Materials Research Proceedings 69 (2026) 1220-1224 https://doi.org/10.21741/9781644904251-210

A Multi-Fidelity Gaussian Process Framework for Predicting Process-Induced Deformations in Composites

K. JOHNSON[1,a], M. SANTORI[2,b], E. ZAPPINO[2,c], N. ZOBEIRY[1,d] and M. PETROLO[2,e *]

[1]Department of Materials Science and Engineering, University of Washington, Seattle, WA, 98195, USA

[2]MUL2 Lab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, 10129 Turin, Italy

[a]kaj1402@uw.edu, [b]martina.santori@polito.it, [c]enrico.zappino@polito.it, [d]navidz@uw.edu, [e]marco.petrolo@polito.it

Keywords: Multi-Fidelity Surrogate Models, Gaussian Process Regression, Bayesian Optimization, Carrera Unified Formulation, Process-Induced Deformations, Composite Design

Abstract. This work presents a multi-fidelity machine learning strategy for efficiently modeling the mechanical response of composite laminates during curing. A Gaussian Process Regression (GPR) framework is adopted to link the laminate stiffness parameters (B_{11}, D_{22}) to the stress component σ_{xx}, combining data from numerical models of different accuracy. High- and low-fidelity datasets are generated using Layer-Wise (LW) and Equivalent Single Layer (ESL) theories derived from the Carrera Unified Formulation (CUF) and finite-element analysis. The proposed approach employs Bayesian Optimization to adaptively select the most informative high-fidelity samples, thereby achieving accurate predictions at reduced computational cost. The model's predictive validity for stress distributions is demonstrated for L-shaped cross-ply composite laminates.

Introduction

Process-induced deformations and residual-stress development in composite laminates during curing pose significant challenges for ensuring structural performance and dimensional reliability in aerospace components [1]. Low-fidelity models provide fast, cost-effective estimates of deformation and stress distributions but often fail to capture detailed thickness variations. Conversely, high-fidelity simulations can resolve detailed stress fields across the laminate thickness but are computationally expensive, limiting their practicality for extensive design exploration of more complex part geometries.

For these reasons, there is growing interest in developing multifidelity simulation tools to analyze the composite manufacturing process [2]. The objective of this work is to minimize reliance on high-fidelity simulations by utilizing lower-fidelity data to train a surrogate model that accurately predicts stresses in composite components.

To bridge this gap, we propose a multi-fidelity surrogate modeling framework that combines two fidelity tiers: Equivalent Single Layer (ESL) theories and Layer-Wise (LW) simulations based on the Carrera Unified Formulation (CUF) [3]. A probabilistic surrogate model is then constructed using a Gaussian Process Regression (GPR) [2]. Two laminate stiffness parameters ($\mathbf{B_{11}}$, $\mathbf{D_{22}}$), derived from classical lamination theory, serve as GPR inputs and are mapped to the target stress. Bayesian Optimization (BO) is employed to allocate high-fidelity simulations to regions of high predictive uncertainty selectively.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

Materials Research Proceedings 69 (2026) 1220-1224 https://doi.org/10.21741/9781644904251-210

The effectiveness of the ML method is evaluated by predicting the stress $\boldsymbol{\sigma_{xx}}$ in L-shaped laminates made of AS4/8552 material.

Methodology

A multi-fidelity Gaussian Process Regression (GPR) framework was developed to model the relationship between laminate stiffness parameters B_{11} and D_{22} and the simulated stress component σ_{xx}. The goal was to quantify the amount of high-fidelity (HF) data required to achieve accurate predictions when abundant low-fidelity (LF) data are available. The GPR approach was chosen for its ability to provide both a mean prediction and an uncertainty estimate, enabling the model to use Bayesian optimization (BO) to identify regions where substitutive HF data would be most beneficial for improving model accuracy. The LF and HF datasets share the same inputs but differ in accuracy and computational cost. LF simulations are faster but less accurate, whereas HF simulations are more representative of the physical system but more computationally expensive to generate.

The distinction between fidelity levels arises from the structural modeling approach used to compute the training data. The high-fidelity (HF) data were obtained using Layer-Wise (LW) models, in which an independent kinematic expansion is used to describe each composite layer. This formulation enables a detailed through-thickness description, yielding a piecewise linear displacement field and a more accurate representation of interlaminar stress distributions. Conversely, the low-fidelity (LF) data are generated from Equivalent Single Layer (ESL) models, in which all layers share a single global kinematic expansion in the thickness direction. While each ply still retains its own material properties, the through-thickness displacement field in ESL models is linear and continuous. Both the ESL and LW datasets were generated using a higher-order model based on the Carrera Unified Formulation (CUF) combined with the finite element method (FEM) [4].

To appropriately represent the difference in accuracy between the two data sources, each fidelity level was assigned its own noise value (α). The noise term quantifies the uncertainty the model ascribes to each data point; larger values indicate that the model should place less trust in the data. Because LF simulations are less precise compared to HF simulations, the LF points were intentionally assigned higher noise levels. The optimal noise configuration was identified through a grid search over possible LF/HF combinations, ensuring that the LF-only model captures the general trends in the data without being overly optimistic. Once the baseline noise levels were established, the GPR model was iteratively trained using a composite kernel designed to capture both linear and nonlinear relationships between the stiffness parameters and stress response. The kernel combined a linear term, a Radial Basis Function (RBF), and a Rational Quadratic (RQ) component, along with a small fixed white noise term for numerical stability:

$$k(x,x') = C_1^2(x \cdot x') + C_2^2 RBF(L_1) + C_3^2 RationalQuadratic(\alpha, L_2) + C_4^2 WhiteNoise \quad (1)$$

$$k(x,x') = C_1^2(x \cdot x') + C_2^2\, exp\left(-\frac{\|x-x'\|^2}{2L_1^2}\right) + C_3^2\left(1+\frac{\|x-x'\|^2}{2\alpha L_2^2}\right)^{-\alpha} + C_4^2\delta(x,x') \quad (2)$$

where:

C_1^2 = Signal variance associated with the linear term
C_2^2 = Signal variance associated with the RBF term
L_1 = Characteristic length scale of the RBF term
C_3^2 = Signal variance of the RQ term
α = Shape parameter of the RQ term
L_2 = Length scale of the RQ term

Materials Research Proceedings 69 (2026) 1220-1224 https://doi.org/10.21741/9781644904251-210

C_4^2 = Noise variance
$\delta(x, x')$ = Kronecker delta function

Next, a BO framework was employed to determine which LF data points should be replaced with HF values to maximize model performance. The acquisition function utilized to accomplish this task was Expected Improvement (EI):

$$EI(x) = (\mu(x) - f^* - \xi)\Phi(z) + \sigma(x)\phi(z) \quad (3)$$

and

$$z = \frac{\mu(x) - f^* - \xi}{\sigma(x)} \quad (4)$$

where:
$\mu(x)$= Predicted mean of the Gaussian Process at candidate point x
$\sigma(x)$= Predicted standard deviation (uncertainty) of the Gaussian Process at x
f^*= Current best observed value of the objective function (here, validation R^2)
ξ= Small positive constant that controls exploration–exploitation balance
$\Phi(z)$= Cumulative distribution function (CDF) of the standard normal distribution
$\phi(z)$= Probability density function (PDF) of the standard normal distribution

At each iteration, BO ranks all candidate LF points by their potential contribution to model improvement. The LF samples with the greatest expected improvement are then replaced by their HF counterparts, and the GPR is retrained using the same kernel and noise configuration. This iterative substitution continues until the validation accuracy converges or all available HF data are used.

Results

The proposed approach was applied to an L-shaped composite structure comprising eight plies of unidirectional AS4/8552 material mounted on an Invar tool. The geometry, material properties, and curing cycle adopted are described in detail in [3]. Finite Element (FE) simulations based on the CUF higher-order model were conducted for all 256 possible cross-ply stacking sequences, employing both ESL and LW theories. The total degrees of freedom of the LW and ESL models of the L-shape structure were 21861 and 7455, respectively. The resulting dataset for the GPR analysis comprised 174 training points and 84 validation points. Initially, the GPR model was trained solely on the LF (ESL-based) dataset. In subsequent iterations, however, the 10 LF points with the greatest expected improvement are replaced with their corresponding HF counterparts. In Fig. 1 the updated GPR surface is evaluated over a grid to visualize how the predicted surface evolved as more HF data points are incorporated.

Materials Research Proceedings 69 (2026) 1220-1224 https://doi.org/10.21741/9781644904251-210

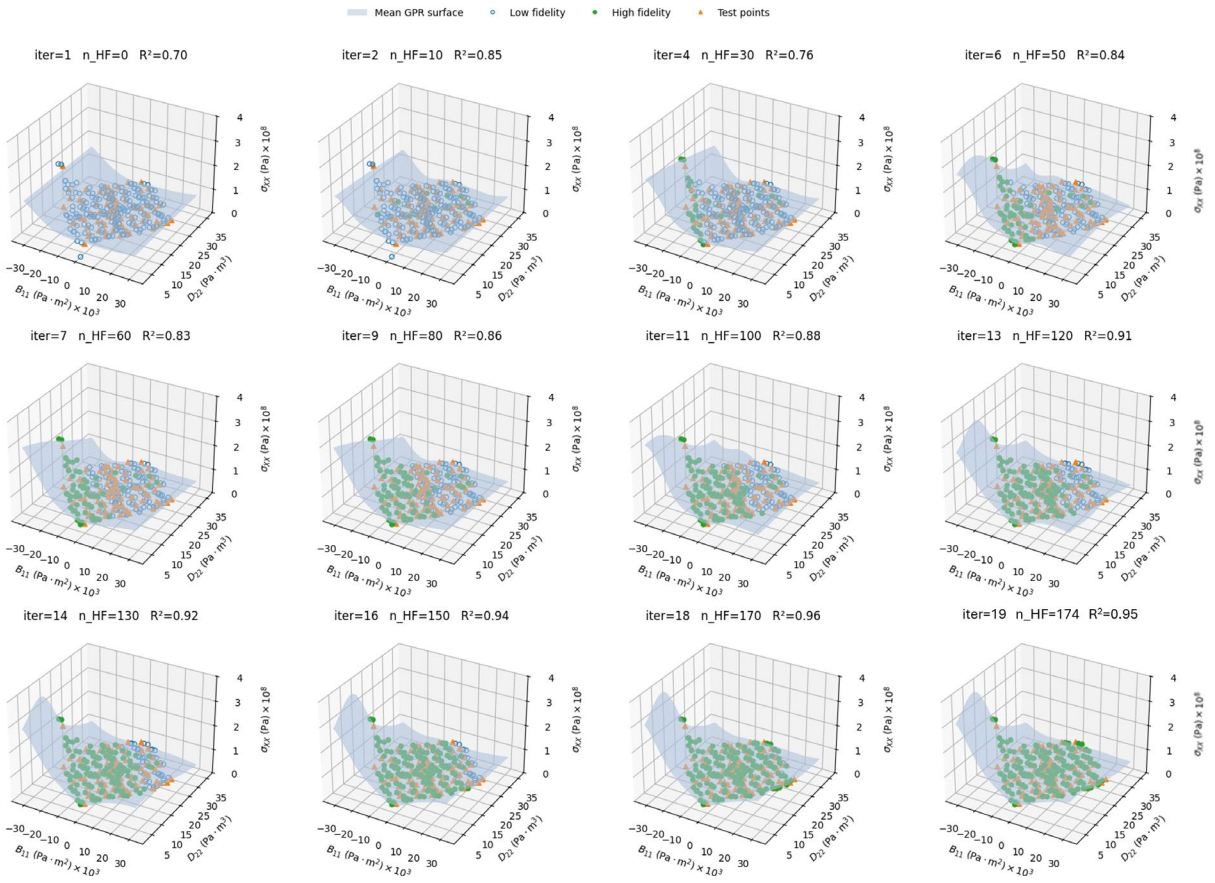


Figure 1: Evolution of the predicted GPR surface over the input parameter grid as high-fidelity data points $\boldsymbol{n_{HF}}$ are progressively incorporated

These results show the progressive refinement of the multi-fidelity GPR model as LF data points are gradually replaced with HF values. The term n_{HF} is representative of the number of HF points at a given iteration. At iteration 1 (iter = 1), the model trained only on LF data achieves a moderate accuracy (R^2 = 0.70), capturing general trends but missing finer variations in the 3D space. As additional HF points are incorporated, the predicted surface becomes smoother and more physically consistent with the observation data, and the validation accuracy steadily improves from roughly R^2 = 0.70 to R^2 = 0.95 by the final iteration. This also demonstrates that, within six iterations, the model increased accuracy by 14% and maintained that level for the remaining iterations. This means that, very quickly, it is possible to create an accurate machine learning model by replacing only 50 HF points (~29% of the total LF data). The results in Fig. 2 illustrate how the model evolves as HF data are selectively introduced through BO. At iteration 1, the GPR surface relies solely on LF data, producing a smooth but relatively uncertain trend with a moderate accuracy of R^2 = 0.70. By iteration 6, several strategically chosen LF points have been replaced by their HF counterparts, tightening the 95% confidence band and improving correlation to R^2 = 0.84. Evidence that the model is learning the actual stress distribution more effectively. Finally, at iteration 14, with 130 HF points now in training, the model achieves R^2 = 0.92, with the predictive surface aligning closely with the HF validation points and uncertainty collapsing in multiple zones of the domain.

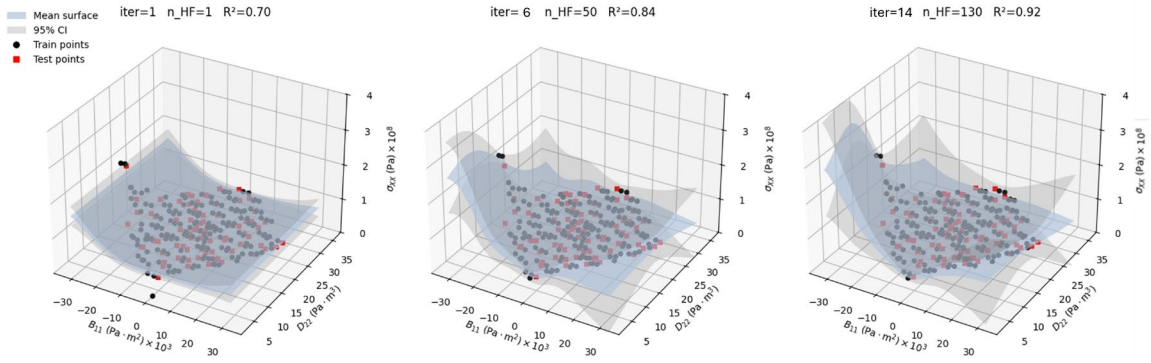


Figure 2: Multi-fidelity GPR predictive surface with iterative replacement of low-fidelity (LF) points by high-fidelity (HF) points, and comparison with the 95% confidence bands

Conclusion

In this work, a multi-fidelity GPR framework was applied to predict the stress field in composite laminates by combining data from models of different fidelity. The approach integrates low-fidelity (ESL) and high-fidelity (LW) simulations, formulated within the Carrera Unified Formulation (CUF), and solved using finite-element analysis. The results obtained for the L-shaped AS4/8552 composite structure show that, through Bayesian Optimization, a limited number of HF samples are sufficient to improve prediction accuracy significantly. After only a few iterations, the model achieved an accuracy $\mathbf{R^2}$ of 0.95, demonstrating that selective replacement of LF data is an effective strategy for balancing computational costs and accuracy.

Acknowledgement

This work was partly supported by the Italian Ministry of Foreign Affairs and International Cooperation, grant number US23GR12.

References

[1] E. Zappino, M. Santori, R. Masia, N. Zobeiry, M. Petrolo. Numerical Analysis of the Impact of Process Parameters on the Residual Stress of a Flat Composite Part. Aerotecnica Missili & Spazio (2025). https://doi.org/10.1007/s42496-024-00231-7

[2] C. Schoenholz, E. Zappino, M. Petrolo, N. Zobeiry. Efficient analysis of composites manufacturing using multi-fidelity simulation and probabilistic machine learning. Composites Part B: Engineering, Vol. 280, 111499 (2024). https://doi.org/10.1016/j.compositesb.2024.111499

[3] E. Zappino, N. Zobeiry, M. Petrolo, R. Vaziri, E. Carrera, A. Poursartip. Analysis of process-induced deformations and residual stresses in curved composite parts considering transverse shear stress and thickness stretching. Composite Structures, Vol. 241 (2020). https://doi.org/10.1016/j.compstruct.2020.112057

[4] E. Carrera, M. Cinefra, M. Petrolo, E. Zappino: Finite Element Analysis of Structures through Unified Formulation. John Wiley & Sons (2014). https://doi.org/10.1002/9781118536643