AI-Based Surrogate Models of Digital Twins for Industrial Processes

LWELE, Emmanuel and SHENFIELD, Alex <http://orcid.org/0000-0002-2931-8077>

AI-based surrogate models of digital twins for Industrial Processes

1st Emmanuel Lwele
Sheffield Hallam University
National Centre of Excellence for Food Engineering
Sheffield S9 2AA, United Kingdom
e.lwele@shu.ac.uk

2nd Alex Shenfield
Sheffield Hallam University
National Centre for Excellence in Food Engineering
Sheffield S9 2AA, United Kingdom
a.shenfield@shu.ac.uk

***Abstract*—Digital twins represent virtual replicas of physical systems, integrating real-time data and advanced analytics to monitor, simulate, and optimize industrial processes. This research delves into the application of AI-based surrogate models to improve the efficiency and accuracy of digital twins for industrial processes. The study employs machine learning techniques to develop computationally efficient models that maintain high accuracy. The integration of advanced sampling techniques and challenges related to data quality and interpretability are highlighted, proposing solutions to improve model robustness and reliability.**

***Index Terms*—AI-based surrogate models, Digital twins, industrial processes, Machine learning, Neural networks, Decision trees, Ensemble methods, Predictive maintenance, Process optimization, Sampling techniques, and Computational efficiency.**

I. Introduction

Surrogate models are instrumental in providing virtual representations that mirror physical objects or processes, yet they serve distinct purposes in the realm of simulations and digital transformation. Surrogate models, also known as proxy or emulator models, are simplified mathematical or computational models designed to approximate the behaviour of complex systems or processes. They serve as surrogates for computationally expensive or time-consuming simulations, enabling rapid and efficient analysis, optimization, or decision-making. Surrogate models provide a trade-off between accuracy and computational cost, making them valuable for uncertainty quantification and optimization tasks [1]

Digital twins go beyond approximation and emulation. They are comprehensive virtual representations of physical entities, integrating real-time data, physics-based models, and historical information to mirror the behaviour and evolution of their physical counterparts. Digital twins are interactive and dynamic, capable of continuously learning and adapting to real-world conditions, enabling predictive maintenance, performance monitoring, and decision-making support [1]. According to [2], surrogate models are approximations of complex, often computationally expensive, mathematical models or simulations. They are used to predict the behaviour of a system or process based on limited input data, making them valuable in scenarios where the original model is too resource-intensive or time-consuming to use directly.

Surrogate models play a crucial role in various applications, such as optimization, uncertainty quantification, and real-time decision-making. They are utilized to efficiently explore design spaces in engineering, accurately estimate system responses, or enable rapid what-if analyses.

The use of deep learning (DL) and machine learning (ML) methods has further extended the capabilities of surrogate modelling, as discussed in [3]. DL and ML enable the creation of highly interconnected models that approximate the underlying physics of the system with a high level of accuracy, replacing many expensive physics-based simulations. Deep Learning models, like long short-term memory (LSTM) networks, have been extensively leveraged to build surrogate models for sequence-to-sequence prediction, such as in aerodynamics, elastoplastic finite element simulation, and microstructure evolution applications [3].

Digital twins and surrogate models both serve as virtual representations of physical systems, but they differ in their fundamental purposes and methodologies. While digital twins strive to create a real-time, dynamic replica of physical entities, surrogate models seek to approximate complex, often computationally expensive systems or simulations with simpler, easier-to-evaluate models. Digital twins are developed to mirror the behaviour and performance of physical objects or processes, facilitating continuous monitoring, analysis, and prediction, and allowing for interaction with their physical counterparts.

Surrogate models are created to approximate the behaviour or response of complex models or systems using simpler, more computationally efficient models. They are often employed in scenarios where the original model is too resource-intensive or time-consuming to use directly. Surrogate models are valuable for optimization, uncertainty quantification, and rapid what-if analyses, among other applications [4].

A. Contributions to the Field

This research significantly advances the field of industrial process modelling in several ways:

- Enhanced Model Precision: The paper offers a framework for improving the accuracy and dependability of forecasts in industrial applications by incorporating cutting-edge machine-learning techniques.

- Methodological Advancements: The use of gradient boosting and ensemble approaches has raised the bar for model building in industrial settings by handling complicated datasets with multiple input variables in a methodologically novel way.
- Practical Implications: The results highlight the useful advantages of AI in the renewable energy sectors by directly enhancing operational efficiency and cutting costs in wind power management.

B. Problem Statement

Because of the intricate nature of the systems involved and the dynamic environmental factors that affect them, modelling complex industrial processes presents substantial problems. Industrial processes are highly variable and influenced by a multitude of unpredictable elements, particularly those that depend on natural resources like solar or wind power. Accurate outcome prediction is challenging due to these complications, which is important for maximising operational dependability and efficiency.

C. Objective

With an emphasis on wind turbine power generation, the main goal of this research is to present AI-based surrogate models as cutting-edge approaches to improving the predictability and efficiency of industrial processes. The purpose of this study is to show how these sophisticated models can greatly enhance operational strategies and decision-making in the renewable energy industry.

II. Literature Review

A. Traditional Modeling Techniques

In traditional industrial process simulation, physical and experimental approaches are specifically designed for understanding and anticipating the behaviour of complicated industrial systems and represent a significant part of current modelling techniques.

- Physical-Based Models: These models use the basic principles of physics, chemistry and engineering for describing systems behaviour. Common examples include thermodynamic models for energy systems, fluid dynamics models for chemical reactors, and finite element models (FEM) for structural analysis. The physical models can be very precise because they are built on established science laws [13].
- Empirical models: Empirical models are developed based on historical data and observed relationships, and do not necessarily understand the underlying phenomena. Typical examples are techniques such as regression analysis, polynomial models and surface response methodology [14]. Empirical models are faster to build than those of the Physical Model because they do not need as much detailed system information.

Traditional modelling techniques face several challenges and limitations, despite their widespread use. Physical models are often a combination of complicated equations, requiring substantial computing resources to be used in particular for large-scale industrial applications, which make them costly and time-intensive. While physical models are accurate for known conditions, the accuracy of their calculations is reduced when applied to situations outside the scope of their initial assumptions. Because they are typically designed for specific scenarios and cannot adapt well to changes in process parameters [16].

Empirical models struggle with generalization across different systems or conditions, the models are frequently tailored to the particular datasets they are trained on, which restricts their usefulness to only comparable scenarios [15]. The availability and quality of historical data is also another major factor for empirical models. Unreliable models may result from inadequate or poor data, leading to decisions that tack a solid foundation of information [15].

B. Surrogate Modeling

Surrogate models play a critical role in various applications, including optimization, uncertainty quantification, and rapid what-if analyses [4]. They are utilized to efficiently explore design spaces in engineering, accurately estimate system responses, or enable rapid analyses and decision-making for various processes and systems [4].

The conceptual and functional differences between surrogate models and traditional simulations lie in their underlying principles, objectives, and applications. AI-based surrogate models of digital twins are virtual representations of physical objects or systems that combine real-time data, physics-based models, and artificial intelligence to simulate, predict, and optimize the behaviour and performance of their physical counterparts [7]. In contrast, traditional simulations typically focus on specific aspects or subsystems of a physical object or process, using mathematical models and input data to predict behaviour and outcomes based on defined parameters [8]

The use of deep learning (DL) and machine learning (ML) methods has extended the capabilities of surrogate modelling, enabling the creation of highly interconnected models that approximate the underlying physics of the system with a high level of accuracy [3]. Surrogate models act as efficient stand-ins for more intricate and demanding simulations, facilitating faster analysis and decision-making for a wide array of processes and systems [3]. In digital twin technology, surrogate models are essential for enabling predictive and interpretable digital twins, especially in scenarios where the original model is too resource-intensive or time-consuming to use directly [5]. The importance of surrogate models lies in their ability to provide efficient approximations of complex systems for various analytical and computational purposes, thus enhancing the speed and efficiency of digital twin operations [3].

The ability of digital twins to provide integrated, multi-physics and multiscale simulation of complex products using the best available physical models and sensor updates to replicate their physical counterpart's life cycle is crucial [6]. Through digital twins, the product design process can be effectively divided into conceptual design, detailed design, and virtual verification, enabling a comprehensive approach to product development [6].

In the conceptual design phase, digital twins facilitate a quick understanding of areas for improvement by integrating various types of scattered data in the product's physical space, providing a streamlined source of information for designers [6]. Furthermore, digital twins act as faithful mappings of physical products, enhancing transparent and rapid communication between clients and designers, and enabling the utilization of customer feedback to guide the improvement of new products in the design phase [6].

During the detailed design stage, digital twins offer a solution to the lack of real-time data and environmental-impacted data, as they co-evolve with physical objects throughout their lifecycle, recording all product data and environmental influences [6]. This capability ensures that digital twins can support repeated simulation tests and provide insights into the impact of environmental parameters on product performance [6], and overall provide comprehensive insights throughout the product design process, ultimately contributing to more efficient and effective product development

The significance of surrogate models in enhancing the efficiency of digital twins in manufacturing is underscored by their ability to provide rapid and accurate approximations of complex systems or processes. Surrogate models play a vital role in reducing computational costs and time associated with complex simulations, thereby enabling efficient analysis, optimization, and decision-making processes within the digital twin framework [1]. By providing simplified yet accurate representations of physical systems or processes, surrogate models contribute to the overall efficiency of digital twins. These simplified models facilitate rapid data analysis and support timely decision-making by offering a trade-off between accuracy and computational cost [1].

Furthermore, surrogate models enable the integration of real-time data collection and prediction capabilities within digital twins, thereby enhancing the overall performance and responsiveness of digital twin systems [1].

C. Artifical Intelligence Techniques and Machine Learning in Surrogate Modelling

The concept of machine learning-based surrogates is an emerging area that leverages advanced technologies to create a digital replica of physical systems. This approach integrates machine learning techniques to enhance the capabilities of surrogate models in various domains such as predictive maintenance, simulation modelling, and intelligent decision-making [10]. According to [9], the concept of machine learning-based surrogate model frameworks comprises four modules: a physics-based nominal model, a data collection module, an algorithm for real-time update of the surrogate model, and a module for predicting the future state.

One key aspect highlighted in [10], is the utilization of a gray-box modelling approach, which combines both physics-based and data-driven frameworks. This approach allows the surrogate model to generalize and predict future responses. Furthermore, the gray box modelling framework is developed by coupling Bayesian filtering and machine learning algorithms. The study specifically employs Gaussian processes as the machine learning regression algorithm.

In [11], AI-based approaches for surrogate models are highlighted in various applications such as intelligent architecture for automated systems, trial and error testing reduction, comparison studies, and new framework proposals. For instance, the use of machine learning in trial and error testing reduction for 3D printing models is exemplified by Mukherjee and DebRoy in 2019 and the proposal of a new framework for AI-based surrogate models of digital twin systems based on machine learning models by Alexopoulos et al. in 2020.

In [12], the integration of artificial intelligence with surrogate models for structural health monitoring using deep learning is demonstrated, as well as the utilization of probabilistic graphical models and cloud-based surrogate models for structural health monitoring.

D. Applications of AI-Based Surrogate Models in Wind Turbine Modeling

The field of wind turbine modelling has experienced significant progress through the integration of AI-based surrogate models, which aim to improve the predictability and efficiency of wind generating systems. Surrogate models are a more efficient alternative to complex physical models, as they greatly reduce computing cost and time while still retaining acceptable levels of accuracy.

Historically, wind turbine power prediction models have mostly utilised physical and empirical approaches. Physical models, such as Blade Element Momentum (BEM) theory, offer an in-depth understanding of the aerodynamic characteristics of turbines but require significant computer resources [17]. Empirical models utilise historical data to build and establish statistical relationships. However, they frequently fall short in capturing the non-linear interactions among different atmospheric and operational elements [18]. Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) have been used to forecast wind power production under different circumstances. These algorithms utilise historical data to learn complex patterns, potentially enhancing predictions compared to traditional methods [19].

Despite their advantages, AI-based surrogate models also present certain limitations:

- Data Dependency: The performance of these models heavily relies on the quantity, quality, and variety of the training data. Inadequate or noisy data can significantly impair model accuracy.
- Generalizability: AI models often struggle to generalize beyond the specific conditions or geographic locations they were trained on. This poses challenges in deploying these models in different wind farms with distinct environmental characteristics.
- Interpretability: Many AI models, particularly deep learning networks, act as ”black boxes”, providing little insight into the underlying mechanisms driving their predictions.

This lack of transparency can hinder their acceptance and troubleshooting in operational settings.
- Computational Demands: While less intensive than detailed physical simulations, training sophisticated AI models can still demand substantial computational resources, especially when handling large datasets or performing real-time analysis [20]

While the promise of AI-based surrogate models for predicting the power of wind turbines in real-world scenarios is undeniable, addressing these limitations is essential for their practical applicability and reliability in real-world scenarios.

III. Methodology

A. Data Collection and Preprocessing

The model presented in this research was created utilising data from a Supervisory Control and Data Acquisition (SCADA) system in an operational wind turbine located in Turkey. SCADA systems are used in wind turbine power systems to track parameters and evaluate the operating state of the turbines. SCADA systems in wind turbines record and measure data at 10-minute intervals, including wind direction, speed, and generated power. To guarantee quality and reliability for modelling purposes, the first dataset has undergone rigorous preprocessing steps. The 50,530 values in the dataset utilised for this research have significant attributes including Date/Time, which provide timestamps demonstrating the exact moment the measurements were taken. Because of the time series nature of wind data, lost values have been identified and corrected with linear interpolation. To focus on variables directly affecting the efficiency of turbines, feature selection was based on correlation coefficients and knowledge in the area concerned. Lastly, to aid convergence during model training, all features have been normalized to have zero mean and unit variance.

B. Exploratory Data Analysis (EDA)

The exploratory data analysis revealed several insights into the underlying dynamics of wind turbine operations. Scatter plots in Figure 1 were used to identify relationships between power output and wind speed, showing a non-linear dependency. Univariate analysis in Figure 2 highlighted seasonal variations in wind speeds and directions. The correlation matrix in Figure 3, visualized through heatmaps, assisted in understanding the multicollinearity between features.

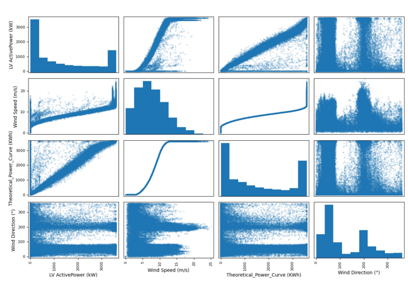

Fig. 1. Exploratory Data Analysis of Wind Turbine Operations: Scatter plots.

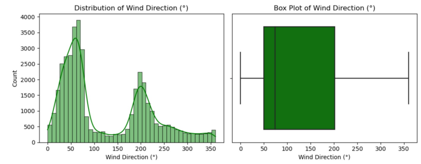

Fig. 2. Distribution and Box Plot of Wind Direction (°).

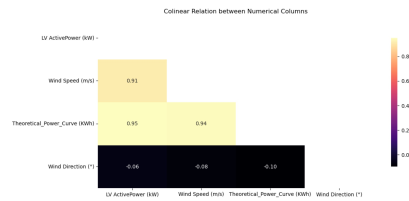

Fig. 3. Correlation Matrix of Numerical Columns.

C. Model Development

Advanced AI-based surrogate models were developed to enhance prediction accuracy and reliability. To ensure optimal performance, these models use robust regression techniques and sophisticated machine learning algorithms which are equipped with stringent hyperparameter tuning and validation methods.

This study examined several types of machine learning algorithms, chosen for their distinct advantages in handling large datasets and complex nonlinear patterns:

- Gradient Boosting Regressor: This ensemble model sequentially increases prediction accuracy by optimising for loss functions by combining decision trees through a boosting strategy.
- Support Vector Regression (SVR): SVR is useful in high-dimensional spaces and fits the error inside a predetermined threshold by utilising the principles of Support Vector Machines (SVM) for regression applications.
- Random Forest Regressor: The Random Forest Regressor is an ensemble technique that employs several decision trees and averages their forecasts to reduce overfitting and enhance overall prediction accuracy.
- Linear Regression: Provides a straightforward approach for modelling relationships between dependent and independent variables through a linear equation.
- Extra Trees Regressor: This model lowers the variance by averaging the results of random splits of all observations, much as random forests.
- AdaBoost Regressor: The AdaBoost Regressor increases decision tree performance by concentrating on challenging instances and adding models until a threshold is met for accuracy or the number of models.
- Decision Tree Regressor: Tree-like representation of decisions and their potential outcomes. Although Decision Tree Regressions are easy to comprehend and interpret, overfitting is a possibility.
- XGBoost Regressor: The XGBoost Regressor is a fast and efficient gradient-boost decision tree solution that works

particularly effectively with large datasets.

- CatBoost Regressor: This algorithm applies symmetric trees to handle categorical variables and optimises computations over high-dimensional data.

Grid search and randomised search methods were combined to optimise the above models. In order to identify the optimal configurations, these techniques investigated a broad variety of model parameters, including the number of trees, tree depth, learning rate, and regularisation terms. For example, the predictive accuracy of the CatBoostRegressor was much improved by fine-tuning it using RandomizedSearchCV over parameters like as learning rate, tree depth, and iterations.

D. Model Training and Evaluation

To guarantee the accuracy and dependability of the models, a thorough training and assessment methodology is incorporated into the creation and validation of AI-based surrogate models for wind turbine power production prediction. Data is first divided into training and testing sets, allocating 80% of the data for training and 20% for evaluation. This maintains a subset for objective performance evaluation while guaranteeing adequate data availability for model training. To determine which factors have the greatest influence after training, a feature importance evaluation is carried out. This helps to improve the models and comprehend how they make decisions.

K-fold cross-validation was used to thoroughly validate the models, guaranteeing their resilience and generalizability across various data segments. With this approach, the data is divided into 'k' subsets. The remaining subset is used for testing, and the model is iteratively trained on 'k-1' subsets. This procedure averages the model's accuracy over several folds, which reduces overfitting and aids in evaluating the model's performance across a range of data samples.

The effectiveness of the surrogate models was quantitatively assessed using R^2 (coefficient of determination) and RMSE (root mean squared error). The CatBoostRegressor emerged as the top performer, achieving an R^2 score of approximately 98.25%, indicating an excellent fit to the data variability. The evaluation highlighted the capability of AI-based models to predict wind turbine power output with high precision, providing a critical tool for optimizing the operation and maintenance of wind energy systems.

IV. Results and Discussion

A. Model Performance

The surrogate models employed in the study demonstrated exceptional performance metrics, as illustrated in Figure 4. The CatBoostRegressor emerged as the top performer with an R^2 score of 98.25% and an RMSE of 172.62, signifying outstanding predictive accuracy. This model, along with others like the XGBRegressor and ExtraTreesRegressor, which also showcased high R^2 scores above 97%, significantly outperformed traditional modelling techniques. These results underscore the efficacy of advanced regression and ensemble methods in enhancing the precision of power output predictions in wind turbines.

	Model-Name	R2_score	RMSE
9	CatBoostRegressor	98.253759	172.616020
7	XGBRegressor	97.770481	195.044879
4	ExtraTreesRegressor	97.760676	195.473301
2	RandomForestRegressor	97.399257	210.657873
6	DecisionTreeRegressor	95.528001	276.235730
0	GradientBoostingRegressor	94.777728	298.509905
8	XGBRFRegressor	94.474649	307.049873
3	LinearRegression	90.607095	400.339970
1	SVR	88.491109	443.144887
5	AdaBoostRegressor	86.857825	473.546488

Fig. 4. Comparative Performance of AI-Based Surrogate Models

B. Visualization of Results

The efficacy of the AI-based surrogate models in forecasting wind turbine power production is illustrated in Figure 5. The plot shows how the predicted power (orange dots), real power output (blue dots), and theoretical power curve (purple line) vary with wind speed using the CatBoostRegressor model. The visualisation shows how accurate the model is at predicting, especially when it comes to how the predicted values capture the variability observed in real power outputs while staying relatively close to the theoretical power curve.

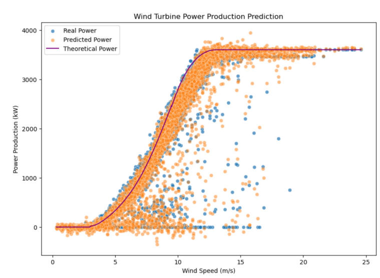

Fig. 5. Wind Turbine Power Production Prediction

The graph in Figure 5 shows the regions where the model performs well and where there are differences, especially at higher wind speeds when certain predicted values differ from measured values. These kinds of insights play a critical role in assessing the robustness and reliability of the model and demonstrating its practical applications. The model's ability to accurately represent dynamic environmental interactions is supported by the data points' distribution over the speed spectrum and its distinct trend lines, which aid in understanding the impact of wind speed on power output.

C. Discussion

The analysis of the models reveals their robust capability to handle complex nonlinear patterns in wind speed and power output data, which are often challenging for traditional models. The superior performance of the CatBoost and XGB models can be attributed to their algorithmic efficiency in managing categorical features and large datasets.

D. Implications for Industrial Processes

The integration of these AI-based models into wind turbine power prediction processes promises significant enhancements in operational efficiency and energy management. By accurately forecasting power outputs, these models can help optimise maintenance schedules and reduce unnecessary expenditures, thereby boosting the overall profitability and sustainability of wind energy projects.

V. Conclusion

A. Summary of Findings

The research study successfully demonstrated how AI-based surrogate models can forecast wind turbine power output with a high degree of accuracy.

- High Predictive Accuracy: When compared to traditional models, models like the CatBoostRegressor got a R^2 value of 98.25%, suggesting a strong fit to the data and superior predictive accuracy.
- Handling complicated Non-linear Relationships Effectively: One important factor in the variable conditions typical of wind energy production is how well the models handled the complicated non-linear relationships between wind speed and power output.
- Visual Validation: The study's plots and graphs visually verified that the models' predictions, especially when considering different wind speeds, closely match real data and theoretical expectations.

B. Future Work

Future research directions and potential improvements include:

- Real-time Data Integration: Adding operational and environmental data in real-time to the models could improve their accuracy and flexibility, especially when predicting in quickly changing settings.
- Advanced Algorithmic Approaches: To solve the constraints encountered at higher wind speeds and other complicated circumstances, research into the usage of hybrid models that integrate machine learning and physical laws, as well as deeper learning architectures, may be conducted.
- Interdisciplinary Studies: By working together across disciplines to incorporate more mechanical and meteorological insights into the modelling process, it may be possible to further improve accuracy and provide a more comprehensive method for forecasting and controlling the performance of wind turbines.

Rights Retention

References

[1] Q. Qi, F. Tao, T. Hu, N. Anwer, A. Liu, Y. Wei, L. Wang, and A. Y. C. Nee, "Enabling technologies and tools for digital twin," Journal of Manufacturing Systems, vol. 58, pp. 3–21, 2021, Elsevier.

[2] T. Tripura, A. S. Desai, S. Adhikari, and S. Chakraborty, "Probabilistic machine learning based predictive and interpretable digital twin for dynamical systems," Computers & Structures, vol. 281, pp. 107008, 2023, Elsevier.

[3] A. Thelen, X. Zhang, O. Fink, Y. Lu, S. Ghosh, B. D. Youn, M. D. Todd, S. Mahadevan, C. Hu, and Z. Hu, "A comprehensive review of digital twin—part 1: modelling and twinning enabling technologies," Structural

[4] M. Liu, S. Fang, H. Dong, and C. Xu, "Review of digital twin about concepts, technologies, and industrial applications," Journal of Manufacturing Systems, vol. 58, pp. 346–361, 2021, Elsevier.

[5] M. G. Kapteyn and K. E. Willcox, "From physics-based models to predictive digital twins via interpretable machine learning," arXiv preprint arXiv:2004.11356, 2020.

[6] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," The International Journal of Advanced Manufacturing Technology, vol. 94, pp. 3563–3576, 2018, Springer.

[7] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentafat, and S. Bakiras, "The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities," IEEE Access, vol. 9, pp. 32030–32052, 2021, IEEE.

[8] M. Atalay, U. Murat, B. Oksuz, A. M. Parlaktuna, E. Pisirir, and M. C. Testik, "Digital twins in manufacturing: Systematic literature review for physical–digital layer categorization and future research directions," International Journal of Computer Integrated Manufacturing, vol. 35, no. 7, pp. 679–705, 2022, Taylor & Francis.

[9] S. Garg, A. Gogoi, S. Chakraborty, and B. Hazra, "Machine learning based digital twin for stochastic nonlinear multi-degree of freedom dynamical system," Probabilistic Engineering Mechanics, vol. 66, pp. 103173, 2021, Elsevier.

[10] Z. Huang, Y. Shen, J. Li, M. Fey, and C. Brecher, "A survey on AI-driven digital twins in industry 4.0: Smart manufacturing and advanced robotics," Sensors, vol. 21, no. 19, pp. 6340, 2021, MDPI.

[11] S. M. Zayed, G. M. Attiya, A. El-Sayed, and E. Hemdan, "A review study on digital twins with artificial intelligence and internet of things: concepts, opportunities, challenges, tools and future scope," Multimedia Tools and Applications, pp. 1–27, 2023.

[12] X. Lai, L. Yang, X. He, Y. Pang, X. Song, and W. Sun, "Digital twin-based structural health monitoring by combining measurement and computational data: An aircraft wing example," Journal of Manufacturing Systems, vol. 69, pp. 76–90, 2023.

[13] R. Smith, "Chemical Process Design and Integration," John Wiley, 2005.

[14] D. Jones and A. Harrison, "The Relevance of Empirical Modeling in Production Engineering," International Journal of Production Research, 2000.

[15] T. Ambreen, N. Ikram, M. Usman, and M. Niazi, "Empirical research in requirements engineering: trends and opportunities," Requirements Engineering, vol. 23, pp. 63–95, 2018, Springer.

[16] Y. Cengel and J. Cimbala, "Ebook: Fluid mechanics fundamentals and applications (SI units)," McGraw Hill, 2013.

[17] J. G. Slootweg, S. W. H. De Haan, H. Polinder, and W. L. Kling, "General model for representing variable speed wind turbines in power system dynamics simulations," IEEE Transactions on Power Systems, vol. 18, no. 1, pp. 144–151, 2003, IEEE.

[18] M. Lydia, S. S. Kumar, A. I. Selvakumar, and G. E. P. Kumar, "A comprehensive review on wind turbine power curve modeling techniques," Renewable and Sustainable Energy Reviews, vol. 30, pp. 452–460, 2014, Elsevier.

[19] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The State-Of-The-Art in Short-Term Prediction of Wind Power, A Literature Overview, Joint deliverable report of ANEMOS. plus and SafeWind European projects," PhD thesis, MINES Paris Coordinator of ANEMOS. plus and SafeWind projects, 2011.

[20] K. L. Wagstaff et al., "Machine Learning for Space Weather: As Applied to the Sun to the Earth," Space Weather, 2012.