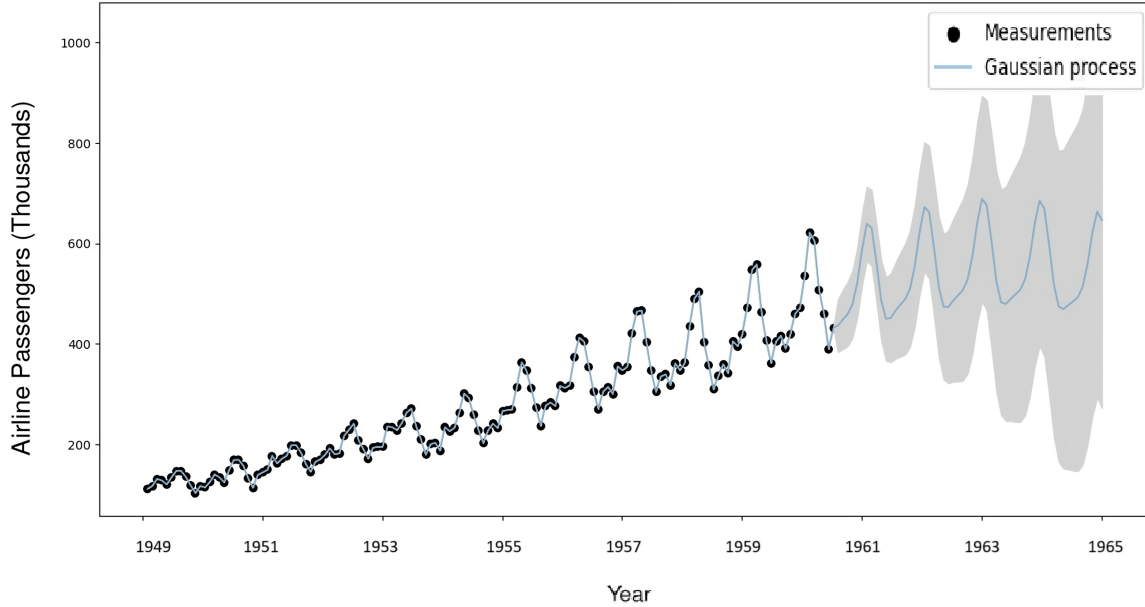


Figure 4: GP approximation with a kernel given by Eq.(21) on passengers database

5 CONCLUSION

In this paper we presented a functional framework to create and combine new kernels for GP in the python library SMT. This expands the scope of functions that can be modeled with this method and also provides a way to use Kriging for prediction and extrapolation with the use of the periodic kernel, which is usually not possible. However, there are still some gray areas. For complex data, finding the right kernel can be a complex endeavor and we can note that all the tests done here are in 1 dimension only, as the complexity increases a lot for complex kernel with the dimension. Indeed the number of hyperparameters in the kernel is multiplied by the dimension making the duration of modelisation quickly prohibitive.

The development of this user-defined kernel library opens several avenues for future work. First this work was limited to sum and product of kernels and can be extended to any PSD matrix manipulation such as ANalysis Of VAriance (ANOVA) [19]. Also, it was limited to spatial classical kernels whereas some more complex ones are better tailored for timeseries forecasting such as the Mercer or spectral kernels [22, 23]. Future updates may include automated techniques for selecting the most appropriate kernel configuration based on data characteristics, utilizing methods such as Bayesian model selection or evolutionary algorithms. Enhancing kernel hyperparameter tuning through gradient-based methods and automatic differentiation will further improve computational efficiency, especially for large-scale problems. The new framework can be directly applied to complex aeroelastic analyses of high aspect ratio wings, such as HALE (High Altitude, Long Endurance) solar aircraft subjected to gust perturbations.

ACKNOWLEDGEMENTS

This work is part of the activities of ONERA - ISAE - ENAC joint research group. The research presented in this paper has been performed within the framework of the COLOSSUS project (Collaborative System of Systems xExploration of Aviation Products, Services and Business Models) and has received funding from the European Union Horizon Programme under grant agreement n°101097120. The authors acknowledge the research project MIMICO funded in France by the Agence Nationale de la Recherche (ANR, French National Research Agency), grant number ANR-24-CE23-0380.

REFERENCES

- [1] A. Chan, A. F. Pires, and T. Polacsek. Trying to elicit and assign goals to the right actors. In *Conceptual Modeling: 41st International Conference, ER 2022*, 2022.
- [2] N. Bartoli, T. Lefebvre, S. Dubreuil, M. Panzeri, R. D’Ippolito, K. Anisomov, and A. Savelyev. Robust nacelle optimization design investigated in the agile european project. In *19th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2018. doi:10.2514/6.2018-3250.
- [3] A. Lambe and J. R. R. A. Martins. A unified description of mdo architectures. In *9th World Congress on Structural and Multidisciplinary Optimization*, 2011.
- [4] J. S. Gray, J. T. Hwang, J. R. R. A. Martins, K. T. Moore, and B. A. Naylor. OpenMDAO: An open-source framework for multidisciplinary design, analysis, and optimization. *Structural and Multidisciplinary Optimization*, 59:1075–1104, 2019. doi:10.1007/s00158-019-02211-z.
- [5] P. Saves, Y. Diouane, N. Bartoli, T. Lefebvre, and J. Morlier. A general square exponential kernel to handle mixed-categorical variables for gaussian process. In *AIAA Scitech 2022 Forum*, 2022.
- [6] A. F. Hernandez and M. A. Grover. Stochastic dynamic predictions using gaussian process models for nanoparticle synthesis. *Computers & Chemical Engineering*, 34(12):2042–2052, 2010.
- [7] N. R. Costa and J. Lourenço. Gaussian process model – an exploratory study in the response surface methodology. *Quality and Reliability Engineering International*, 32(3):935–944, 2016.
- [8] P. Saves, R. Lafage, N. Bartoli, Y. Diouane, J. H. Bussemaker, T. Lefebvre, J. T. Hwang, J. Morlier, and J. R. R. A. Martins. SMT 2.0: A surrogate modeling toolbox with a focus on hierarchical and mixed variables Gaussian processes. *Advances in Engineering Software*, 188:1–15, 2024.
- [9] P. Saves, Y. Diouane, N. Bartoli, T. Lefebvre, and J. Morlier. A mixed-categorical correlation kernel for gaussian process. *Neurocomputing*, 532:556–568, 2023.
- [10] R. Priem, Y. Diouane, N. Bartoli, S. Dubreuil, and P. Saves. High-dimensional bayesian optimization using both random and supervised embeddings. *AIAA Journal*, 63(1):162–173, 2025.

- [11] D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6):119–139, 1951.
- [12] C. K. Williams and C. E. Rasmussen. *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006.
- [13] H. Lee. *Gaussian Processes*, chapter 5, pages 575–577. Springer Berlin Heidelberg, 2011.
- [14] R. Conde Arenzana, A. López-Lopera, S. Mouton, N. Bartoli, and T. Lefebvre. Multi-fidelity Gaussian process model for CFD and wind tunnel data fusion. In *ECCOMAS Aerobest*, 2021.
- [15] M. A. Bouhlel, J. T. Hwang, N. Bartoli, R. Lafage, J. Morlier, and J. R. R. A. Martins. A python surrogate modeling framework with derivatives. *Advances in Engineering Software*, 135, 2019. doi:10.1016/j.advengsoft.2019.03.005.
- [16] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.
- [17] R. Rossi. *Mathematical statistics: an introduction to likelihood based inference*. John Wiley & Sons, 2018.
- [18] D. Duvenaud, J. Lloyd, R. Grosse, J. Tenenbaum, and G. Zoubin. Structure discovery in nonparametric regression through compositional kernel search. In *International Conference on Machine Learning*, pages 1166–1174. PMLR, 2013.
- [19] O. Roustant, E. Padonou, Y. Deville, A. Clément, G. Perrin, J. Giorla, and H. Wynn. Group kernels for gaussian process metamodels with categorical inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):775–806, 2020.
- [20] C. D. Keeling. The influence of mauna loa observatory on the development of atmospheric co2 research. In *Mauna Loa Observatory. A 20th Anniversary Report. (National Oceanic and Atmospheric Administration Special Report*, pages 36–54, 1978.
- [21] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [22] A. Wilson and R. Adams. Gaussian process kernels for pattern discovery and extrapolation. *International Conference on Machine Learning*, 2013.
- [23] A. Gorodetsky and Y. Marzouk. Mercer kernels and integrated variance experimental design: connections between gaussian process regression and polynomial approximation. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):796–828, 2016.