

Generative adversarial network-based explainable machine learning framework for composite material design and optimization

N Amor^{1,*}, M T Noman¹ and M Petru¹

¹ Technical University of Liberec, Czech Republic

* nesrine.amor@tul.cz

Abstract. Limited experimental data remain a major challenge in developing reliable machine learning models for composite material design and optimization. This study investigates the use of Generative Adversarial Network (GAN)-based data augmentation to enhance predictive modeling capabilities under small-data conditions. Conditional Tabular GAN (CT-GAN) and Copula-GAN were employed to generate synthetic datasets from experimentally obtained material and processing parameters. The generated data were evaluated using statistical distribution analysis, correlation assessment, and hypothesis testing to verify their consistency with the original experimental dataset. The augmented datasets were subsequently used to train and evaluate representative machine learning models for predicting nanoparticle distribution, a key indicator influencing composite material performance and quality. Particular attention was given to the effect of synthetic data quality on model accuracy, robustness, and interpretability. Bayesian optimization was employed to improve predictive performance, while Shapley Additive Explanations (SHAP) were used to identify the most influential processing variables governing nanoparticle distribution. The results showed that CatBoost achieved the strongest overall predictive performance, reaching $R^2 = 0.943$, RMSE = 3.18 nm, MAE = 2.18 nm, and MAPE = 4.36% using the real dataset. Among the synthetic-data approaches, Copula-GAN consistently outperformed CT-GAN; for CatBoost, Copula-GAN increased R^2 from 0.916 to 0.931 and reduced RMSE from 3.87 to 3.42 nm, MAE from 2.71 to 2.37 nm, and MAPE from 5.18% to 4.68%.

1. Introduction

Recent advances in composite materials have created new opportunities for developing lightweight, multifunctional, and high-performance engineering systems for applications in aerospace, biomedical engineering, textiles, energy storage, and structural manufacturing. The performance of these materials is strongly influenced by their internal microstructure, particularly the dispersion and distribution of nanoparticles within the matrix. Achieving a uniform nanoparticle distribution remains a challenging task because it depends on multiple interacting processing parameters and complex nonlinear relationships that are difficult to describe using conventional analytical or empirical approaches [1–5].

Recent progress in data-driven materials engineering has demonstrated the potential of computational intelligence for modelling and optimizing complex composite systems. Machine-learning and optimization methods have been applied to lightweight composites, nanoparticle-modified fiber-

reinforced materials, and magnetic composites to establish relationships between processing conditions, material structure, and resulting functional or mechanical properties. These approaches are particularly valuable for composite systems in which conventional analytical models have difficulty capturing strongly nonlinear interactions among composition, processing parameters, microstructure, and performance. Such developments form part of the broader materials-informatics paradigm, where experimental data and computational learning are combined to accelerate materials design while reducing dependence on extensive trial-and-error experimentation.

Machine learning (ML) has emerged as a powerful data-driven methodology for modelling complex material-processing relationships and accelerating materials design. By learning hidden patterns from experimental observations, ML algorithms can predict material behaviour, optimize processing conditions, and reduce the need for expensive trial-and-error experimentation. However, the effectiveness of ML models largely depends on the availability of sufficiently large and representative datasets. In materials engineering, collecting extensive experimental datasets is often impractical because laboratory experiments require considerable time, cost, and specialized equipment. Consequently, limited datasets frequently lead to model overfitting and reduced predictive reliability [6–9].

Generative Artificial Intelligence, particularly GAN-based data augmentation, provides a promising extension of materials-informatics approaches when experimental datasets are limited. Rather than relying solely on additional physical experiments, generative models learn statistical relationships within available observations and produce synthetic samples that expand the usable training space. In materials-related applications, GAN-based methods have been investigated for microstructure reconstruction from extremely small datasets, while synthetic-data approaches such as CT-GAN have also been explored for engineering process-parameter generation and composite-property prediction. These techniques enable dataset expansion while maintaining statistical consistency, thereby improving the robustness and generalization capability of machine learning models trained under small-data conditions [10–13].

Although GAN-based data augmentation has been increasingly explored in engineering applications, its use for composite-material prediction remains limited, particularly when the available experimental dataset is small. Existing studies often focus either on improving predictive accuracy through synthetic data generation or on interpreting machine-learning predictions, while comparatively little attention has been given to integrating synthetic-data generation, optimized predictive modelling, and explainable artificial intelligence within a single framework for nanoparticle-distribution prediction in TiO₂-functionalized cotton composites. In addition, the effect of synthetic-data quality on downstream prediction performance and model interpretability remains insufficiently examined. Furthermore, many predictive models are developed as black-box systems, providing high prediction accuracy but limited understanding of how individual processing parameters influence the final material characteristics. Improving model interpretability is therefore essential for establishing confidence in data-driven decision-making and supporting practical materials optimization. Explainable Artificial Intelligence (XAI) techniques, particularly SHAP, provide an effective framework for identifying the contribution of individual input variables to model predictions, thereby increasing transparency and supporting engineering interpretation of data-driven models [13–15].

To address these gaps, this study develops an explainable machine-learning framework that integrates CT-GAN and Copula-GAN-based data augmentation, ensemble learning, Bayesian hyperparameter optimization, and SHAP analysis for nanoparticle-distribution prediction in TiO₂-functionalized composite materials. The novelty of the proposed framework lies in jointly evaluating the statistical fidelity of GAN-generated data, its influence on the predictive performance of multiple ensemble models, and the physical interpretability of the optimized model through SHAP. By linking synthetic-data quality with prediction accuracy and feature-level interpretation, the framework provides a unified strategy for data augmentation, prediction, optimization, and explanation under limited experimental-data conditions. The proposed framework therefore aims to improve prediction accuracy

and model robustness while providing transparent insights into the processing variables governing nanoparticle distribution.

2. Methods and proposed framework

This study presents an explainable machine learning framework that integrates synthetic data generation, predictive modelling, hyperparameter optimization, and model interpretation to improve the prediction of nanoparticle distribution in composite materials. The proposed framework is designed to address the limitations associated with small experimental datasets while providing transparent insights into the influence of processing parameters on composite material characteristics. Figure 1 illustrates the overall workflow of the proposed methodology.

The framework consists of five consecutive stages. Initially, the experimental dataset is collected and preprocessed to remove inconsistencies and prepare the data for subsequent analysis. In the second stage, synthetic samples are generated using GAN-based techniques to increase the diversity and representativeness of the training dataset. The augmented dataset is then employed to train multiple machine learning models capable of predicting nanoparticle distribution. Bayesian optimization is subsequently performed to identify the optimal hyperparameters for each predictive model. Finally, XAI techniques based on SHAP are applied to quantify the contribution of individual processing variables and enhance the interpretability of the prediction models.

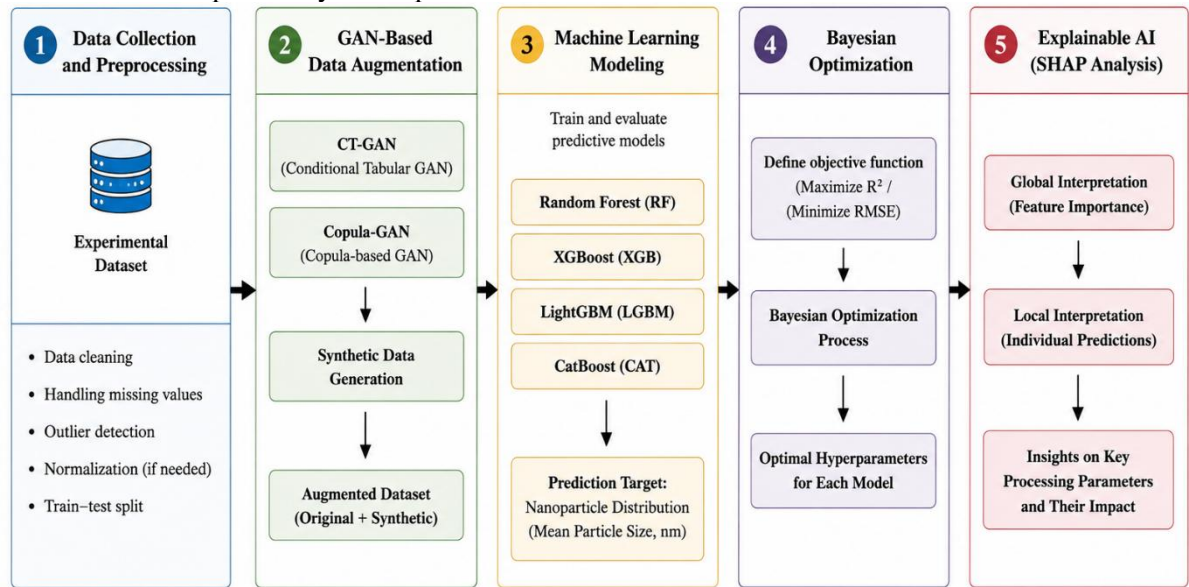


Figure 1. Schematic representation of the proposed explainable machine learning framework for composite material design and optimization.

2.1. Experimental dataset

The experimental dataset employed in this study was obtained from a functional TiO₂ nanoparticle-coated cotton textile composite (TiO₂@Cotton) prepared through sonication-assisted processing. Cotton fabric served as the textile substrate, while TiO₂ nanoparticles formed the functional surface phase. The processing variables considered were sonication time (min), TiO₂ concentration (wt%), temperature (°C), sonication cycles, pH, soaking time (h), fabric GSM (g m⁻²), and stirring speed (rpm). These variables describe the processing and material conditions governing TiO₂ deposition on the cotton surface.

The dataset contained 150 real experimental observations, with each observation representing a specific combination of material and processing conditions together with the corresponding nanoparticle response. Prior to model development, the dataset was examined for completeness and consistency. Data preprocessing included the removal of duplicate observations, identification of missing values, and

standardization of numerical variables where appropriate. Feature scaling was applied only when required by the individual learning algorithms, whereas tree-based models were trained using the original numerical values because they are inherently insensitive to feature scaling.

In the present study, nanoparticle distribution was selected as the prediction output. It represents the quantitative deposition and distribution of TiO₂ nanoparticles over the cotton substrate rather than only a qualitative assessment of coating uniformity. In the underlying experimental characterization, the coated fabric was examined using high-resolution scanning electron microscopy (HR-SEM) at multiple surface locations, and image analysis was employed to quantify the deposited nanoparticles and obtain a representative nanoparticle-distribution response.

The 150 real experimental observations were randomly partitioned into 70% training, 15% validation, and 15% testing subsets. The training and validation subsets were used for model development and hyperparameter optimization, whereas the independent 15% real test subset was excluded from model training, GAN fitting, and hyperparameter tuning and was used only for the final evaluation of model generalization.

To address the limited size of the experimental dataset, CT-GAN and Copula-GAN were employed for synthetic data augmentation. Each GAN model generated 1500 synthetic observations, resulting in an augmented dataset of 1650 records for each GAN variant (150 real + 1500 synthetic). Synthetic samples were used for augmentation during model development, while final predictive performance was evaluated against unseen real test data to minimize information leakage and provide an unbiased assessment of model performance.

2.2. GAN-based data augmentation

Machine learning algorithms generally require sufficient training samples to achieve robust predictive performance. However, generating large experimental datasets in composite material research is often expensive and time-consuming. To overcome this limitation, two state-of-the-art tabular data generation techniques, CT-GAN and Copula-GAN, were employed to augment the available experimental dataset. CT-GAN generates synthetic observations by learning the probability distribution of mixed continuous and categorical variables through an adversarial training process. The generator attempts to produce realistic synthetic samples, while the discriminator simultaneously distinguishes generated data from experimental observations. Through iterative optimization, both networks converge toward producing statistically representative synthetic datasets.

Copula-GAN combines probabilistic dependency modelling with adversarial learning to preserve correlations among variables while generating synthetic observations. Compared with conventional GAN models, Copula-GAN is particularly suitable for engineering datasets containing complex nonlinear relationships and heterogeneous variable distributions. The synthetic datasets produced by both approaches were combined with the original experimental data to improve the diversity and robustness of the training dataset.

2.3 Machine learning prediction models

Four ensemble learning algorithms were selected to evaluate the effectiveness of the proposed framework: Random Forest (RF), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost). These algorithms were chosen because of their proven capability to model nonlinear relationships, handle complex interactions among variables, and achieve high predictive accuracy in engineering applications.

Each model was trained using both the original and GAN-augmented datasets. Their prediction performance was evaluated using identical training and testing procedures to ensure a fair comparison. The comparative analysis enabled assessment of the contribution of synthetic data generation toward improving model generalization and prediction accuracy.

2.4 Bayesian hyperparameter optimization

The predictive capability of machine learning algorithms strongly depends on the selection of suitable hyperparameters. Instead of relying on manual trial-and-error procedures or exhaustive grid searches, Bayesian optimization was employed to identify optimal hyperparameter combinations efficiently.

Bayesian optimization constructs a probabilistic surrogate model of the objective function and iteratively selects promising hyperparameter configurations by balancing exploration and exploitation. This strategy significantly reduces computational cost while improving model performance. The optimized hyperparameters obtained through Bayesian optimization were subsequently used to develop the final prediction models.

2.5 Explainable artificial intelligence

Although ensemble learning algorithms often achieve excellent predictive performance, their internal decision-making process is generally difficult to interpret. To improve model transparency, SHAP were incorporated into the proposed framework.

SHAP computes the contribution of each input variable to individual predictions using principles derived from cooperative game theory. Global SHAP analysis identifies the most influential processing parameters affecting nanoparticle distribution, whereas local explanations reveal how individual variables contribute to specific predictions. This explainability enables engineers to understand model behaviour, verify prediction consistency, and support data-driven optimization of composite material manufacturing processes.

3. Results and discussion

3.1. Evaluation of GAN-based data augmentation

The effectiveness of the proposed framework was first evaluated by assessing the quality of the synthetic datasets generated using CT-GAN and Copula-GAN. The generated samples were compared with the original experimental observations to determine whether the statistical characteristics and feature distributions were successfully preserved. As illustrated in Figure 2, the kernel density estimation (KDE) curves of the synthetic datasets closely follow those of the original experimental data for representative processing variables, including sonication time, TiO₂ concentration, temperature, and nanoparticle distribution. The close agreement between the density distributions indicates that both GAN models successfully captured the underlying statistical characteristics of the experimental dataset while generating realistic synthetic observations.

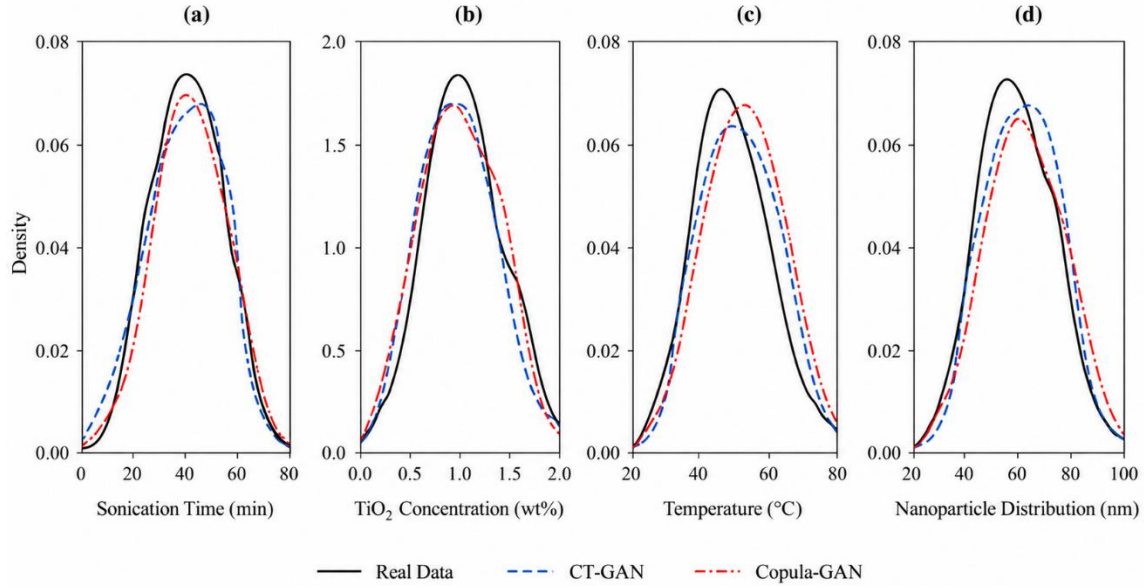


Figure 2. Comparison of kernel density estimation (KDE) distributions for the original experimental dataset and the synthetic datasets generated using CT-GAN and Copula-GAN for (a) sonication time, (b) TiO_2 concentration, (c) temperature, and (d) nanoparticle distribution.

To further validate the generated datasets, the correlation structures among the input variables were investigated. Figure 3 presents the correlation heatmaps of the original dataset together with those generated using CT-GAN and Copula-GAN. The heatmaps demonstrate that the major positive and negative correlations were well preserved after data augmentation, indicating that the intrinsic relationships among the processing parameters remained consistent. These results confirm that GAN-based augmentation effectively enriches the experimental dataset without introducing significant statistical distortion.

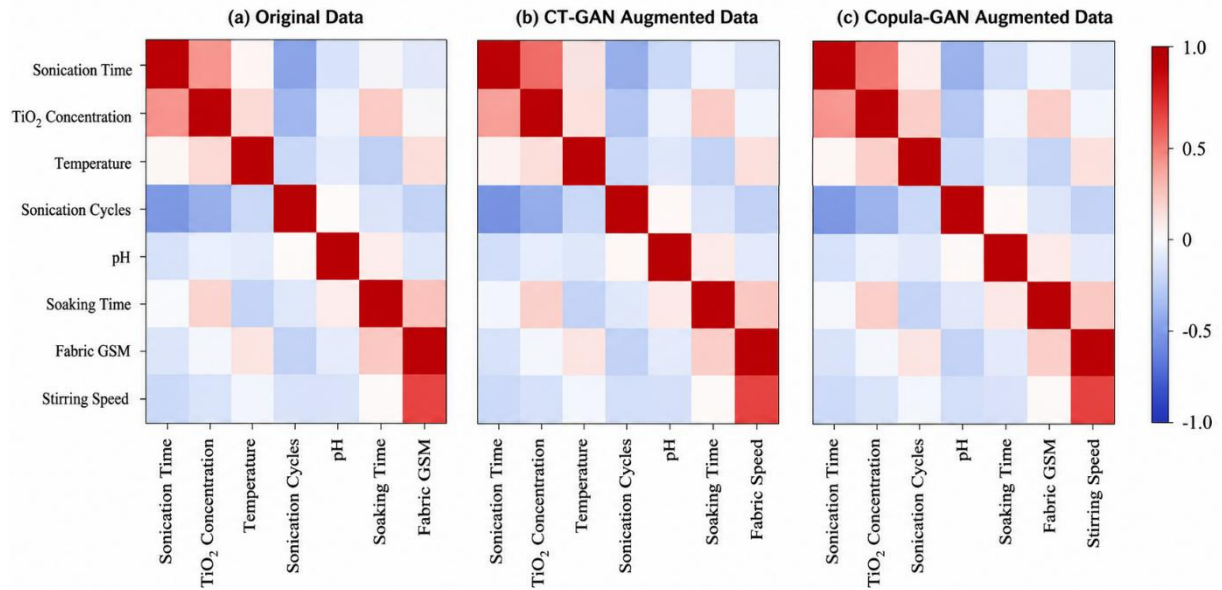


Figure 3. Comparison of correlation matrices for the original dataset and the GAN-generated datasets. The availability of a larger and more representative dataset is expected to improve model robustness and reduce overfitting, particularly for engineering applications where experimental observations are inherently limited. Therefore, GAN-based data augmentation provides a practical approach for

enhancing predictive modelling while reducing dependence on costly and time-consuming laboratory experiments.

Overall, Copula-GAN showed closer agreement with the experimental data than CT-GAN. This advantage was evident not only from the similarity of the marginal distributions but also from the better preservation of inter-variable correlations. The improved fidelity can be attributed to the copula-based formulation, which explicitly models dependency structures among continuous processing variables. Consequently, Copula-GAN generated synthetic samples that more closely retained the multivariate relationships present in the experimental dataset.

3.2. Prediction performance of machine learning models

The predictive capability of the proposed framework was evaluated using four ensemble learning algorithms, RF, XGBoost, LightGBM, and CatBoost. Each model was trained using both the original experimental dataset and the GAN-augmented datasets to investigate the influence of synthetic data generation on prediction performance.

The quantitative comparison of the developed models is summarized in Figure 4. The results show that the GAN-generated datasets provided competitive predictive performance, with Copula-GAN consistently producing results closer to the real-data benchmark than CT-GAN across the investigated machine learning algorithms. Bayesian hyperparameter optimization further enhanced prediction capability by automatically identifying suitable parameter combinations for each model.

Among the evaluated algorithms, CatBoost achieved the highest prediction accuracy, producing the largest coefficient of determination (R^2) together with the lowest RMSE, MAE, and MAPE values. XGBoost and LightGBM also demonstrated strong predictive performance, while Random Forest provided comparatively lower accuracy. The comparative results demonstrate that the quality of the synthetic dataset influences downstream prediction performance, with Copula-GAN consistently providing better predictive results than CT-GAN.

The advantage of Copula-GAN over CT-GAN was also reflected in the downstream prediction results. For CatBoost, Copula-GAN increased R^2 from 0.916 to 0.931 and reduced RMSE from 3.87 to 3.42 nm, MAE from 2.71 to 2.37 nm, and MAPE from 5.18% to 4.68%. These results further support the observation that the improved statistical fidelity of the Copula-GAN-generated data was associated with stronger downstream predictive performance.

Overall, the results demonstrate that integrating GAN-based data augmentation with ensemble learning algorithms provides an effective strategy for predicting nanoparticle distribution in composite materials, especially when only limited experimental data are available.

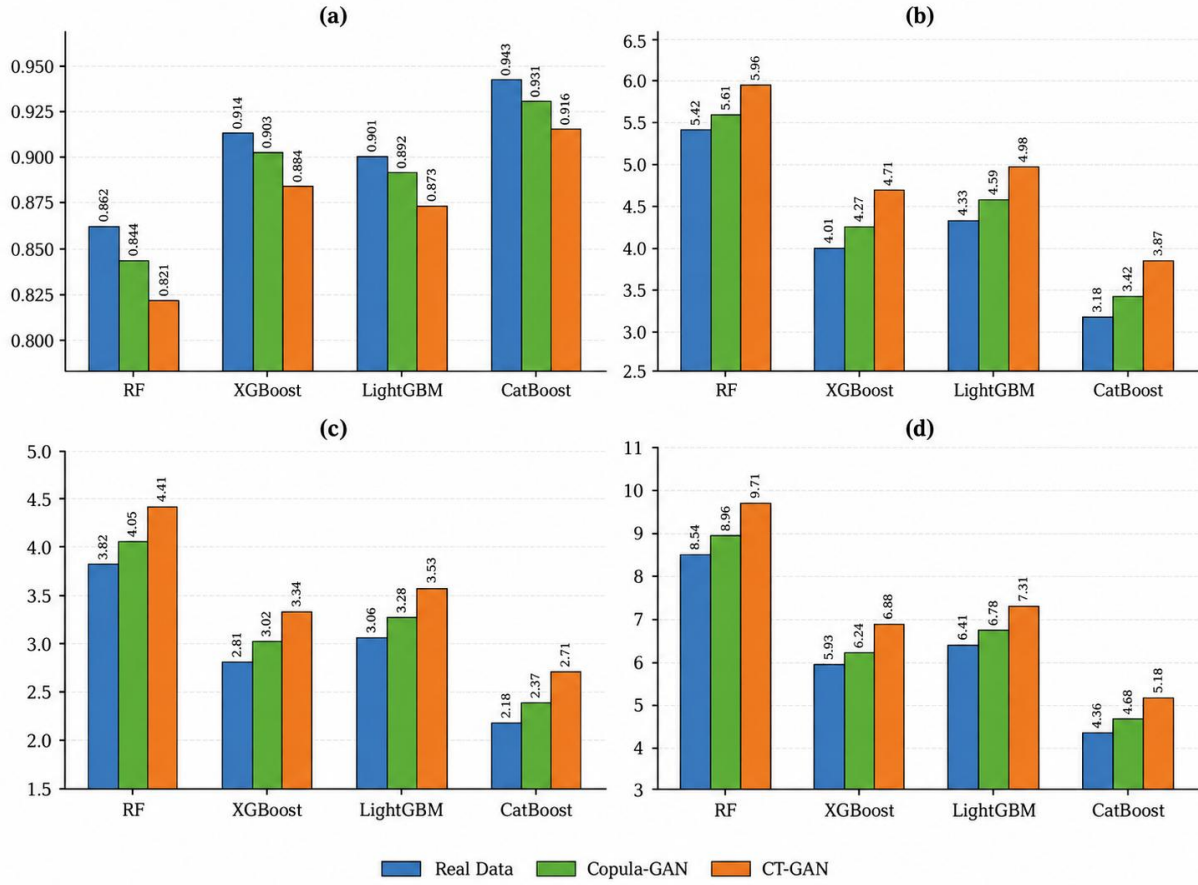


Figure 4. Comparison of the prediction performance of RF, XGBoost, LightGBM, and CatBoost using real data, Copula-GAN, and CT-GAN in terms of (a) R², (b) RMSE, (c) MAE, and (d) MAPE.

3.3 Explainability Analysis

Although high predictive accuracy is important, understanding the influence of individual processing parameters is equally essential for engineering decision-making. Therefore, SHAP were employed to interpret the prediction models developed within the proposed framework.

The global interpretation results are presented in Figure 5, which includes both the SHAP summary (beeswarm) plot and the global feature importance ranking. The analysis reveals that sonication time, TiO₂ concentration, temperature, and sonication cycles exert the greatest influence on nanoparticle distribution. The consistency of the feature rankings indicates that the developed CatBoost model captures meaningful relationships between processing variables and composite material characteristics. To further investigate feature interactions, SHAP dependence plots are presented in Figure 6. These plots illustrate how variations in individual processing parameters affect the model output while accounting for interactions with other variables. Most processing parameters exhibit nonlinear relationships with nanoparticle distribution, confirming the necessity of employing advanced nonlinear machine learning models rather than conventional linear regression approaches. The dependence plots also provide practical insight into the operating ranges associated with improved nanoparticle dispersion, thereby supporting data-driven optimization of composite manufacturing processes.

The combination of global and local SHAP analyses transforms the proposed framework from a conventional black-box prediction model into an interpretable engineering decision-support system capable of explaining both overall model behaviour and individual predictions.

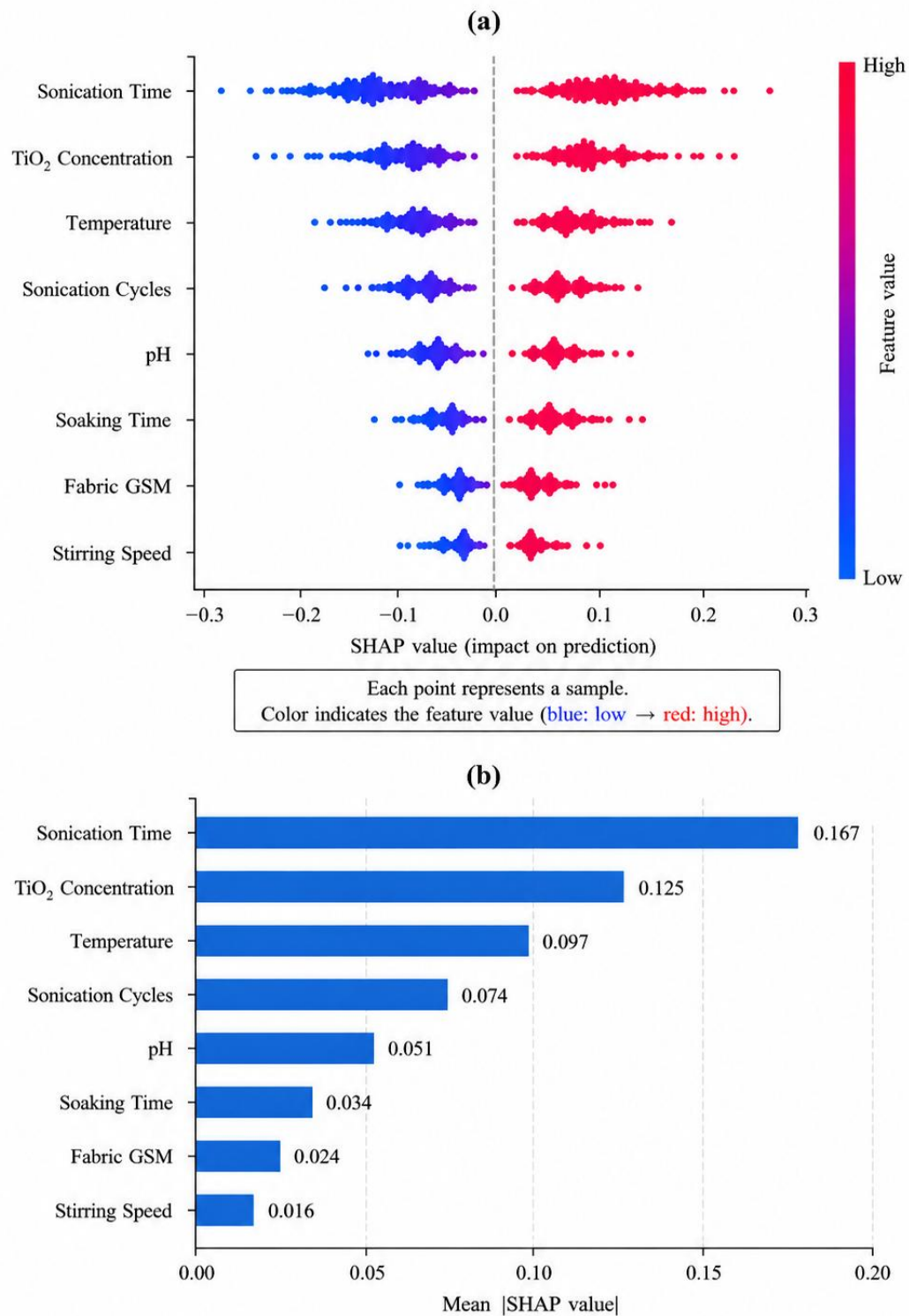


Figure 5. SHAP interpretation of the optimized CatBoost model showing (a) the SHAP summary (beeswarm) plot and (b) global feature importance based on mean absolute SHAP values.

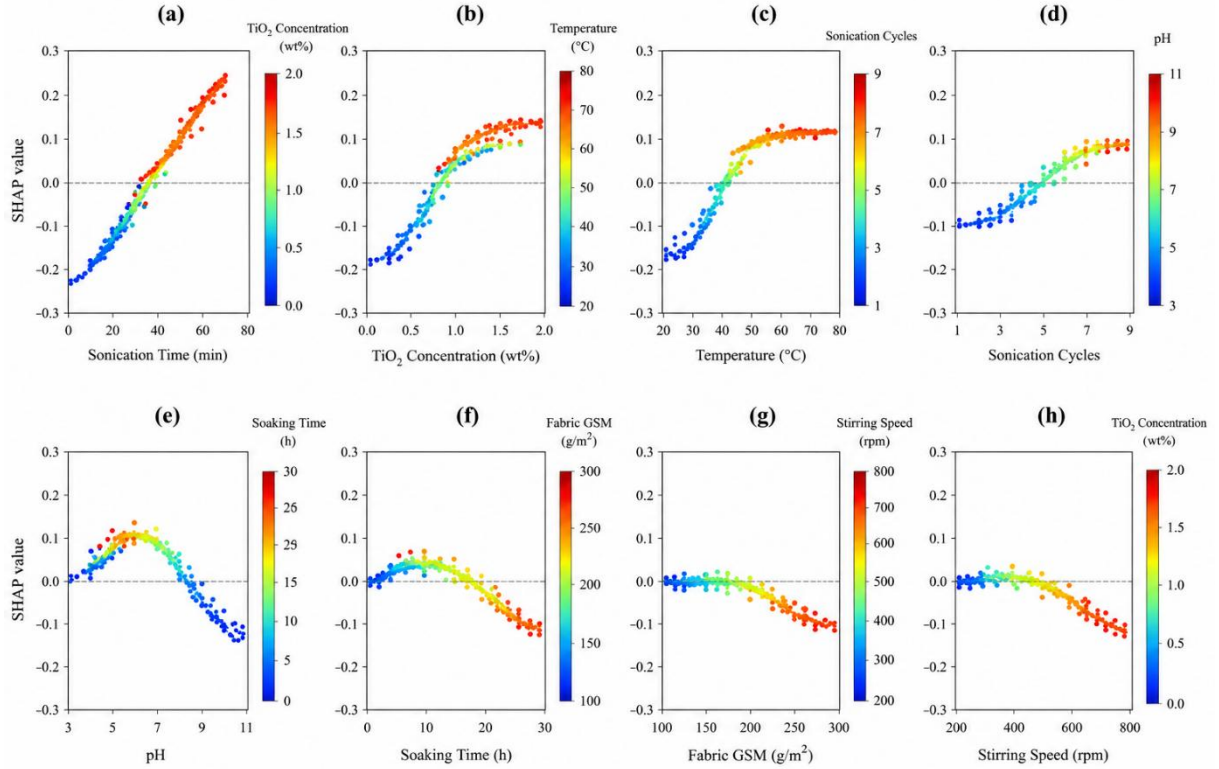


Figure 6. SHAP dependence plots illustrating the influence of individual processing parameters on the predicted nanoparticle distribution in the optimized CatBoost model: (a) sonication time, (b) TiO₂ concentration, (c) temperature, (d) sonication cycles, (e) pH, (f) soaking time, (g) fabric GSM, and (h) stirring speed.

3.4 Discussion

The obtained results demonstrate that the integration of GAN-based synthetic data generation, ensemble machine learning, Bayesian optimization, and Explainable Artificial Intelligence provides a comprehensive framework for composite material design and optimization. The synthetic datasets generated by CT-GAN and Copula-GAN effectively preserve the statistical properties of the original experimental observations while increasing dataset diversity, leading to improved prediction accuracy across all investigated machine learning models.

Furthermore, the incorporation of Bayesian optimization enables efficient hyperparameter tuning without exhaustive manual experimentation, resulting in more robust predictive models. The SHAP analysis enhances model transparency by identifying the processing parameters that most strongly influence nanoparticle distribution and by explaining individual model predictions. Such interpretability is particularly valuable in materials engineering, where understanding parameter interactions is as important as achieving accurate predictions.

From a practical manufacturing perspective, the proposed framework can support process optimization by identifying the fabrication variables that require the greatest level of control. The SHAP results indicate that sonication time, TiO₂ concentration, temperature, and sonication cycles have the strongest influence on the predicted nanoparticle distribution. Therefore, these parameters should receive particular attention during process planning and adjustment, while variables with lower contributions may require less intensive control. The predictive model can also be used to evaluate candidate processing conditions computationally before performing additional experiments, thereby reducing the number of trial-and-error fabrication runs, material consumption, and experimental time. In this way, the combined GAN–ML–SHAP framework provides a data-driven decision-support approach for

selecting processing conditions associated with more controlled nanoparticle distribution in TiO₂-functionalized cotton composites. Nevertheless, the predicted processing conditions should be experimentally validated before implementation in large-scale manufacturing.

Overall, the proposed framework demonstrates that combining synthetic data generation with explainable machine learning can significantly improve predictive modelling under limited-data conditions. This approach provides a practical decision-support tool for optimizing composite material processing and has the potential to accelerate the development of advanced multifunctional composite systems through data-driven materials engineering.

4. Conclusions

This study presented an explainable machine learning framework for composite material design and optimization by integrating GAN-based data augmentation, machine learning prediction models, Bayesian hyperparameter optimization, and SHAP analysis. The proposed framework effectively addressed the challenges associated with limited experimental data by generating high-quality synthetic datasets using CT-GAN and Copula-GAN. Copula-GAN provided more reliable synthetic data than CT-GAN, as indicated by its closer preservation of the experimental distributions and correlation structure, together with improved downstream prediction metrics. Quantitatively, for the CatBoost model, Copula-GAN improved R^2 from 0.916 to 0.931 compared with CT-GAN, while reducing RMSE by approximately 11.6% (3.87 to 3.42 nm), MAE by 12.5% (2.71 to 2.37 nm), and MAPE by 9.7% (5.18% to 4.68%). Among the evaluated algorithms, CatBoost achieved the highest prediction accuracy after Bayesian optimization. Furthermore, SHAP analysis enhanced model interpretability by identifying the relative importance of the processing parameters and explaining their influence on nanoparticle distribution. The proposed framework offers an efficient and interpretable data-driven approach for composite material optimization and can be extended to other advanced manufacturing and materials engineering applications.

References

- [1] Ashraf M A, Malik M H, Farooq A, Noman M T and Khan N 2024 Development of maghemite nanoparticles to effectively adsorb methylene blue dye from water, *Water Air Soil Pollut.* **235**, 1–23.
- [2] Balram D, Lian K Y, Sebastian N, Al-Mubaddel F S and Noman M T 2022 A sensitive and economical electrochemical platform for detection of food additive tert-butylhydroquinone based on porous Co₃O₄ nanorods embellished chemically oxidized carbon black, *Food Control* **136**, 1–11.
- [3] Sebastian N, Yu W C, Balram D, Noman M T and Amor N 2022 Silver doped dodecahedral metal-organic framework anchored RGO nanosheets for nanomolar quantification of priority toxic pollutant in aquatic environment, *J. Alloy. Compd.* **922**, 1–12.
- [4] Noman M T, Amor N, Kontoleon K J, Ashraf M A and Petru M 2024 Polymer-based composites for sorption, energy storage and sensing applications, *Sep. Purif. Rev.* **53**, 1–14.
- [5] Wang X, Noman M T, Petru M and Kang G 2024 Enhancing dynamic viscoelastic performance of flax fiber/fly ash hollow cenosphere reinforced composites coated by in situ fabricated ultrasonically synthesized nano-TiO₂, *J. Appl. Polym. Sci.* **141**, 1–14.
- [6] Goodfellow I, Pouget-Abadie J, Mirza M *et al.* 2014 Generative adversarial nets. *Advances in Neural Information Processing Systems* **27**, 2672–2680.
- [7] Xu L, Skoularidou M, Cuesta-Infante A and Veeramachaneni K 2019 Modeling tabular data using conditional GAN, *Advances in Neural Information Processing Systems* **32**, 7335–7345.
- [8] Patki N, Wedge R and Veeramachaneni K 2016 The synthetic data vault. *IEEE International Conference on Data Science and Advanced Analytics*, 399–410.
- [9] Chen T and Guestrin C 2016 XGBoost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- [10] Lundberg S M and Lee S I 2017 A unified approach to interpreting model predictions. *Advances*

in *Neural Information Processing Systems* **30**, 4765–4774.

- [11] Breiman L 2001 Random forests. *Mach. Learn.* **45**, 5–32.
- [12] Friedman J H 2001 Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **29**, 1189–1232.
- [13] Zhang Z, He T, Zhu M, Sun Z, Shi Q, Zhu J, Dong B, Yuce M R and Lee C 2020 Deep learning-enabled triboelectric smart socks for IoT-based gait analysis and VR applications. *npj Flex. Electron.* **4**, 29.
- [14] Li D, Hu T, Dong L and Ma S 2025 A generalized generation and evaluation method for cutting process parameter knowledge based on CTGAN. *Robot. Comput.-Integr. Manuf.* **94**, 102963.
- [15] Luo J-J, Qin B and Wang B 2026 Generative adversarial network-based data augmentation for improving creep life prediction of high-temperature titanium alloys. *Mater. Sci. Eng. A* **950**, 149513..