MECSOL 2022 - Proceedings of the 8th International Symposium on Solid Mechanics
M.L. Bittencourt, J. Labaki, L.C.M. Vieira Jr. and E. Mesquita (Editors), Campinas SP, Brazil, October 17th to 19th, 2022

Surrogate-Based Optimization of Functionally Graded Plates using Multi-Fidelity Models

Leonardo Gonçalves Ribeiro[1], Evandro Parente Junior[1] and Antônio Macário Cartaxo de Melo[1]

[1] Laboratório de Mecânica Computacional e Visualização, Departamento de Engenharia Estrutural e Construção Civil, Universidade Federal do Ceará, Fortaleza-CE, Brasil

Abstract. Optimization methods can be employed to find the optimal material designs in Functionally Graded (FG) structures. This is often performed by the use of bio-inspired algorithms, even though these may require a large number of function evaluations. For a more efficient process, surrogate models can be used to provide a cheaper estimate for the structural response. Also, on structural optimization of complex structures, analysis models with multiple levels of fidelity (via coarser mesh discretization or simplification of analysis theory) can be easily created. Thus, Multi-Fidelity models can be employed for a more accurate approximation. To further increase the optimization process effectiveness, the Sequential Approximate Optimization (SAO) can be employed, where the approximate surface is iteratively improved by the addition of new points in regions of interest. In this work, SAO will be employed in the optimization of Functionally Graded Plates. The multi-fidelity Hierarchical Kriging model will be employed. For comparison purposes, results using the single-fidelity Kriging model will be shown. These approaches will be compared in terms of efficiency and accuracy. Results show that the Hierarchical Kriging can greatly reduce the number of expensive evaluations required to find the optimal material gradation.
Keywords: Surrogate-Based Optimization, Functionally Graded Plates, Multi-Fidelity Models, Adaptive sampling.

INTRODUCTION

Functionally Graded Materials are a class of composites where material properties change smoothly over the structure domain (Shen, 2009). This allows for efficient designs while also avoiding disadvantages such as stress concentrations or delamination, which are common in the design of laminate structures.

Studies over Functionally Graded Plates (FGPs) have been gaining interest in the last few years. Regarding these structures, material gradation can be given in one direction (e.g. thickness direction) or in multiple directions. An appropriate scheme to define this gradation is usually chosen beforehand, such as univariate or trivariate B-Spline functions (Ribeiro *et al.*, 2020).

For an efficient design, one may perform the optimization of such structures. This can be done by the use of bio-inspired algorithms, such as Particle Swarm Optimization (PSO). However, these usually require hundreds or even thousands of function evaluations. Thus, the computational cost involved in the entire process might be too high, especially when numerical methods (e.g. FEM or IGA) are used for structural analyses. For a more efficient process, one may use Surrogate-Based Optimization (Forrester, Sobester, and Keane, 2008). Based on a small number of data points, the expensive response is approximated by a surrogate model. The model is then able to provide an approximate assessment of the true function, and it is used to guide the optimization process. For instance, Do, Lee, and Lee (2019) and Do, Nguyen-Xuan and Lee (2020) have employed a Deep Neural Network (DNN) to fit an approximated surface for the buckling load and fundamental frequency of FGPs. The DNN was then used to conduct a very efficient optimization process.

When a fixed surrogate is used, the user relies entirely on the model accuracy to find the optimum response. However, it might be important to add new data points in regions of interest, so that, our model becomes more accurate near the optimum. To that end, an adaptive sampling scheme may be adopted, where new data points are defined by the optimization of a chosen acquisition function. Ribeiro *et al.* (2020) and Maia *et al.* (2021) successfully used Bayesian models (RBF and Kriging, respectively) for optimization of FGPs.

However, when the analysis model is complex, single fidelity approaches may still present a high processing time. This is especially relevant for high-dimensional problems, where a high number of data points are required to provide sufficient accuracy. Thus, Multi-Fidelity Models can be used to improve the computational efficiency by evaluating some data points using a lower fidelity source (Han, 2012). This source may be derived using a coarser mesh or a simplified analysis theory. This allows for a better exploration of the design space, without requiring a large number of expensive analyses.

In this work, Hierarchical Kriging will be used for the Surrogate-Based Optimization of FGPs (Han, 2012). Adaptive sampling will be performed using adequate acquisition functions, such as Expected Improvement (EI), Probability of Improvement (PI), and variable-fidelity versions of such methods. For comparison purposes, results will also be shown for a single-fidelity Kriging model. The obtained results will be compared in terms of accuracy, the number of evaluations, and computational efficiency.

FUNCTIONALLY GRADED PLATES

Functionally Graded Plates are composite structures made of two materials, usually a ceramic c and a metal m, whose material composition changes inside the structure domain. In this work, we perform the optimization of unidirectional FGPs, where material varies only in the structure thickness. To enhance design flexibility, B-Spline functions can be employed to define the material gradation (Ribeiro *et al.*, 2020):

$$V_c(\xi) = \sum_{i=1}^{n_{cp}} B_{i,p}(\xi)\, V_{c,i}, \quad V_m(\xi) = 1 - V_c(\xi) \quad \xi \in [0,1] \tag{1}$$

where n_{cp} is the number of control points, $V_{c,i}$ is the fraction of the ceramic volume of the i-th control point, $B_{i,p}(\xi)$ is the corresponding B-Spline base, p is the base degree and ξ is the parametric coordinate. The B-Spline base is evaluated by the Cox-de Boor formula using a given a knot vector $\Xi = [\xi_1, \xi_2, ..., \xi_{n+p+1}]$.

Effective material properties at a given point of the structure can be performed using adequate homogenization schemes. In this work, the Mori-Tanaka scheme will be employed. Here, the structure, made of a material m, is assumed to be reinforced by spherical particles from material c (Shen, 2009). Mechanical properties can then be evaluated by:

$$\frac{K(z) - K_m}{K_c - K_m} = \frac{V_c(z)}{1 + \dfrac{3\,V_m(z)\,(K_c - K_m)}{3\,K_m + 4\,G_m}} \quad \text{and} \quad \frac{G(z) - G_m}{G_c - G_m} = \frac{V_c(z)}{1 + \dfrac{V_m(z)\,(G_c - G_m)}{G_m + f_m}} \tag{2}$$

where:

$$f_m = \frac{G_m\,(9\,K_m + 8\,G_m)}{G_m + f_m} \tag{3}$$

The Young's modulus (E) and Poisson's ratio (ν) are computed according to standard expressions (Shen, 2009).

SURROGATE MODELING

Structural responses of FGPs are often evaluated using numerical methods of analysis (such as Isogeometric Analysis). However, these may become very expensive, especially for complex structures with many degrees of freedom. In that context, surrogate models can provide an approximate response surface with a much faster evaluation cost.

This approximate response surface can be fitted based on a small set of sampling points. If no prior information about the true function's behavior is available, a Design of Experiments (DoE) technique can be employed to pick the initial sampling points. For instance, the Latin Hypercube Sampling (LHS) is a stratified random sampling technique commonly employed for Surrogate-Based Optimization.

In this work, we aim at using Multi-Fidelity models, in particular the Hierarchical Kriging, to approximate an expensive structural response. For comparison purposes, results will also be shown for the single-fidelity Kriging model. In the following, we explain how both these models can be fitted based on a set of data points.

Kriging

The Kriging model is given by the sum of a global trend g and its autocorrelated localized deviations Z (Forrester, Sobester and Keane, 2008):

$$\hat{y}(\mathbf{x}) = g(\mathbf{x}) + Z(\mathbf{x}) \tag{4}$$

For the Ordinary Kriging, $g(\mathbf{x}) = \mu$ is constant, where μ is the process mean. The localized deviations $Z(\mathbf{x})$ can be assumed as the realization of a stochastic process with mean zero and covariance:

$$\mathrm{cov}(\mathbf{y}, \mathbf{y}) = \sigma^2\, \boldsymbol{\Psi} \tag{5}$$

where σ^2 is the process variance and $\mathbf{\Psi}$ is a correlation matrix:

$$\Psi_{ij} = \text{cor}[y(\mathbf{x}_i), y(\mathbf{x}_j)] \tag{6}$$

Correlation between the responses is given by a measure of similarity. Usually, the Gaussian correlation function is employed:

$$\text{cor}\left[y(\mathbf{x}_i), y(\mathbf{x}_j)\right] = \exp\left(-\sum_{k=1}^{m} \theta_k \left|x_{i,k} - x_{j,k}\right|^{p_k}\right) \tag{7}$$

The mean and the variance can be defined using the Maximum Likelihood Estimator (MLE). For a given data set $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, the likelihood of a model with mean μ and variance σ is given by:

$$L(\mathbf{y}|\mu,\sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}\left|\mathbf{\Psi}\right|^{1/2}} \exp\left[-\frac{(\mathbf{y}-\mathbf{1}\mu)^T\,\mathbf{\Psi}^{-1}\,(\mathbf{y}-\mathbf{1}\mu)}{2\sigma^2}\right] \tag{8}$$

By differentiation, we obtain the MLE for μ and σ as:

$$\hat{\mu} = \frac{\mathbf{1}^T\mathbf{\Psi}^{-1}\mathbf{y}}{\mathbf{1}^T\mathbf{\Psi}^{-1}\mathbf{1}} \quad \text{and} \quad \hat{\sigma}^2 = \frac{(\mathbf{y}-\mathbf{1}\hat{\mu})^T\mathbf{\Psi}^{-1}(\mathbf{y}-\mathbf{1}\hat{\mu})}{n} \tag{9}$$

To define the MLE for θ_k and p_k, however, we need to use an appropriate optimization method. For simplification, most researchers often consider $p_k = 2.0$. Then, optimization is performed considering only the parameter θ_k (Forrester, Sobester, and Keane, 2008). Substituting Eq. (9) in Eq. (8), taking the natural logarithm, and removing the constant terms, we end up with the concentrated ln-likelihood function:

$$\ln L \approx -\frac{n}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln|\mathbf{\Psi}| \tag{10}$$

Then, the definition of θ_k can be performed by solving the optimization problem:

$$\begin{cases} \text{minimize} & -\ln L(\boldsymbol{\theta}) \\ \text{where} & \boldsymbol{\theta}_L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_U \end{cases} \tag{11}$$

Since this likelihood function is highly multi-modal, in this work we use the Particle Swarm Optimization (PSO) algorithm to solve this problem.

After the model is built, we can evaluate the model prediction using the Kriging estimator (Forrester, Sobester and Keane, 2008):

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \boldsymbol{\psi}^T\,\mathbf{\Psi}^{-1}\,(\mathbf{y}-\mathbf{1}\hat{\mu}) \tag{12}$$

The Kriging model also allows for the assessment of the posterior variance:

$$\hat{s}^2(\mathbf{x}) = \sigma^2\left(1 - \boldsymbol{\psi}^T\mathbf{\Psi}^{-1}\boldsymbol{\psi}\right) \tag{13}$$

which can be interpreted as an estimate of the model uncertainty on its own prediction. Figure 1 shows the prediction and variance of a Kriging model. Note that, using only 3 data points, the model is not able to accurately represent the true function. As expected, model variance is very high in regions where the prediction error is higher.

Hierarchical Kriging

The Hierarchical Kriging was proposed by Han and Gortz (2012) as a powerful Multi-Fidelity model with a simple formulation. Here, alongside the High-Fidelity (HF) data, given by n_h points (located at $\mathbf{x}_h$) evaluated using a higher fidelity source, the model is able to consider Low-Fidelity (LF) data, given by n_l points (located at $\mathbf{x}_l$) evaluated using a lower fidelity source. This allows for a much better exploration of the design space at a lower cost.

For the Hierarchical Kriging, the trend term is given by a Kriging model built based on a set of LF data:

$$\hat{y}(\mathbf{x}) = \beta_0\,\hat{y}_l(\mathbf{x}) + Z(\mathbf{x}) \tag{14}$$

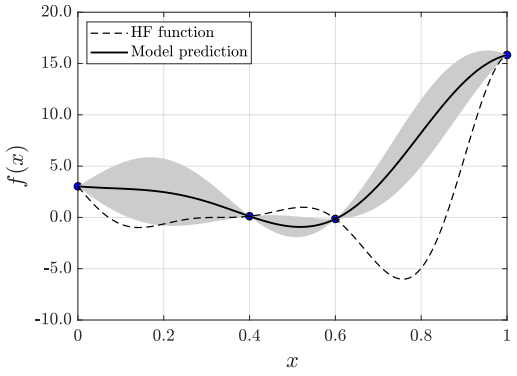


Figure 1: Kriging model prediction and variance ($\hat{y} \pm \hat{s}^2$).

where, again, $Z(\mathbf{x})$ can be assumed to come from a stochastic process with mean zero and covariance given by:

$$\text{cov}(\mathbf{y},\mathbf{y}) = \sigma^2 \boldsymbol{\Psi} \tag{15}$$

Hierarchical Kriging model building starts by fitting a Kriging model to the LF data $\mathcal{D} = \{(\mathbf{x}_{l,i}, y_{l,i})\}_{l,i=1}^{n}$. Thus, we are able to evaluate $\hat{y}_l$ and $\hat{s}_l$ by Eqs. (12) and (13). Then, building of the Hierarchical Kriging model itself can be performed by maximizing its corresponding ln-likelihood function:

$$\ln L = -\frac{n_h}{2}\ln(2\pi) - \frac{n_h}{2}\ln(\sigma^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}(\mathbf{x}_h,\mathbf{x}_h)| - \frac{(\mathbf{y}_h - \beta_0\mathbf{F})^T\boldsymbol{\Psi}(\mathbf{x}_h,\mathbf{x}_h)^{-1}(\mathbf{y}_h - \beta_0\mathbf{F})}{2\sigma^2} \tag{16}$$

where:

$$\mathbf{F} = \hat{y}_l(\mathbf{x}_h) \tag{17}$$

By differentiation, MLEs for β_0 and σ^2 can be defined as:

$$\hat{\beta}_0 = \frac{\mathbf{F}^T\boldsymbol{\Psi}^{-1}\mathbf{y}_h}{\mathbf{F}^T\boldsymbol{\Psi}^{-1}\mathbf{F}} \quad \text{and} \quad \hat{\sigma}^2 = \frac{(\mathbf{y}_h - \beta_0\mathbf{F})^T\boldsymbol{\Psi}(\mathbf{x}_h,\mathbf{x}_h)^{-1}(\mathbf{y}_h - \beta_0\mathbf{F})}{n_h} \tag{18}$$

Substituting Eqs. (18) in Eq. (16), the concentrated ln-likelihood function for the Hierarchical Kriging model can be derived:

$$\ln L \approx -\frac{n_h}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}(\mathbf{x}_h,\mathbf{x}_h)| \tag{19}$$

Then, the rest of hyper-parameters (θ_k) can be evaluated solving the optmization problem:

$$\begin{cases} \text{minimize} & -\ln L(\boldsymbol{\theta}) \\ \text{where} & \boldsymbol{\theta}_L \le \boldsymbol{\theta} \le \boldsymbol{\theta}_U \end{cases} \tag{20}$$

which is very similar to how the Kriging model is built, but now the likelihood function is slightly different, as it considers the influence of the non-constant trend term based on the LF source.

After being built, the Hierarchical Kriging model can be evaluated by:

$$\hat{y}(\mathbf{x}) = \hat{\beta}_0\hat{y}_l(\mathbf{x}) + \boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1}(\mathbf{y}_h - \hat{\beta}_0\mathbf{F}) \tag{21}$$

where $\boldsymbol{\psi}$ is given by:

$$\psi_i(\mathbf{x}) = \exp\left(-\sum_{k=1}^{m}\theta_k|x_{i,k} - x_k|^{p_k}\right) \tag{22}$$

The estimate of the posterior variance for Hierarchical Kriging is given by:

$$\hat{s}^2(\mathbf{x}) = \hat{\sigma}^2\left[1 - \boldsymbol{\psi}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\psi} + \frac{(\hat{y}_l - \mathbf{F}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\psi})^2}{\mathbf{F}^T\boldsymbol{\Psi}^{-1}\mathbf{F}}\right] \tag{23}$$

Hierarchical Kriging also allows for the assessment of the LF model y_h variance, which can be performed similar to a regular Kriging model:

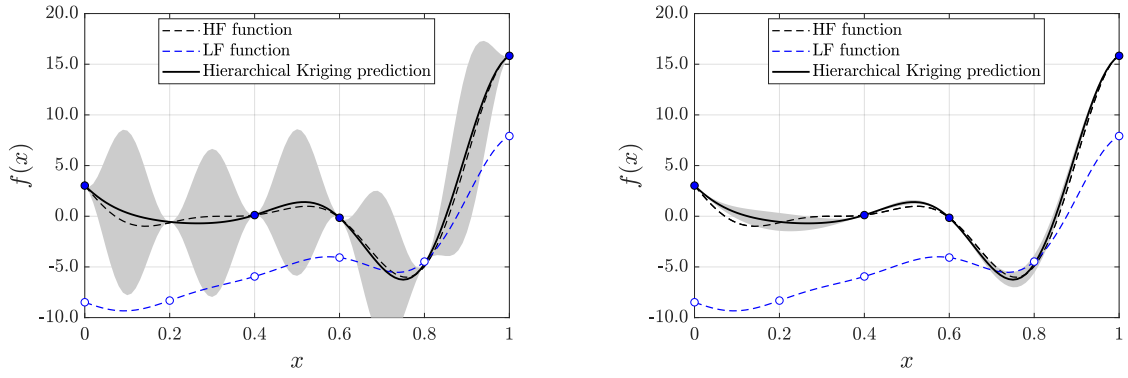
$$\hat{s}_l^2(\mathbf{x}) = \hat{\sigma}_l^2 \left(1 - \boldsymbol{\psi}_l^T \boldsymbol{\Psi}_l^{-1} \boldsymbol{\psi}_l\right) \tag{24}$$

This way, Zhang, Han and Zhang (2018) proposed the assessment of a Variable-Fidelity (VF) posterior variance, given by:

$$\hat{s}^2(\mathbf{x}, f) = \begin{cases} \beta_0^2 \hat{s}_l^2(\mathbf{x}) & \text{, if } f = 1 \\ \hat{s}^2(\mathbf{x}) & \text{, if } f = 2 \end{cases} \tag{25}$$

where $f = 1$ is related to the LF source, and $f = 2$ is related to the HF source. The VF variance is of great importance for the selection and addition of new data points, which will be discussed in the next section.

Figure 2 shows the prediction and variance of a Hierarchical Kriging model. Here, results are shown for the VF variance considering both fidelity sources. Note that, this time, by the consideration of a small number of LF data, the model is able to represent the true function much better.



(a) $\hat{y} \pm \hat{s}^2(\mathbf{x}, 1)$ (b) $\hat{y} \pm \hat{s}^2(\mathbf{x}, 2)$

Figure 2: Prediction and variance for the Hierarchical Kriging model.

SEQUENTIAL APPROXIMATE OPTIMIZATION

For a more accurate process, one may employ the Sequential Approximate Optimization (SAO). Here, an adaptive sampling technique is used, and new promising designs are added to the model during the optimization. The location of such designs can be found by the optimization of an acquisition function, which often depends on model prediction ($\hat{y}$) and variance ($\hat{s}$). Figure 3 shows how SAO is performed iteratively.

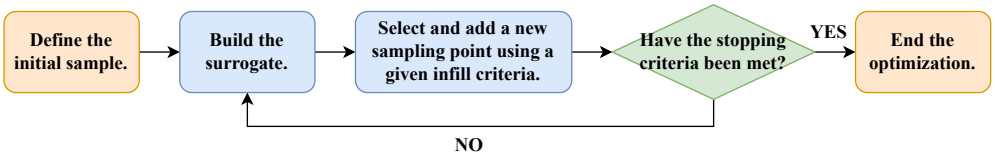
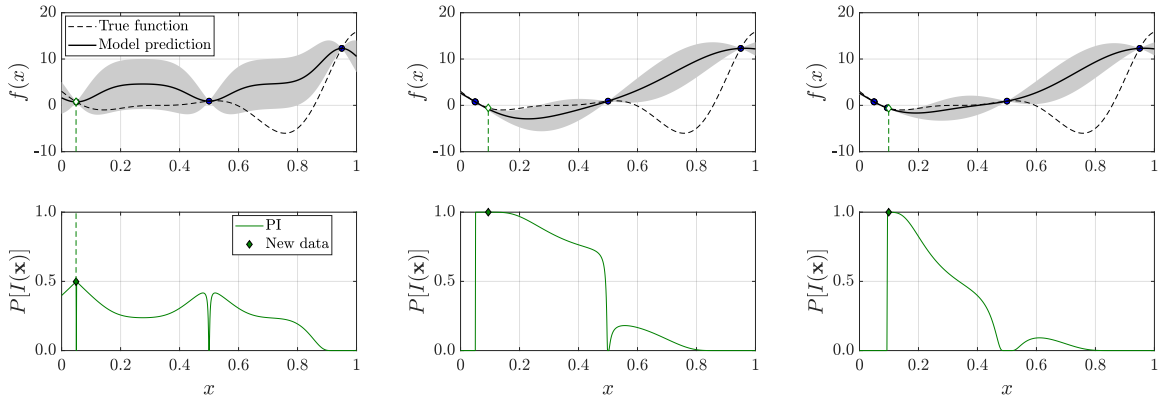


Figure 3: Flowchart for SAO process.

To that end, Kushner (1964) proposed the Probability of Improvement (PI) criterion, which maximizes the probability that a given trial design improves upon the current optimum. The PI can be evaluated by:

$$P[I(\mathbf{x})] = \Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) \tag{26}$$

where Φ is the Cumulative Distribution Function (CDF) for the Normal distribution.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

Figure 4: Probability of Improvement (PI) using the Kriging model.

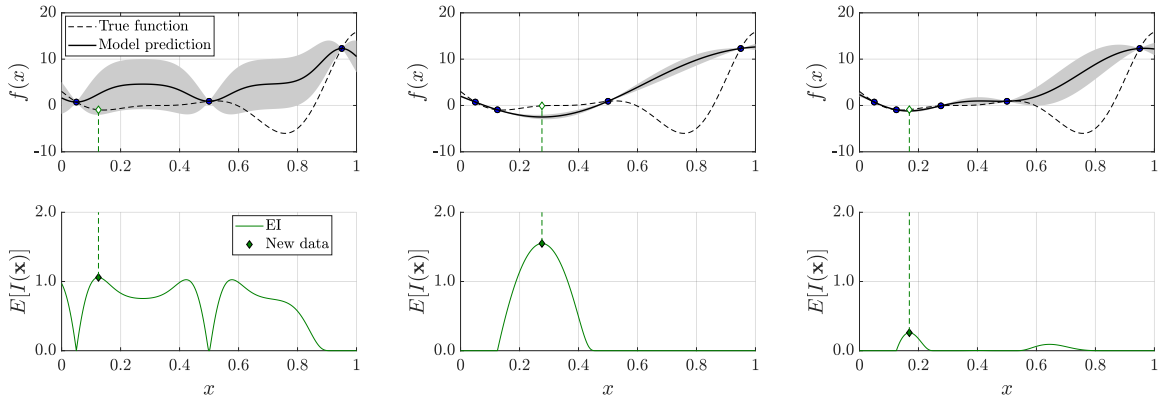
Figure 4 shows how the PI works for the first three iterations of a test problem. Here, the single-fidelity Kriging model is being used. Unfortunately, the PI heavily favors exploitation, and it is usually not able to find promising unexplored regions. Furthermore, it may present large constant regions, where $P[I(\mathbf{x})] = 0.0$ or $P[I(\mathbf{x})] = 1.0$.

Alternatively, based on the theory of expected utility, Mockus (1974) proposed the Expected Improvement (EI) criterion. This method considers not only the probability that the response is an improvement, but also the magnitude of this improvement. The approach was popularized by Jones, Schonlau, and Welch (1998) after the EI was employed in the well-known Efficient Global Optimization (EGO) algorithm. The EI is given by:

$$E[I(\mathbf{x})] = (y_{min} - \hat{y}(\mathbf{x}))\,\Phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) + \hat{s}(\mathbf{x})\,\phi\left(\frac{y_{min} - \hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) \tag{27}$$

where Φ and ϕ are the Cumulative Distribution Function (CDF) and the Probability Density Function (PDF) for the Normal distribution, respectively.

Figure 5 shows how the EI works for the first three iterations of a test problem. Again, the single-fidelity Kriging model is considered. Note that, while the method was not able to find the optimum region in the first three iterations, it does select new trial designs in regions with higher uncertainty. Thus, we may expect that, in the following iterations, the method would be able to find the optimum.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

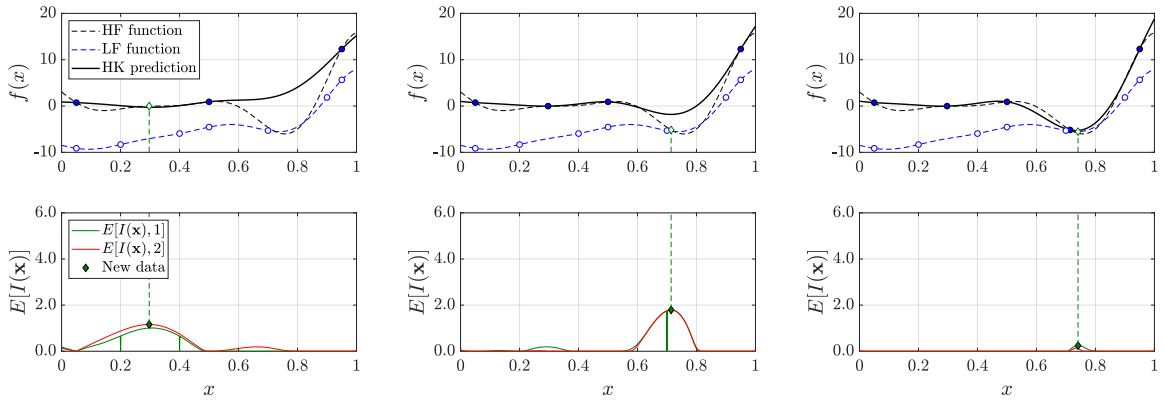
Figure 5: Expected Improvement (EI) using the Kriging model.

Variable-Fidelity approaches

The regular PI and EI can also be employed for Multi-Fidelity (MF) models, as long as model prediction and variance are evaluated accordingly. However, in these cases, along with the definition of the new data location, one should also define the new data fidelity. Using the Hierarchical Kriging model, one may employ Variable-Fidelity (VF) methods to assist with this issue. Here, the VF variance, shown in Eq. (25), is used to evaluate the acquisition function. Then, the source which presents the best acquisition function value is used to evaluate the new trial design. These methods were first used by Zhang, Han, and Zhang (2018), who proposed the VF-EI:

$$E[I(\mathbf{x}),f] = (y_{min} - \hat{y}(\mathbf{x}))\,\Phi\left(\frac{y_{min}-\hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x},f)}\right) + \hat{s}(\mathbf{x},f)\,\phi\left(\frac{y_{min}-\hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x},f)}\right) \tag{28}$$

where $f = 1$ is related to the LF source, and $f = 2$ is related to the HF source. Figure 6 shows how the VF-EI works for the first three iterations of a test problem. Here, the method is able to find the optimum region in the second iteration. A design very close to the optima is then found in the third iteration.



(a) Iteration 1 (b) Iteration 2 (c) Iteration 3

Figure 6: Variable-Fidelity Expected Improvement (VF-EI) using the Hierarchical Kriging model.

Later, Ruan *et al.* (2020) also proposed the VF-PI:

$$P[I(\mathbf{x}),f] = \Phi\left(\frac{y_{min}-\hat{y}(\mathbf{x})}{\hat{s}(\mathbf{x},f)}\right) \tag{29}$$

The authors suggest that the method presents good results when correlation between sources is high.

NUMERICAL EXAMPLE

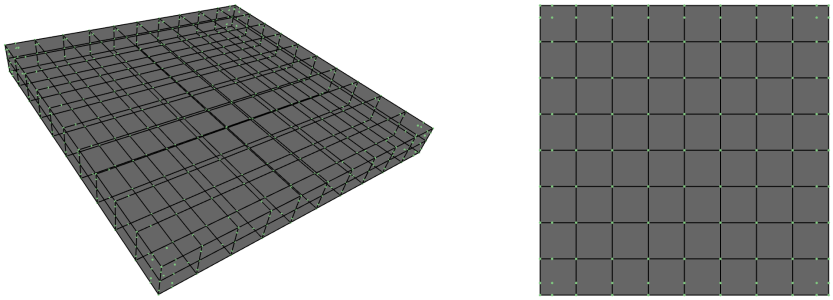
In this section, the maximization of the buckling load of a simply-supported Al/Al_2O_3 square plate will be performed. Material gradation is defined by a symmetric B-Spline with 9 control points, and the design variables are the volume fraction for each control point. Due to symmetry, there are only five design variables. A constraint will be imposed on the maximum ceramic percentage. The optimization problem can be written as:

$$\begin{cases} \text{find} & \mathbf{x} = \{x_1, x_2, \ldots, x_5\} \\ \text{that minimizes} & -\lambda_{norm}(\mathbf{x}) \\ \text{subject to} & g_1(\mathbf{x}) = \dfrac{\overline{V}_c(\mathbf{x})}{\overline{V}_{c,max}} - 1 \leq 0 \\ \text{with} & 0 \leq x_i \leq 1 \end{cases} \tag{30}$$

where the maximum ceramic percentage $\overline{V}_{c,max} = 35\%$. Table 1 presents the properties of each constituent, and effective properties are defined via the Mori-Tanaka model.

Table 1: Material properties.

Material	E (GPa)	ν
Al	70.00	0.30
Al_2O_3	380.0	0.30



(a) HF source. (b) LF source.

Figure 7: Mesh used for each source.

Here, Isogeometric Analysis (IGA) is performed to assess numerical responses. The HF source is given by a 3D model using a $10 \times 10 \times 2$ cubic NURBS mesh, while the LF source is given by an 8×8 cubic NURBS mesh. These two meshes are depicted in Figure 7.

Figure 8 shows the correlation between responses. Here, the correlation factor is very high ($R^2 = 0.99$), and the average relative difference between sources is close to 2.27%. The 3D model offers slightly lower critical buckling load factors than those obtained via the 2D model. Thus, the LF source, although cheaper, may lead to less conservative designs when compared to the HF source. Each HF evaluation takes, on average, 28.75 s, while the LF evaluation takes only 0.64 s. Thus, the relative expense is $C_r = C_l/C_h = 2.2\%$.

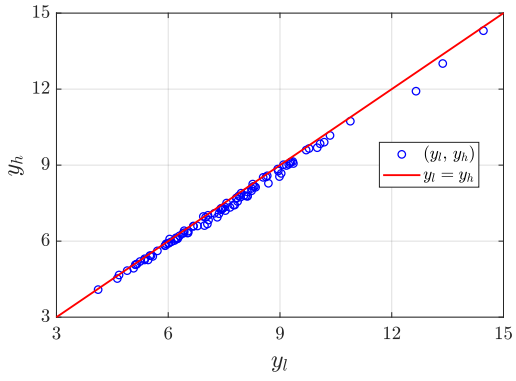
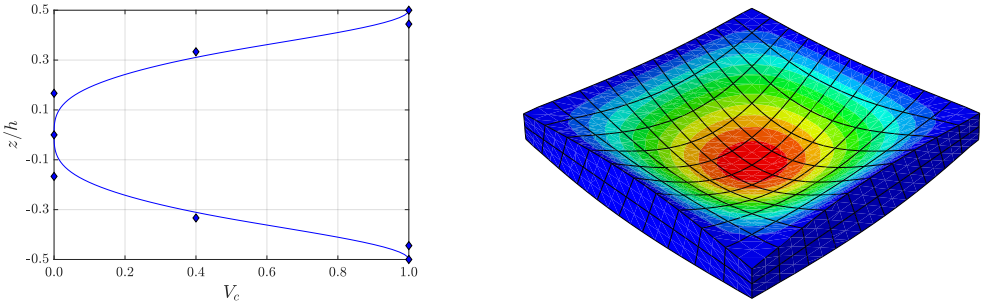


Figure 8: Correlation between sources for the buckling load maximization of the square FGP.

The optimum of the problem is found at $\mathbf{x} = [1.0, 1.0, 0.4, 0.0, 0.0]$. Figure 9 shows the optimum gradation, as well as the buckling mode for the HF model.

This problem was first proposed by Do, Lee, and Lee (2019). The authors used 100,000 data points to fit and validate a Deep Neural Network (DNN) and, then, optimized the DNN itself to find the optimum. Considering our HF source, it would take a month to evaluate this amount of data points. Later, Ribeiro *et al.* (2020) showed that a bio-inspired algorithm would require, on average, 3750 evaluations, which would take almost 30 h to evaluate.

In this work, the problem will be solved using SAO. Kriging (KRG) and Hierarchical Kriging (HKRG) models will be considered. These methods were implemented in BIOS, an in-house optimization software developed in LMCV to perform structural optimization using meta-heuristics. Results will be shown the PI, the EI, the VF-PI, and the VF-EI. For the KRG model, 5 initial HF data points will be considered. For the HKRG, only 3 HF points will be considered for



(a) Ceramic material gradation.

(b) Buckling mode.

Figure 9: Optimum design for the buckling load maximization of the square FGP.

the initial sample, along with 20 LF data points. The maximum number of sampling points is 100, but the algorithm may also stop earlier after a maximum number of consecutive iterations with no improvement $It_{stall} = 10$.

Results will be compared in terms of the average Normalized Root Mean Squared Error (NRMSE), the total number of HF and LF evaluations (n_h and n_l), and the average Wall-Clock Time (WCT) spent for each optimization. In all cases, 10 optimizations were performed. All numerical computations were performed on a computer with a Core i7-5500U CPU of 2.40 GHz clock speed and 16 GB of RAM. No parallelization procedure was used.

Table 2 presents the obtained results. Results show that the KRG-PI was not able to find the optimum, on average. Also, even though the correlation between sources is very high, the HKRG-VF-PI also struggled in finding the optimum design. On the other hand, both the EI and the VF-EI showed very accurate results. In terms of efficiency, the HKRG-VF-EI was able to find the optimum with only 7 HF evaluations, on average, while KRG-EI required 21 evaluations. That way, in terms of computational time, the HKRG-VF-EI was able to perform the optimization process 2.7 times faster than the single-fidelity algorithm.

Table 2: Results found for the buckling load maximization of the square FGP.

Model	Infill criterion	λ	NRMSE	n_h	n_l	WCT (s)
KRG	PI	9.823	0.70%	27	-	772
	EI	9.893	0.00%	21	-	607
HKRG	VF-PI	9.761	1.34%	29	46	927
	VF-EI	9.892	0.01%	7	31	225

It is worth noting that, using a DNN to assist in the optimization process, Do, Lee, and Lee (2019) did not manage to find the optimum design. On average, the best design found by the authors presented an NRMSE of 0.49%. Both the KRG-EI and the HKRG-VF-EI managed to find much better designs, and at a much lower number of HF evaluations.

Figure 10 shows a comparison of the time spent for each approach for each phase of the process. Note that, while the Hierarchical Kriging shows a higher model building cost, the total process using the HKRG-VF-EI is much faster due to the lower number of HF evaluations required. On the other hand, the HKRG-VF-PI also showed a very poor efficiency, as it struggled to converge and, still, did not find the optimum of this problem.

CONCLUSION

In this paper, we performed the Sequential Approximate Optimization of a Functionally Graded Plate. We used two different models, Kriging and Hierarchical Kriging. The latter is a Multi-Fidelity (MF) model, able to consider information from multiple fidelities. Different acquisition functions were considered: Probability of Improvement (PI) and Expected Improvement (EI), as well as Variable-Fidelity (VF) versions of these methods for the Hierarchical Kriging. All results were compared in terms of accuracy and computational efficiency.

It was shown that the Hierarchical Kriging with the VF-EI is able to find very accurate optimal results, achieving outstanding efficiency. In comparison to the regular Kriging model with the EI, the method required 66% fewer HF evaluations. The total process was up to 2.7 times faster. We also showed that the VF-PI is not able to present similar results, as the method struggles to find the optimum.

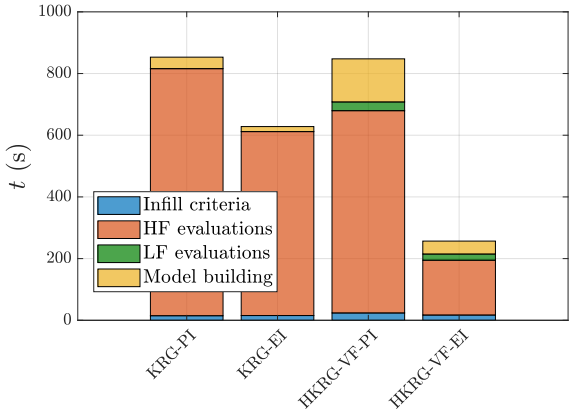


Figure 10: Cost of each phase of the process for the buckling load maximization of the square FGP.

REFERENCES

Do, D., Lee, D., Lee, J., 2019, "Material optimization of functionally graded plates using deep neural network and modified symbiotic organisms search for eigenvalue problems", Composites Part B: Engineering, v. 159, p. 300–326.

Do, D., Nguyen-Xuan, H., Lee, J., 2020, "Material optimization of tri-directional functionally graded plates by using deep neural network and isogeometric multimesh design approach", Applied Mathematical Modelling, v. 159, p. 300–326.

Forrester, A. I. J., Sobester, A., Keane, A. J., 2008, "Engineering design via surrogate modelling: A pratical guide.", Wiley.

Han, Z. H., Gortz, S., 2012, "Hierarchical Kriging model for variable-fidelity surrogate modeling". AIAA Journal, v. 50, n. 9, p. 1885-1896. Available from Internet: https://doi.org/10.2514/1.J051354.

Jones, D. R., Schonlau, M., Welch, W. J., "Efficient Global Optimization of expensive black-box functions", Journal of Global Optimization, p. 455-492, 1998.

Kushner, H. J., 1964, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise". Journal of Basic Engineering, v. 86, n. 1, p. 97–106.

Maia, M. A., Parente, E., Melo, A. M. C., "Kriging-based optimization of functionally graded structures". Structural and Multidisciplinary Optimization, 2021.

Mockus, J., 1974, "On Bayesian methods for seeking the extremum", Optimization Techniques IFIP, Technical COnference Novosibirsk, July 1-7, Berlin, Heidelberg: Springer Berlin Heidelberg, p. 400–404.

Ribeiro, L. G. *et al.*, 2020, "Surrogate based optimization of Functionally Graded Plates using Radial Basis Functions". Composite Structures, v. 252.

Ruan, X. *et al.*, 2020, "Variable-fidelity probability of improvement method for efficient global optimization of expensive black-box problems", Structural and Multidisciplinary Optimization, Springer, v. 62, n. 6, p. 3021–3052.

Shen, H. S. , 2009, "Functionally Graded Materials: Nonlinear analysis of plates and shells.", CRC Press.

Zhang, Y., Han, Z. H.; Zhang, K. S., 2018, "Variable-fidelity expected improvement method for efficient global optimization of expensive functions", Structural and Multidisciplinary Optimization, Springer, v. 58, n. 4, p. 1431–1451.

RESPONSIBILITY NOTICE

The authors are the only parties responsible for the printed material included in this paper.