

Effat University Repository

Wind Farm Layout Optimization: Genetic Algorithms, Machine Learning, and Bibliometric Insights

Authors	Brahimi, Tayeb;Asfour, Rawan;Fathy El-Amin, Mohammed
Citation	NA
DOI	https://www.taylorfrancis.com/chapters/edit/10.1201/9781003470649-10/wind-farm-layout-optimization-genetic-algorithms-machine-learning-bibliometric-insights-tayeb-brahimi-rawan-asfour-mohammed-fathy-el-amin
Publisher	Taylor & Francis
Download date	2026-05-16 08:09:37
Link to Item	https://repository.effatuniversity.edu.sa/handle/20.500.14131/2653

Advancing Wind Farm Layout Optimization: Genetic Algorithms, Machine Learning, and Bibliometric Insights

Tayeb Brahimī Rawan Asfour Mohamed F. El Amin

Effat College of Engineering Energy and Technology Research Center, Effat University, Jeddah, Saudi Arabia

Abstract

The optimization of wind farm layouts (WFLO) is pivotal in maximizing energy production while reducing operational costs and environmental impact. This chapter presents a comprehensive framework that investigates Genetic Algorithms (GA) and Machine Learning (ML) techniques for WFLO. A parametric GA-based simulation explores the influence of key parameters, such as population size, hub height, surface roughness, and cost models, on layout efficiency and power output. To enrich this computational foundation, a bibliometric analysis was performed covering 2015-2025 using cleaned and merged data from Scopus and Web of Science, involving 483 unique documents. The analysis classifies WFLO research into major optimization categories such as GA, ML, PSO, and hybrid techniques, and reveals an increasing reliance on data-driven models and hybrid frameworks. A keyword heatmap confirms the growing prominence of ML algorithms like ANN, SVM, and Random Forest in WFLO. The results highlight the evolution of optimization strategies, the synergy between GA and ML, and the emergence of intelligent layout designs that combine predictive accuracy with computational efficiency.

Keywords— Genetic Algorithm, Wind Farm, WFLO, Machine Learning, Optimization.

1 Introduction

The global urgency for climate action has significantly accelerated the deployment of renewable energy, with wind energy experiencing substantial growth over the past decade. Installed wind capacity increased from 349.42 GW in 2014 to more than 1 TW globally in 2024 [1], [2], reflecting over 185% growth. Offshore wind, a rapidly expanding segment, grew from 7.1 GW in 2014 to 75.2 GW in 2023, and now accounts for 7.4% of total wind capacity [1], [3]. Despite this progress, market concentration in key regions such as China, the EU, the US, and India highlights the need for broader global participation, particularly from the Global South.

Wind farms, comprising multiple turbines strategically placed to capture wind energy, are central to this expansion. As of 2024, there are 41,850 operational wind farms worldwide, with an additional 1,271 offshore projects planned [4], [5]. The year 2025 is expected to set a record for offshore wind installations [5]. However, optimizing wind farm layouts (WFLO) remains a critical challenge due to wake effects, where upstream turbines reduce wind speed and turbulence intensity for downstream turbines. These interactions can result in power generation losses of up to 20%, and in some configurations, even up to 30% [6]. This substantial loss underscores the importance of accurate modeling of wake interactions and the design of efficient wind farm layouts that can mitigate such effects.

To address these challenges, this chapter presents a Genetic Algorithm (GA)-based WFLO framework designed to optimize turbine placement and enhance overall wind farm performance. A parametric study is conducted to assess the influence of key factors, including population size, hub height, and surface roughness, on the effectiveness and adaptability of the GA optimization process. These findings offer valuable insights into how evolutionary algorithms can refine wind farm design and support the development of more efficient computational models.

In parallel with GA, recent advancements in machine learning (ML) have introduced new possibilities for data-driven optimization in wind energy systems [7]–[9]. ML techniques—particularly deep learning and predictive modeling—are increasingly being employed to simulate complex wake dynamics, forecast wind behavior, and enhance layout design strategies. In light of these developments, the chapter also includes a bibliometric analysis that highlights emerging research trends in ML applications for wind energy. The analysis reveals a growing shift

toward intelligent hybrid strategies, where ML is integrated with traditional optimization frameworks to improve accuracy, scalability, and adaptability.

The remainder of the chapter is structured as follows. Section 2 provides a comprehensive review of state-of-the-art techniques in wind farm layout optimization, covering both deterministic and computational intelligence methods, as well as recent machine learning developments. Section ?? outlines the methodology used in this study, detailing the GA-based WFLO model and the bibliometric approach. Section 5 presents the results of the optimization and parametric analyses, along with insights from the bibliometric study. Finally, Section 6 summarizes the key findings and suggests directions for future research, particularly in combining ML with GA-based optimization frameworks.

2 Literature Review

Wind Farm Layout Optimization (WFLO) is a complex, multidisciplinary research area that plays a critical role in the evolution of modern wind energy systems. The process of selecting optimal wind turbine placement is influenced by numerous factors, including environmental conditions, wake interactions, and site-specific constraints. As electricity demand rises globally, wind energy has emerged as one of the fastest-growing renewable energy sources. Its expansion is supported by global efforts to reduce reliance on fossil fuels and transition to sustainable energy systems. To assess the state-of-the-art in Wind Farm Layout Optimization (WFLO) and the emerging integration of Machine Learning (ML) and Artificial Intelligence (AI) techniques, a systematic literature review and bibliometric analysis were conducted covering publications from 2015 to 2024 in the Web of science database. The search query was designed to capture the intersection of machine learning and optimization techniques with wind farm layout challenges. Only peer-reviewed journal articles and conference proceedings were included to ensure the quality and relevance of the data. The search yielded a robust set of publications that reflect both the classical and emerging computational approaches to Wind Farm Layout Optimization (WFLO), especially highlighting the increased use of machine learning and hybrid techniques over the past decade.

The literature reveals that Wind Farm Layout Optimization (WFLO) has traditionally been tackled with traditional deterministic methods [10], [11] such as linear programming (LP) and mixed-integer linear programming (MILP), which are often used to design WFLO. These methods depend on basic assumptions that do not fully account for variability in critical parameters such as wind speed, the impact of turbine wakes, or the nature of the terrain complexity. They typically use simplified wake models and fixed parameters, which limit their flexibility and scalability, especially in large-scale wind farm layouts. As a result, they can be computationally expensive and may produce less-than-ideal layouts for large-scale wind farms [12]. In contrast, Computational Intelligence (CI) techniques [13] such as Genetic Algorithms (GA) [14], Particle Swarm Optimization (PSO) [15], Ant Colony Optimization (ACO) [16], Simulated Annealing (SA) [17], and Artificial Neural Networks (ANN) [18] adopt a more adaptive and nature-inspired approach and offer higher flexibility and scalability. For example, evolutionary algorithms such as Genetic Algorithms (GA) [19], [20] test and evolve many turbine arrangements over time, adapting to variations in wind, wakes, and terrain to find practical wind farm layouts that effectively balance energy output and cost. In fact, wake effects, caused by the interaction of upstream and downstream turbines, remain one of the most critical limitations in WFLO. Modern research emphasizes refined wake modeling using computational fluid dynamics (CFD) and machine learning to capture atmospheric stability, turbulence intensity, and turbine arrangement impacts. The review article by Wang et al [21] underscores the transition from analytical models to sophisticated CFD and hybrid approaches, reflecting the ongoing efforts to optimize wind farm performance through improved wake modeling techniques. Studies have shown that turbine spacing and alignment significantly impact the energy yield, with staggered layouts often performing better than aligned layouts by minimizing wake overlap and turbulence intensities. Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Whale Optimization Algorithm (WOA) are emerging as prominent tools for handling the complex, nonlinear nature of WFLO problems [20], [22]–[24]. These stochastic methods are well-suited for navigating the multimodal search spaces typical in WFLO, where traditional deterministic showed low performance [25]. Hybrid methods combining genetic algorithms with reinforcement learning (RLGA) [26] demonstrate a convergence that is three times faster than traditional GAs for complex layouts. El Mestari [27] underscored the importance of systematic parameter tuning in wind farm layout optimization by applying a genetic algorithm enhanced through a full factorial Design of Experiments (DOE) approach. The study achieved notable improvements in power output and turbine placement efficiency, validating that data-driven tuning methods outperform traditional trial-and-error approaches, especially in constrained environments with wake effects. The review study by ElJaadi [28] identified hybrid metaheuristics as a promising future direction in wind farm layout optimization (WFLO). The review emphasizes that combining multiple algorithms can help overcome individual limitations, such as local optima or slow convergence, seen in single-method approaches. The authors noted that

hybrid methods consistently outperform standalone algorithms in terms of efficiency, wake integration, solution quality, and convergence speed. These methods are especially advantageous for complex, multi-objective WFLO problems and are adaptable to different wind scenarios, such as constant vs. variable wind speed and direction, allowing for more realistic and effective WFLO. According to comprehensive benchmarking study Thomas et al. [29], they stated that while metaheuristic methods offer robust performance for WFLO, no single method is universally best. Instead, hybrid or ensemble approaches may provide better balance between accuracy, speed, and robustness for practical wind farm layout design.

Recent advancements in machine learning (ML) have further enhanced the capabilities of these optimization techniques, enabling better wake modeling and real-time adaptation to environmental conditions [8], [30]. Bouabdallaoui et al. [31] applied four AI models—SVM, Decision Trees, ANFIS, and ANN to predict wind energy output at a wind farm in Yalova, Turkey, using wind speed and direction as inputs. The author's objective was to improve the weaker models by integrating them with metaheuristic optimization algorithms. SVM performed best with a R^2 of 0.95. Ayele et al. [32] compared six machine learning models, models—Elastic Net, Random Forest, SARIMA, XGBoost, Prophet, and a combined Prophet-XGBoost, to forecast long-term power production at the Adama II wind farm using six years of SCADA data (Supervisory Control and Data Acquisition). The results showed that SARIMAX excelled in short-term forecasting while the combined Prophet-XGBoost model achieved the best long-term accuracy for yearly predictions. In a review of applications of AI algorithms in wind farms, Wang et al. [33] reported that machine learning techniques, especially ANNs, SVMs, and hybrid models are highly effective for prediction-based optimization in wind farms. They noticed that ANNs are favored for their adaptive learning, fault tolerance, and high predictive accuracy, making them ideal for nonlinear, uncertain environments like wind farms, while SVMs are more precise with small datasets but require careful parameter tuning, and hybrid methods outperform individual models due to better feature extraction and parameter optimization. The recent review study of various ML techniques used for wind power prediction by Liu et al. [34], highlights the effectiveness of hybrid models that combine different ML algorithms or integrate ML with statistical methods. It also discussed several challenges faced in applying ML, including the variability of wind, the lack of high-quality data, and the complexity of modeling nonlinear relationships.

To synthesize the findings from the reviewed literature and clarify the landscape of optimization techniques applied to WFLO, Figure 1 provides a comprehensive classification of methods. It categorizes existing approaches into four major groups: heuristic algorithms, metaheuristic algorithms, artificial intelligence (AI)-based techniques, and hybrid methods. This hierarchical organization reflects the methodological evolution in the field from basic rule-based algorithms to sophisticated hybrid frameworks that integrate machine learning and evolutionary computation. The inclusion of emerging strategies such as Reinforcement Learning (RL), Whale Optimization Algorithm (WOA), and CFD-ML hybrids illustrates the growing complexity and adaptability of modern WFLO approaches. This flowchart serves as a visual summary of the tools currently shaping the optimization of wind farm layouts and offers a foundation for identifying research gaps and future directions.

Overall, a wide range of optimization techniques have been developed for Wind Farm Layout Optimization (WFLO), several important gaps remain. Many existing studies rely on idealized assumptions, with limited consideration of flowfield complexity, site-specific parameters, or terrain-induced variability. In particular, aspects such as population size, hub height, and surface roughness—which are crucial for real-world performance—are often overlooked or simplified. Additionally, there is a lack of systematic parametric analysis to understand how such factors influence the performance and adaptability of optimization algorithms, especially in large-scale scenarios. Another key gap lies in the disconnect between traditional optimization methods and emerging machine learning (ML) techniques. While Genetic Algorithms (GA) and other evolutionary algorithms have shown strong capabilities in navigating the nonlinear search space of WFLO problems, their predictive modeling capacities remain limited. In contrast, ML models offer enhanced potential for learning wake behavior, forecasting power output, and adapting to complex environmental conditions. However, these approaches have yet to be fully integrated into mainstream WFLO research. The integration of machine learning for real-time WFLO and adaptive control remains a significant research opportunity [8].

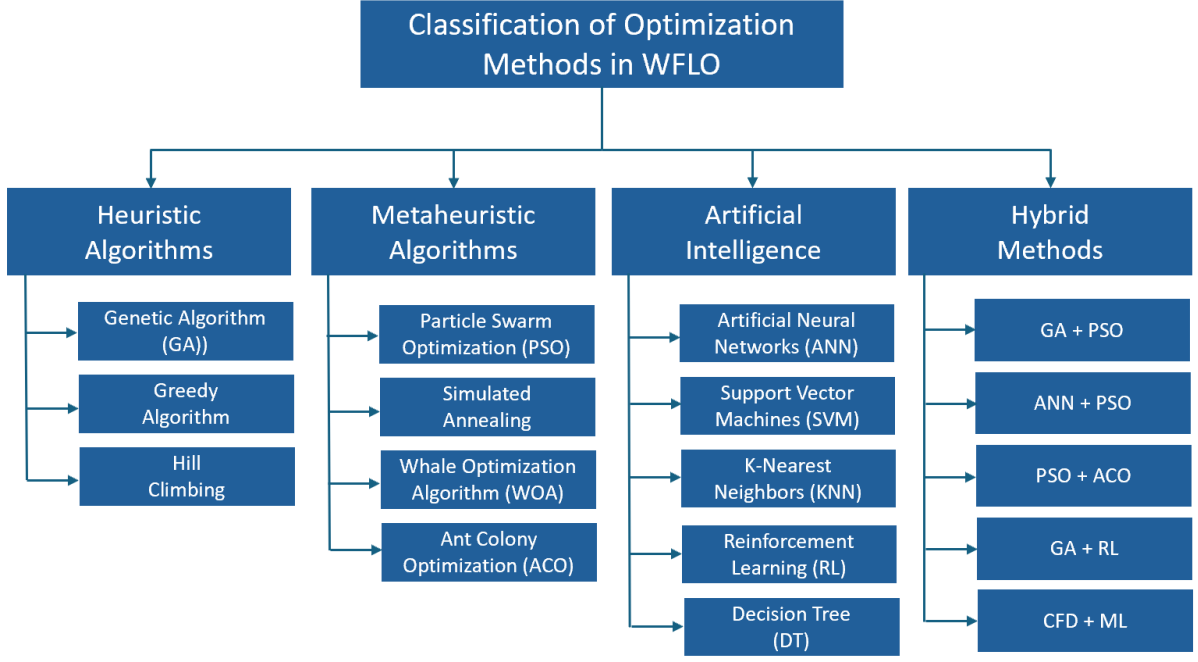


Figure 1: Classification of optimization methods found in the reviewed papers

To address these gaps, this chapter presents a GA-based framework tailored to wind farm layout optimization, with a focus on evaluating its performance under varying parameter settings such as population size, hub height, and surface roughness. The findings offer valuable insight into the operational behavior and flexibility of GA in real-world conditions. Furthermore, this study includes a bibliometric analysis that captures the recent surge of interest in machine learning applications within wind energy optimization. By synthesizing trends and emerging methods, the analysis supports the development of future hybrid strategies that combine GA with ML for enhanced prediction, optimization, and overall energy system sustainability.

3 Mathematical Modeling and Optimization Method

This study presents a comprehensive mathematical and computational framework to optimize the layout of onshore horizontal-axis wind turbines (HAWTs) in wind farms. The objective is to maximize total energy output while minimizing the cost of energy (COE). The proposed methodology integrates wake modeling, cost modeling, and genetic algorithm (GA)-based optimization, all implemented within a MATLAB environment.

3.1 Wake Effect Modeling

To estimate the wake losses caused by upstream turbines, the analytical Jensen wake model is applied **jensen1983**. The model, illustrated in Figure 2, is based on momentum conservation and assumes a linear expansion of the wake downstream of the turbine. The wind speed u_j experienced at a location j behind turbine i , after accounting for the wake deficit, is given by:

$$u_j = u_0 (1 - u_{\text{def},ij}) = u_0 \left[1 - \left(\frac{2a}{(1 + \alpha \frac{x_{ij}}{r})^2} \right) \right] \quad (1)$$

where: u_0 is the free-stream wind speed, a is the axial induction factor, α is the wake decay (entrainment) constant, x_{ij} is the downstream distance between turbines i and j , r is the initial wake radius (equal to the rotor radius).

The axial induction factor a is calculated from the thrust coefficient C_T as:

$$a = \frac{1 - \sqrt{1 - C_T}}{2} \quad (2)$$

The wake expansion constant α , which determines the growth rate of the wake, depends on the hub height h and surface roughness length z_0 of the terrain:

$$\alpha = \frac{0.5}{\ln(h/z_0)} \quad (3)$$

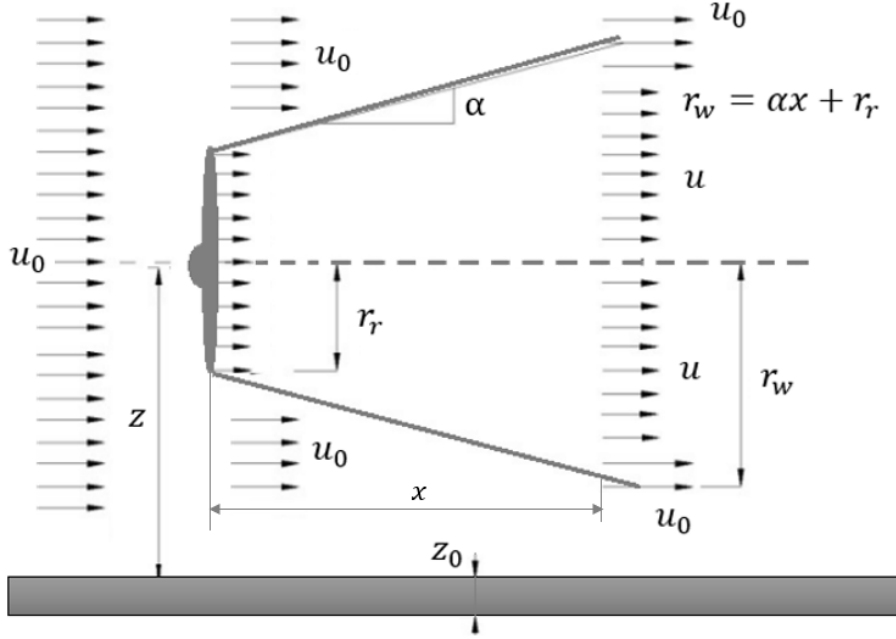


Figure 2: Schematic representation of the Jensen wake model.

When multiple upstream wakes influence a turbine, the effective wind speed u_i at turbine i is obtained by superimposing the kinetic energy deficits from all relevant upstream turbines using the following expression:

$$u_i = u_{0i} - \left(\sum_{j=1}^{N_T} (u_{0j} - u_{ij})^2 \right)^{1/2} \quad (4)$$

where: N_T is the number of upstream turbines affecting turbine i , u_{0j} is the free-stream wind speed before turbine j , u_{ij} is the wake-influenced wind speed from turbine j at the location of turbine i .

This model enables the prediction of wake-induced energy losses in large wind farms and is often used within optimization algorithms to improve turbine placement and farm efficiency.

Figure 3 provides a 3D visualization of wake velocity in the axial direction, offering insight into how wind speed recovers with downstream and radial distance from a turbine. This surface plot captures the velocity field resulting from wake expansion and mixing with ambient flow, with wind speed increasing as the flow moves away from the turbine. The gradient from blue to red illustrates this velocity recovery, moving toward the undisturbed free-stream wind speed. The observed behavior confirms the fundamental principles of wake aerodynamics, where velocity deficits are greatest near the turbine and diminish further downstream. Such visualizations are critical for assessing wake recovery and guiding optimal spacing strategies in wind farm layout planning.

To further illustrate the behavior of wake interactions within a wind farm, Figure 4 presents a simulation generated using the Jensen wake model. In this model, the wake expands linearly downstream, and the velocity deficit is derived based on the thrust coefficient C_T , which reflects the energy extracted by the turbine rotor. The color gradient in the figure indicates the variation in wind speed, with cooler tones representing areas of higher velocity deficit. For this simulation, key parameters include a thrust coefficient $C_T = 0.75$, representative of modern utility-scale turbines, a wake decay constant $k = 0.0813$ associated with moderate onshore turbulence, and a free-stream wind speed $u_0 = 8$ m/s. The wake expansion is modeled as $D_w = D(1 + 2ks)$, where D is the rotor diameter (30 m) and s is the downstream distance normalized by rotor diameter.

The color scale highlights the reduced wind speed, underlining the importance of optimal turbine spacing in minimizing wake losses and improving overall wind farm efficiency. The figure demonstrates how downstream turbines experience a significant reduction in wind speed when located within the wake of upstream units. This highlights the critical importance of wake-aware layout optimization to ensure high energy efficiency and minimal performance loss across the farm.

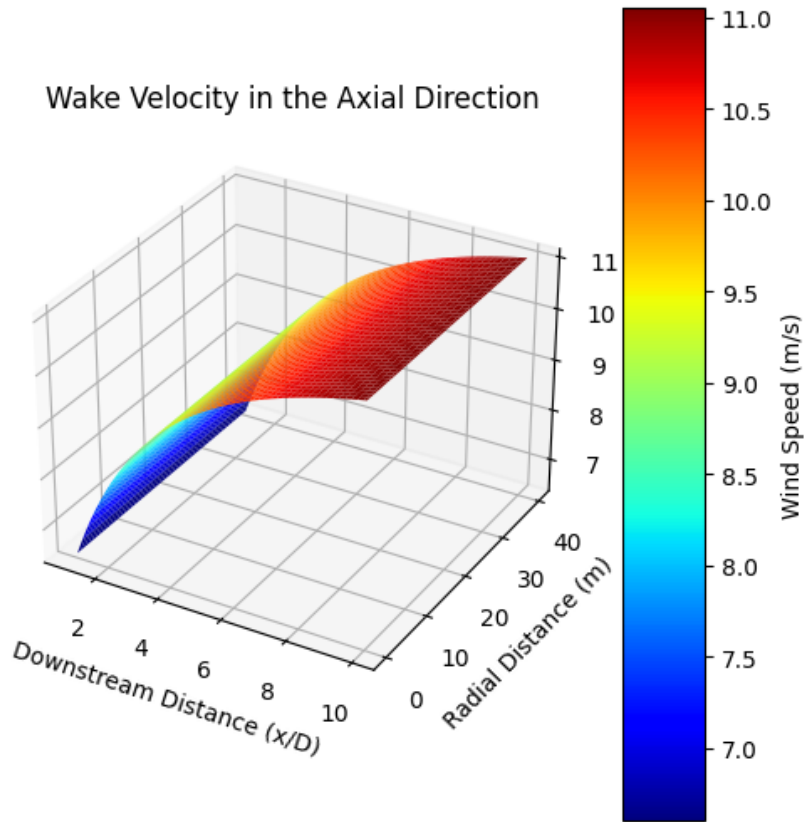


Figure 3: Wake velocity downstream of the wind turbine.

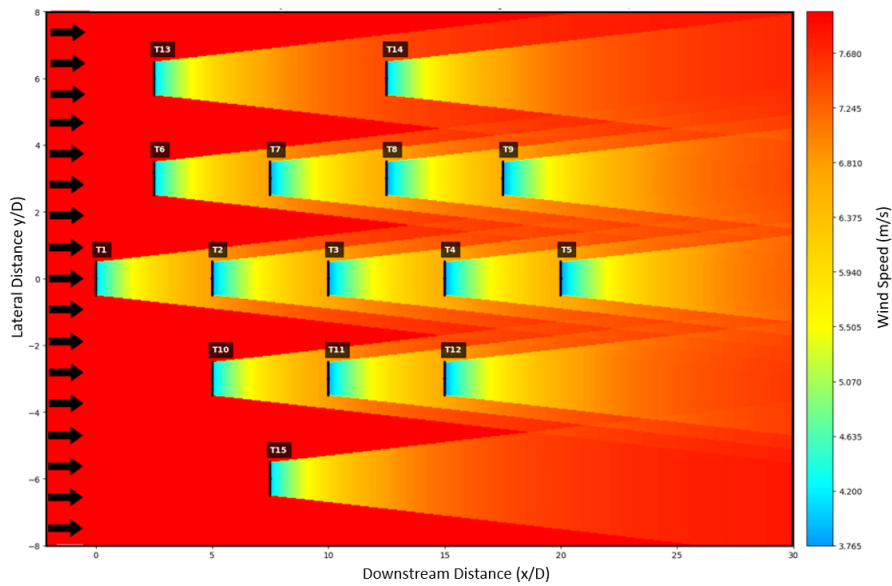


Figure 4: Visualization of multiple turbine wakes using the Jensen model.

3.2 Power and Efficiency Modeling

The total farm power is the sum of individual turbine powers. Between the cut-in speed and rated speed, the power P (in kW) is given by:

$$P_{\text{total}} = \sum_{i=1}^{N_T} 0.3u_i^3 \quad (5)$$

The wind farm efficiency can be calculated as the ratio of the energy produced by N_T turbines in a wind farm to the energy produced by N isolated wind turbines.

$$\text{Efficiency} = \eta = \frac{P_{\text{total}}}{N_T P} = \frac{\sum_{i=1}^{N_T} 0.3 \times u_i^3}{N_T (0.3 \times u_0^3)} \quad (6)$$

3.3 Cost Modeling

Two cost models are used in this study to compare the influence of simplified versus detailed economic formulations on layout optimization outcomes. The first is the simple empirical model proposed by Masetti [20], commonly used in wind farm optimization literature. It estimates cost as a nonlinear function of the total number of turbines N_T :

$$\text{Cost} = N_T \left(\frac{2}{3} + \frac{1}{3} e^{-0.00174N_T^2} \right) \quad (7)$$

This formulation provides a convenient cost approximation for farms using identical turbines and is widely adopted in benchmark studies.

The second, more complex model is adopted from Chen [35]. It is specifically designed to account for multiple types of wind turbines. It incorporates rated turbine power, number of turbines, and their interactions. The cost of the i th wind turbine type is calculated using:

$$\text{Cost}_i = -0.1539 \times P_{r_i} - 0.001 \times N_i + 2 \times P_{r_i} \times N_i + 0.2504 \quad (8)$$

where P_{r_i} is the rated power of the i th wind turbine, and N_i is the number of i th turbine type installed in the wind farm. When multiple turbine types are used, the total cost is computed as

$$\text{Cost}_{\text{tot}} = \sum_{i=0}^n \text{Cost}_i \quad (9)$$

where n is the number of different turbine types considered in the layout.

3.4 Objective Function

The objective function is designed to balance the trade-off between minimizing installation and operational costs and maximizing total power production. It is formulated as the cost per unit of energy generated by the wind farm:

$$\text{Objective} = \frac{\text{Cost}}{P_{\text{total}}} \quad (10)$$

This objective function encourages configurations that not only increase energy output but also penalize excessive turbine counts or suboptimal placements that lead to diminishing returns. By minimizing this ratio, the optimization framework promotes economically efficient wind farm layouts. The genetic algorithm iteratively evolves candidate solutions to reduce this metric, thus achieving a design that is both cost-effective and energy-efficient.

3.5 Genetic Algorithm Framework

A Genetic Algorithm (GA) is a computational method that evolves solutions over time, drawing inspiration from biological evolution [36]. It operates by initializing a population of candidate solutions and refining them through processes similar to natural selection, crossover, and mutation. Originally introduced by John Holland in the 1960s [37] and expanded by David Goldberg in the 1980s [38], GAs are designed to iteratively enhance the quality of solutions across generations. Their success heavily relies on tuning algorithm parameters to the specific problem, allowing effective exploration of complex solution spaces. Often, GAs are integrated with other

optimization methods to increase their flexibility and strength. Each candidate solution is structured like a chromosome, engaging in evolutionary exchanges to improve performance. The adaptability and efficiency of GAs have made them valuable tools across a wide spectrum of optimization challenges.

In this study, a customized Genetic Algorithm (GA) was developed and implemented in MATLAB to optimize wind farm layouts. The optimization process begins by generating an initial population based on a staggered mesh grid, with turbines spaced 5 rotor diameters (5D) apart. Each individual in the population represents a specific turbine configuration, and its fitness is evaluated by taking the inverse of its relative error compared to the minimum objective function value.

The GA operates with the following parameters: a population size of 1500 individuals, 100 generations, a crossover rate of 0.5, and a mutation rate of 0.02. Tournament selection is used to choose parents for reproduction. Crossover is performed using a single-point method, while mutation randomly flips turbine states to introduce diversity into the population.

Figure 5 shows the basic structure of a Genetic Algorithm (GA) begins with the generation of an initial population, either randomly or using heuristic methods. From this population, individuals are selected as parents, and crossover and mutation operators are applied to produce offspring. These new solutions replace existing individuals, and the process repeats iteratively. This evolutionary approach is inspired by natural selection, where fitter individuals are more likely to propagate their traits. The three key genetic operators—selection, crossover, and mutation—drive the exploration and exploitation of the solution space. Figure 3.7 illustrates the main steps of the GA process applied to wind farm layout optimization.

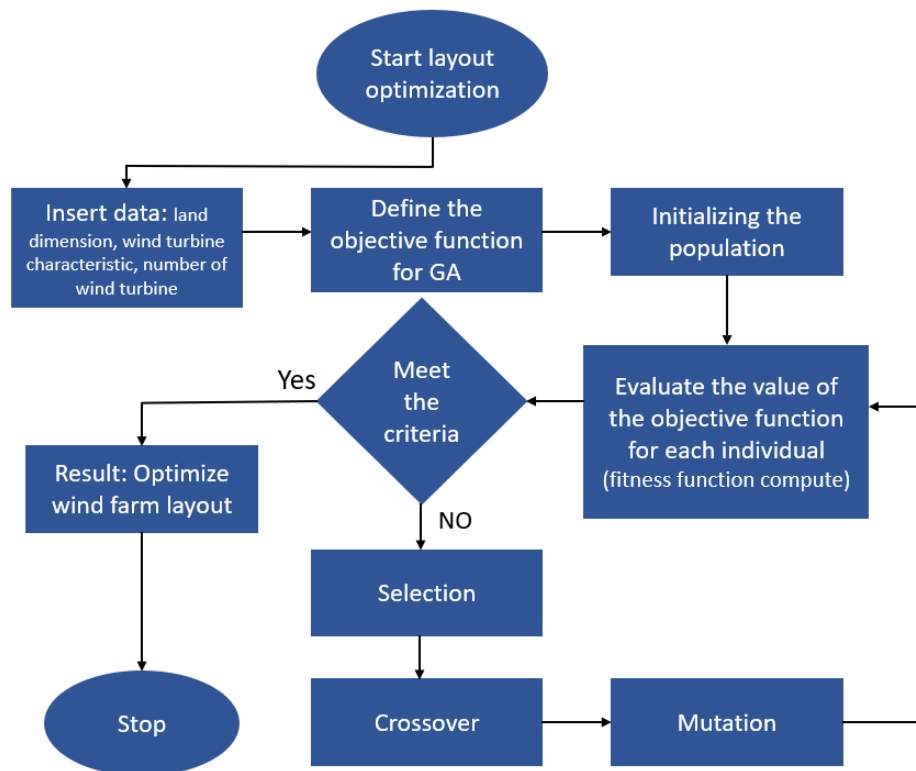


Figure 5: Flowchart of WFLO using GA

The algorithm is tested on a hypothetical 1.6 km x 1.4 km wind farm. Parameters include a rotor radius of 20 m, hub height of 80 m, axial induction factor of 0.33, and an initial wind speed of 12 m/s. Terrain roughness and wake entrainment constants are also included. The model is validated against published benchmarks and shows convergence toward optimal solutions within 55 generations.

Numerical experiments are conducted to assess the influence of four key parameters: population size, cost model, hub height, and surface roughness. These parameters are examined through four structured case studies targeting the objective function, fitness, average power, and wind turbine distribution. Population sizes of 500, 1000, 1500, and 2000 are tested to determine the optimal setting. It is found that a population size of 1500 provides the most effective trade-off between convergence speed and solution quality. Notably, when the number of generations exceeds 55, population sizes of 1500 and 2000 yield very similar fitness values, indicating that increasing the population beyond 1500 does not lead to substantial gains.

The optimized layout achieved a significant increase in energy production and a lower Cost of Energy (COE) compared to traditional layouts. The use of the GA allowed for the strategic placement of additional turbines without increasing wake losses. The study underscores the effectiveness of integrating wake modeling and evolutionary algorithms for wind farm layout optimization, offering a robust method for future real-world applications.

4 Bibliometric Analysis and Data Sources

To complement the technical modeling and simulation presented in this chapter, a bibliometric analysis was conducted to assess the trends, scope, and scholarly impact of wind farm layout optimization (WFLO) using Genetic Algorithms (GA) and Machine Learning (ML) techniques. Bibliometric analysis is a robust quantitative approach that leverages publication metadata—such as titles, keywords, authorship, affiliations, and citations—to map the intellectual structure and thematic development of a research field [39].

The software tool **VOSviewer** was employed to perform this analysis. Developed by Van Eck and Waltman [40], VOSviewer is widely used for constructing and visualizing bibliometric networks such as co-authorship, co-citation, and keyword co-occurrence maps. Prior studies have successfully applied VOSviewer to domains ranging from AI trends [41], rhetorical evolution [42], the metaverse [43], to climate-related studies in the GCC [44].

A structured query was constructed to retrieve relevant records from two major databases: Scopus and Web of Science (WoS) [45]. The query combined terms related to WFLO, optimization algorithms, and ML methods:

```
("wind farm" OR "wind turbine layout" OR "wind farm optimization" OR "wind farm planning" OR "wind turbine placement" OR WFLO OR "wind energy optimization") AND ("genetic algorithm" OR "evolutionary algorithm" OR "greedy algorithm" OR "hill climbing" OR "particle swarm optimization" OR PSO OR "simulated annealing" OR SA OR "whale optimization algorithm" OR WOA OR "ant colony optimization" OR ACO OR "artificial neural network*" OR ANN OR "deep learning" OR "support vector machine*" OR SVM OR "k-nearest neighbor*" OR KNN OR "reinforcement learning" OR RL OR "decision tree*" OR DT OR "random forest" OR "machine learning" OR "computational intelligence" OR "artificial intelligence" OR "hybrid method*" OR "hybrid optimization" OR "CFD" OR "AI" OR "XGBoost")
```

Searches were limited to 2015–2025. From WoS, a total of 241 documents were retrieved from the Web of Science (WoS) database, and 453 documents were obtained from Scopus. To enhance the quality and reliability of the dataset, a data cleaning process was undertaken. This involved removing duplicate records, eliminating irrelevant publications, and filtering out entries with incomplete metadata.

The process followed six main steps, summarized in Figure 6:

- a) Define the research focus and objective.
- b) Select databases (Scopus and WoS).
- c) Construct search queries and collect data.
- d) Clean and filter the dataset (removing duplicates and unrelated entries).
- e) Analyze the data using bibliometric techniques (keyword frequency, co-authorship).
- f) Visualize the research landscape using VOSviewer (clusters, overlays, and timelines).

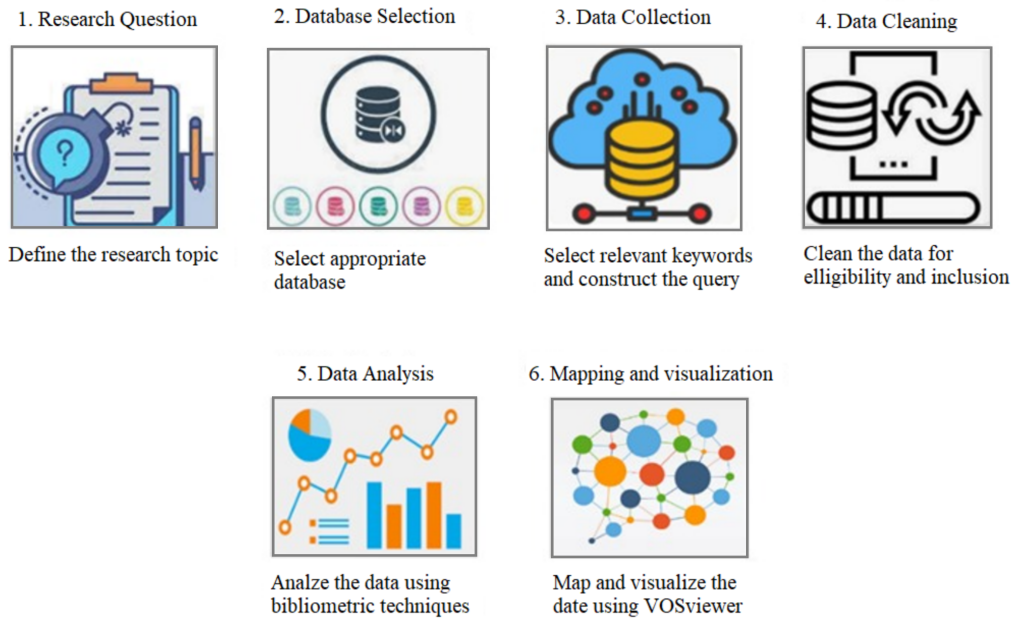


Figure 6: Main stages of the bibliometric analysis.

5 Results and Discussion

5.1 Genetic Algorithm-Based WFLO Simulation

The Genetic Algorithm (GA)-based optimization framework was applied to a hypothetical wind farm layout consisting of 60 turbines within a $1.6 \text{ km} \times 1.4 \text{ km}$ area. The GA was configured with a population size of 1500, a crossover rate of 0.5, a mutation rate of 0.02, and run for 100 generations. The results demonstrate a rapid decrease in the objective function value (cost per unit power) during the initial iterations, with convergence occurring around the 55th generation. This behavior confirms the GA's effectiveness in exploring and exploiting the search space to produce energy-efficient and cost-effective layouts.

To evaluate the adaptability of the GA under varying conditions, a parametric study was conducted focusing on four key factors: population size, hub height, surface roughness, and cost model. Each parameter was systematically tested to assess its influence on the objective function, fitness, total power output, and turbine distribution.

Results showed that increasing the population size from 500 to 1500 significantly improved optimization performance by enhancing solution diversity and reducing premature convergence [46]. However, increasing the size to 2000 did not yield significant improvements beyond those obtained with 1500, indicating that this value is optimal for the current problem scale.

To evaluate the influence of cost modeling, two cost models were compared: the simplified Mosetti model and the more detailed Chen model [20], [35]. As shown in Figure 7, the resulting layouts differ in both turbine distribution and density. Mosetti's model [20] led to more dispersed configurations, while the Chen model favored fewer, more strategically placed turbines. This contrast emphasizes how cost structure impacts optimization outcomes, with Chen's model better capturing long-term economic trade-offs.

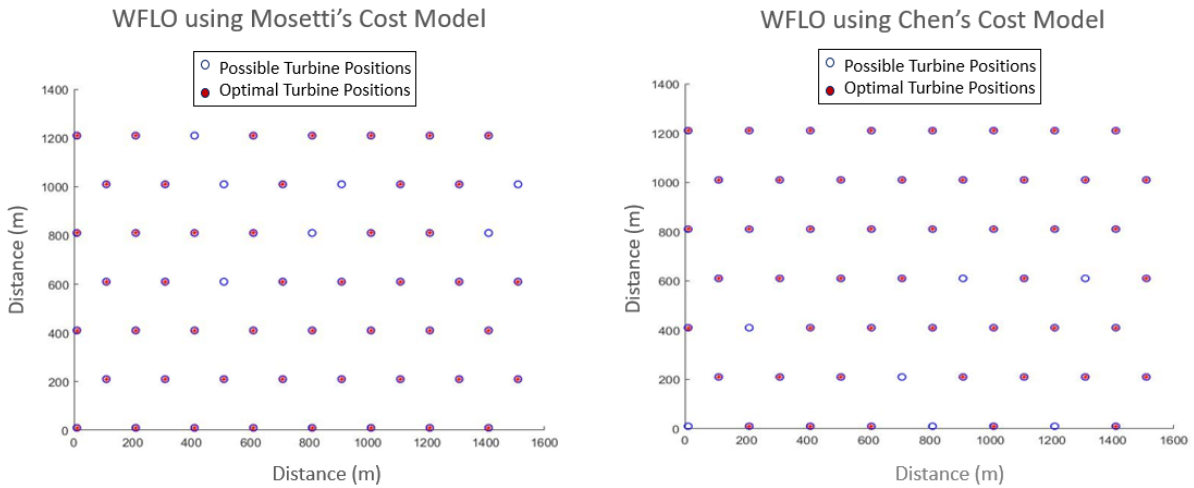


Figure 7: Optimal wind farm layout generated using Mosetti's and Chen's cost models.

In addition to cost modeling, the height of the hub was examined as a key design parameter. As shown in Figure 8, increasing the height of the hub increases the exposure of the turbine to higher wind speeds and reduces the effects of turbulence, leading to improved fitness values. Taller turbines improve energy capture and overall optimization performance, confirming industry practices that favor higher towers in flat or low-roughness sites.

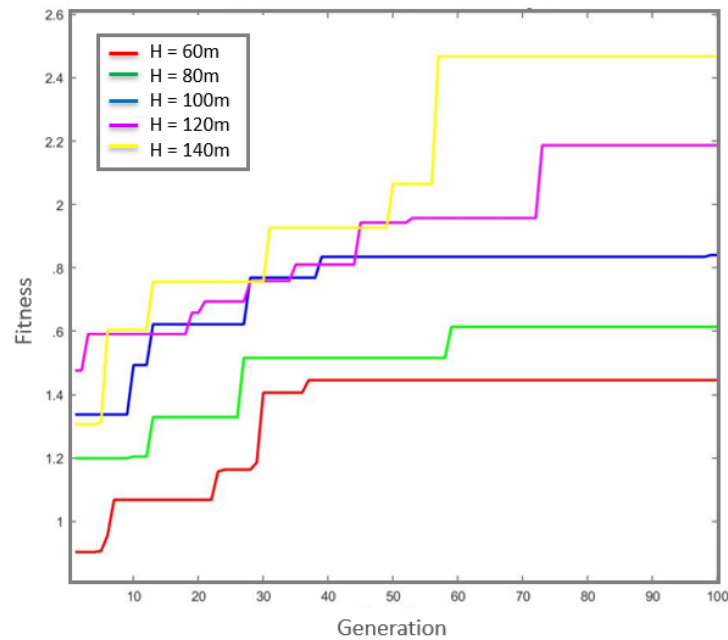


Figure 8: Fitness value across generations for different hub heights.

Surface roughness was also found to significantly impact layout performance. Figures 9 and 11 illustrate how smoother terrains (e.g., $z_0 = 0.005$) lead to faster convergence and higher fitness values, while rougher terrains (e.g., $z_0 = 1.0$) experience slower improvement and increased wake interference. Lower roughness coefficients result in more compact and efficient layouts with minimized wake losses. These findings emphasize the importance of incorporating site-specific terrain characteristics in wind farm design.

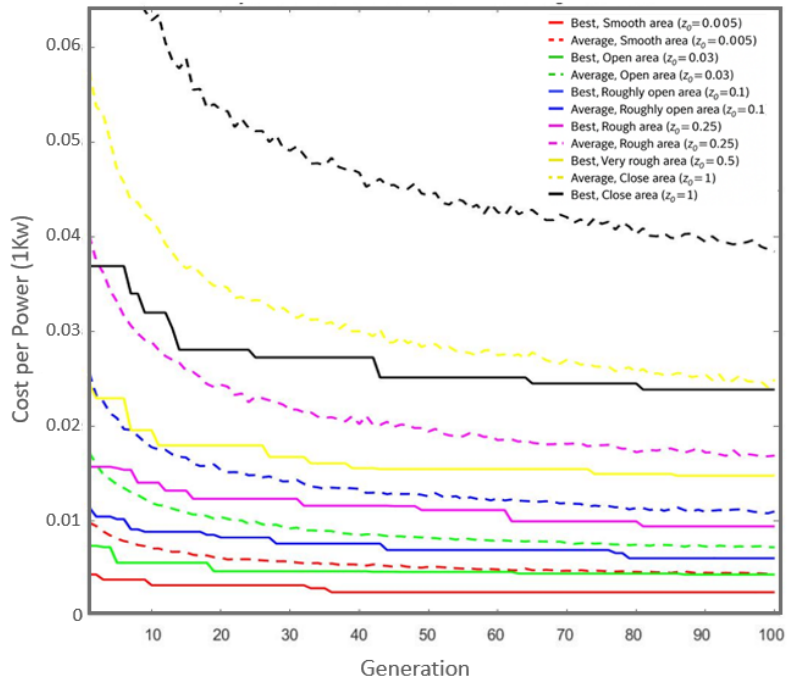


Figure 9: Objective function vs. generation for different z_0 (surface roughness) values.

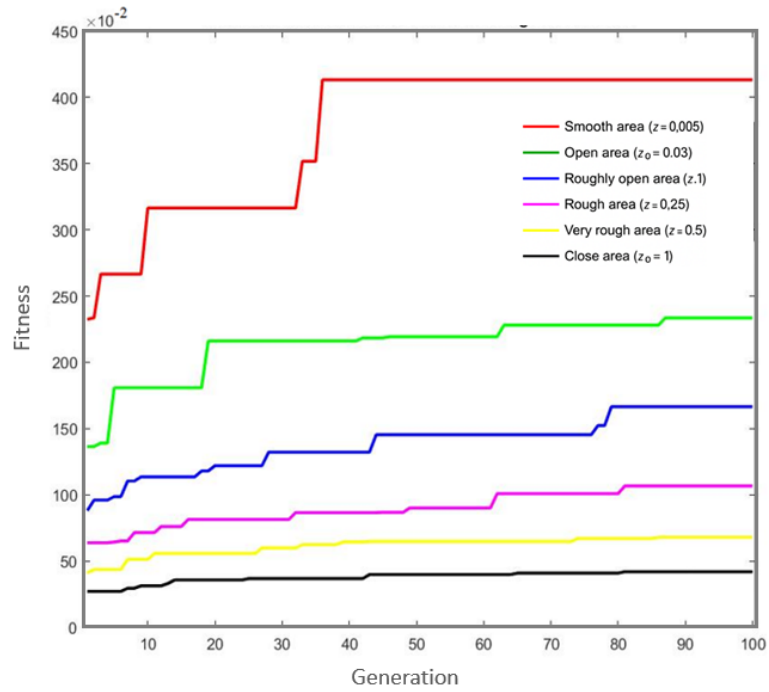


Figure 10: Fitness vs. generation for different z_0 (surface roughness) values.

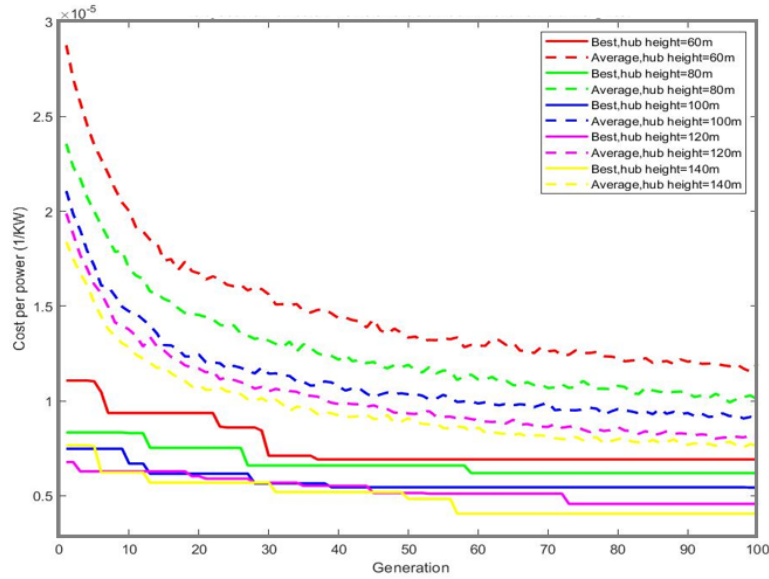


Figure 11: Fitness vs. generation for different z_0 (surface roughness) values.

Overall, the GA-based WFLO framework effectively balances energy maximization and cost minimization while adapting to different environmental and operational parameters. The observed trends validate its scalability and robustness. The results of this study also provide a solid foundation for integrating machine learning in future work to enhance wake modeling accuracy and real-time adaptability further.

5.2 Bibliometric Trends in Annual WFLO Research (2015–2025)

Building on the computational findings, a bibliometric analysis was conducted to examine the scholarly landscape of Wind Farm Layout Optimization (WFLO) research, with a particular focus on the use of Genetic Algorithms (GA) and Machine Learning (ML) techniques. The bibliographic data, collected from Scopus and Web of Science, was cleaned and analyzed using VOSviewer, as detailed in Section ???. Together, the bibliometric insights and GA-based results provide a comprehensive understanding of both the theoretical and scholarly evolution of WFLO, bridging engineering simulations with academic discourse.

In WoS, 241 documents were retrieved (2015–2025, topic field only). In Scopus, 453 documents were retrieved within the same period. The data were filtered to exclude unrelated entries or those missing author data. Figure 12 illustrates the annual publication highlighting the number of WFLO-related publications indexed in Scopus and Web of Science. The data indicates a strong and consistent upward trend in publication output, particularly in the Scopus database, which exhibits a more rapid acceleration in recent years. Between 2015 and 2020, both databases showed a steady increase, but from 2021 onward, the number of Scopus-indexed publications increased significantly, more than doubling by 2025. The findings indicate that the growing literature volume implies a diversification in methodological approaches, including the emergence of hybrid models, deep learning-based wake modeling, and real-time optimization frameworks, all of which are further explored in the thematic keyword and ML trends discussed in the next sections.

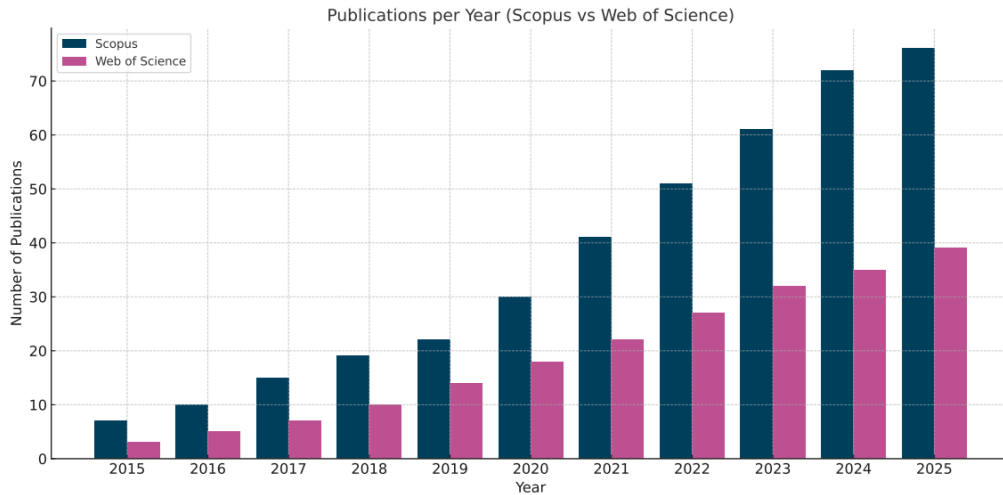


Figure 12: Annual publication trends in Scopus and WoS (2015–2025)

5.3 Categorization, Co-occurrence, Network Visualization of WFLO Evolution

To conduct a comprehensive bibliometric analysis of wind farm layout optimization (WFLO) methods, data were collected from the two major scientific databases, with an initial dataset including a total of 694 documents, 453 papers retrieved from Scopus and 241 from WoS. After merging the datasets, a cleaning process was performed to remove duplicated titles, incomplete records, and documents missing essential metadata such as title or author information. In total, 211 documents were excluded, resulting in a final dataset of 483 unique and eligible documents spanning the years 2015 to 2025. Each document was then carefully categorized based on the primary optimization or modeling technique it employed. The major identified categories included Genetic Algorithms (GA), Machine Learning (ML) techniques such as Artificial Neural Networks (ANN) and Random Forest (RF), Computational Fluid Dynamics (CFD) based modeling, hybrid optimization methods that combine GA or other heuristics with ML, and other metaheuristic algorithms such as Particle Swarm Optimization (PSO) and Simulated Annealing (SA). Minor techniques that appeared infrequently were logically grouped into broader categories to avoid unnecessary fragmentation.

Figure 13 illustrates the annual distribution of WFLO-related publications from 2015 to 2024 across several optimization techniques including Genetic Algorithms (GA), Machine Learning (ML), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Computational Fluid Dynamics (CFD)-based methods, Simulated Annealing, and a category labeled as “Other Methods”. The large number of publications marked as “Uncategorized,” especially after 2019, limits the interpretability of evolving trends; however, our query is based on the title, so any document not explicitly indicating the method in the title is not considered. Also many documents are mainly reviews.

The year 2025 was intentionally excluded from the figure due to the fact that many of its records are early-access or in-press articles that may not reflect finalized publication counts. Overall, the graph shows a steady rise in the use of GA and ML techniques in WFLO.

The heatmap of the emergence of keywords in Figure 14 provides a temporal visualization of these trends, confirming the increasing importance of ML and hybrid techniques in the WFLO literature since 2019. As illustrated in the Figure 14, publication activities in wind farm layout optimization (WFLO) using GA and Machine Learning (ML) techniques has shown substantial growth between 2015 and 2025, reflecting increased academic and industrial interest in these approaches. Notably, ML techniques have seen a pronounced uptick in usage since 2020, with deep learning, random forest, and ensemble models emerging alongside dominant tools like artificial neural networks (ANN). Furthermore, a growing number of studies are exploring hybrid optimization frameworks that combine GA with ML models such as ANN and support vector machines (SVM) to improve layout accuracy and computational efficiency. This trend signals a shift toward intelligent and adaptive wind farm systems.

Categorization of WFLO Papers by Approach (2015–2024)

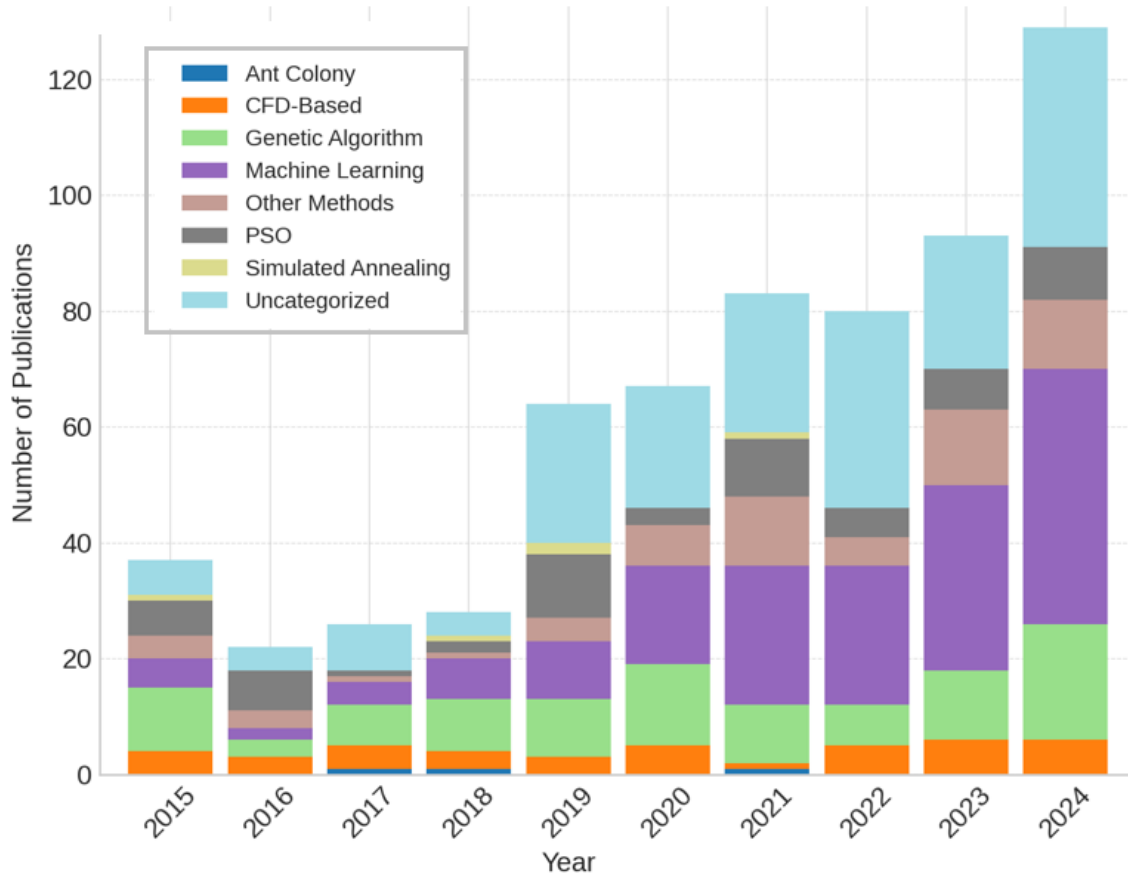


Figure 13: Categorization of WFLO-related publications across several optimization techniques.

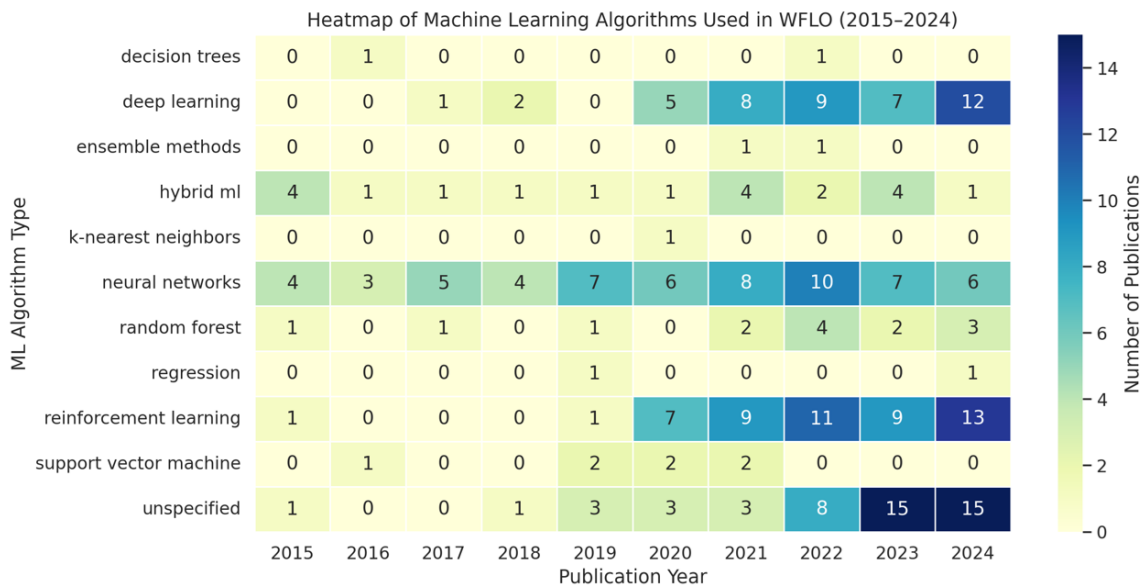


Figure 14: Heatmap of Machine Learning algorithm used in WFLO research (2015–2024).

Co-occurrence network of author keywords generated using VOSviewer, based on documents retrieved through a title-only search using combined Scopus and WoS databases is shown in Figure 15. Each node represents a unique keyword, with size indicating frequency of occurrence. Edges represent co-occurrence links, where

stronger connections indicate frequent joint appearances in the same documents. Keywords are grouped into color-coded clusters reflecting thematic associations. The central yellow cluster, with terms like *wind farm*, *wind power*, and *optimization*, highlights the core research focus on data-driven wind energy optimization. The blue cluster emphasizes aerodynamic modeling and environmental factors (e.g., *wake*, *CFD*, *micro-siting*). The red cluster centers on electrical infrastructure and grid integration (*electric power system*, *load flow*). Green nodes correspond to AI-driven forecasting methods (*deep learning*, *reinforcement learning*), while peripheral clusters touch on operational topics (*maintenance*, *decision trees*) and strategic planning (*infrastructure*, *decision making*). The map underscores the interdisciplinary nature of wind energy research, integrating physical modeling, machine learning, and system-level optimization.

Figure 16 shows the density visualization of author keyword co-occurrence, offering additional insight into the concentration of research themes within the dataset. The color gradient reflects the frequency and significance of terms: warmer colors (yellow to red) indicate high-density areas with frequent and central keywords, while cooler colors (green to blue) highlight lower-density or more specialized terms.

The most intense area, shown in red, is centered on *wind farm*, *machine learning*, and *optimization*, confirming their pivotal role in the literature. Surrounding terms such as *genetic algorithm*, *swarm optimization*, and *power generation* fall within the high-to-medium density zones, suggesting their strong relevance and interconnection with the core topics. In contrast, blue regions contain keywords like *wake loss*, *hub height*, *fault detection*, and *convolutional neural network*, which, while less frequent, indicate emerging or specialized areas of study.

This visualization reinforces the centrality of AI-based optimization approaches in wind energy research and highlights the breadth of supporting topics contributing to this interdisciplinary field.

The overlay visualization highlights emerging areas of research based on the average publication year associated with each keyword, suggesting a growing focus on aerodynamic complexities involved in wind farm layout optimization. Recent research reflects a shift toward solving challenges associated with siting turbines in complex terrains. This trend points to the need for advanced modeling techniques and decision-support tools that incorporate terrain features, wake interactions, and performance losses—particularly in the context of maximizing energy output and minimizing costs under real-world environmental constraints.

Machine learning continues to play a pivotal role in this evolution, offering scalable and adaptive approaches to optimization, prediction, and pattern recognition. Its integration with spatial modeling and aerodynamic simulations is expected to enhance the accuracy and efficiency of wind farm design and operation in increasingly challenging environments.

References

- [1] G. W. E. Council, “Global wind report,” 2024.
- [2] World Wind Energy Association, “WWEA Half-year Report 2024: Strong Wind Power Growth Continues in First Half 2024,” Tech. Rep., Dec. 2024, Accessed: 2025-04-07. [Online]. Available: <https://www.wwindea.org/HYR2024>.
- [3] O. W. in Data, *Data page: Installed capacity for different renewable technologies*, <https://ourworldindata.org/grapher/installed-global-renewable-energy-capacity-by-technology>, Data adapted from IRENA. Retrieved March 21, 2025 from <https://ourworldindata.org/grapher/installed-global-renewable-energy-capacity-by-technology>, 2025.
- [4] ResearchAndMarkets.com, “Global wind farms database 2025,” 2025. [Online]. Available: <https://www.businesswire.com/news/home/20250217267570/en/Global-Wind-Farms-Database-2025-41850-Recorded-Entries-Across-135-Countries---ResearchAndMarkets.com>.
- [5] Evwind, “Record growth in offshore wind energy in 2025,” 2025. [Online]. Available: <https://www.evwind.es/2025/03/19/record-growth-in-offshore-wind-energy-in-2025/104965>.
- [6] R. J. Barthelmie, K. S. Hansen, S. T. Frandsen, *et al.*, “Modelling and measuring flow and wind turbine wakes in large wind farms offshore,” *Journal of Atmospheric and Oceanic Technology*, vol. 26, no. 7, pp. 1662–1684, 2009. DOI: 10.1175/2009JTECHA1258.1.
- [7] G. Ekinici and H. K. Ozturk, “Forecasting wind farm production in the short, medium, and long terms using various machine learning algorithms,” *Energies*, vol. 18, no. 5, p. 1125, 2025.
- [8] Y. Zhang, J. Wang, and P. Wang, “A review of machine learning applications in wind energy,” *Applied Energy*, vol. 307, p. 118 165, 2022.
- [9] N. Bempedelis, F. Gori, A. Wynn, S. Laizet, and L. Magri, “Data-driven optimisation of wind farm layout and wake steering with large-eddy simulations,” *Wind Energy Science*, vol. 9, pp. 869–882, 2024. DOI: 10.5194/wes-9-869-2024. [Online]. Available: <https://doi.org/10.5194/wes-9-869-2024>.
- [10] J. J. Thomas, N. F. Baker, P. Malisani, *et al.*, “A comparison of eight optimization methods applied to a wind farm layout optimization problem,” *Wind Energy Science*, vol. 8, no. 5, pp. 865–891, 2023.
- [11] S. Paaso, “Mixed integer linear programming based algorithms for wind farm layout optimization,” 2022.
- [12] S. Paaso, “Mixed integer linear programming based algorithms for wind farm layout optimization,” Aalto University, 2022. [Online]. Available: <https://aaltodoc.aalto.fi/bitstream/handle/123456789/12345/thesis.pdf>.
- [13] W. Qiao, J. Liang, G. K. Venayagamoorthy, and R. Harley, “Computational intelligence for control of wind turbine generators,” University of Nebraska–Lincoln, DigitalCommons, Faculty Publications from the Department of Electrical and Computer Engineering 1216, 2011, Available at DigitalCommons@University of Nebraska–Lincoln. [Online]. Available: <https://digitalcommons.unl.edu/electricalengineeringfacpub/1216>.
- [14] N. Bhardwaj, A. Vittalaiah, A. Seshapp, *et al.*, “Optimization of wind farm layout using genetic algorithms,” in *E3S Web of Conferences*, EDP Sciences, vol. 581, 2024, p. 01 024.
- [15] T. A. Qureshi and V. Warudkar, “Wind farm layout optimization through optimal wind turbine placement using a hybrid particle swarm optimization and genetic algorithm,” *Environmental Science and Pollution Research*, vol. 30, no. 31, pp. 77 436–77 452, 2023.
- [16] P. Makhampom, J. Waewsak, C. Chancham, S. Chiwamongkhonkarn, and Y. Gagnon, “Optimal wind power plant layout using ant colony optimization:-,” *ASEAN Journal of Scientific and Technological Reports*, vol. 27, no. 1, pp. 44–57, 2024.
- [17] K. Yang and K. Cho, “Simulated annealing algorithm for wind farm layout optimization: A benchmark study,” *Energies*, vol. 12, no. 23, p. 4403, 2019.
- [18] M. El Jaadi, T. Haidi, A. Belfqih, M. Farah, and A. Dialmy, “Optimizing wind farm layout for enhanced electricity extraction using a new hybrid pso-ann method,” *Global Energy Interconnection*, vol. 7, no. 3, pp. 254–269, 2024.
- [19] G. Jones *et al.*, “Genetic and evolutionary algorithms,” *Encyclopedia of Computational Chemistry*, vol. 2, no. 1127-1136, p. 40, 1998.

- [20] G. Mosetti, C. Poloni, and B. Diviacco, "Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 51, no. 1, pp. 105–116, Jan. 1994. DOI: 10.1016/0167-6105(94)90080-9.
- [21] J. Wang, Z. Liu, H. Guo, and Z. Zuo, "Wind turbine wakes modeling and applications: Past, present, and future," *Ocean Engineering*, vol. 304, p. 115 850, 2024. DOI: 10.1016/j.oceaneng.2024.115850. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0029801824018468>.
- [22] J. Kennedy and R. Eberhart, "Particle swarm optimization," vol. 4, pp. 1942–1948, 1995.
- [23] M. Dorigo, "Analyse dun comportement collectif: Le systeme fourmi artificiel," 1992.
- [24] A. Ala, A. Mahmoudi, S. Mirjalili, V. Simic, and D. Pamucar, "Evaluating the performance of various algorithms for wind energy optimization: A hybrid decision-making model," *Expert Systems with Applications*, vol. 221, p. 119 731, 2023, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2023.119731. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417423002324>.
- [25] "A comparison of eight optimization methods applied to a wind farm layout optimization problem," *Wind Energy Science*, 2023. [Online]. Available: <https://wes.copernicus.org/articles/8/865/2023/>.
- [26] G. Dong, J. Qin, C. Wu, C. Xu, and X. Yang, "Reinforcement learning-enhanced genetic algorithm for wind farm layout optimization," *arXiv preprint arXiv:2412.06803*, 2024.
- [27] W. E. Mestari, N. Cheggaga, F. Adli, A. Benallal, and A. Ilinca, "Tuning parameters of genetic algorithms for wind farm optimization using the design of experiments method," *Sustainability*, vol. 17, no. 7, p. 3011, 2025. DOI: 10.3390/su17073011. [Online]. Available: <https://doi.org/10.3390/su17073011>.
- [28] M. E. Jaadi, T. Haidi, and A. Belfqih, "A review of wind farm layout optimization techniques for optimal placement of wind turbines," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 22, no. 3, pp. 763–772, 2024. DOI: 10.12928/TELKOMNIKA.v22i3.25609. [Online]. Available: https://www.researchgate.net/publication/371984527_A_Review_of_Wind_Farm_Layout_Optimization_Techniques_for_Optimal_Placement_of_Wind_Turbines.
- [29] J. J. Thomas, N. F. Baker, P. Malisani, *et al.*, "A comparison of eight optimization methods applied to a wind farm layout optimization problem," *Wind Energy Science*, vol. 8, pp. 865–891, 2023. DOI: 10.5194/wes-8-865-2023.
- [30] X. Chen, Y. Li, and Z. Gao, "Wind farm layout optimization using genetic algorithm and computational fluid dynamics," *Energy*, vol. 223, p. 120 041, 2021.
- [31] D. Bouabdallaoui, T. Haidi, F. Elmariami, M. Derri, and E. M. Mellouli, "Application of four machine-learning methods to predict short-horizon wind energy," *Global Energy Interconnection*, vol. 6, no. 6, pp. 726–737, 2023, ISSN: 2096-5117. DOI: 10.1016/j.gloe.2023.11.006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S209651172300097X>.
- [32] S. T. Ayele, M. B. Ageze, M. A. Zeleke, and T. A. Miliket, "Adama ii wind farm long-term power generation forecasting based on machine learning models," *Scientific African*, vol. 21, e01831, 2023.
- [33] K. Balasubramanian, S. B. Thanikanti, U. Subramaniam, S. Sichilalu, *et al.*, "A novel review on optimization techniques used in wind farm modelling," *Renewable Energy Focus*, vol. 35, pp. 84–96, 2020.
- [34] Z. Liu, H. Guo, Y. Zhang, and Z. Zuo, "A comprehensive review of wind power prediction based on machine learning: Models, applications, and challenges," *Energies*, vol. 18, no. 2, p. 350, 2025. DOI: 10.3390/en18020350. [Online]. Available: <https://www.mdpi.com/1996-1073/18/2/350>.
- [35] Y. Chen, "Commercial wind farm layout design and optimization," Ph.D. Dissertation, Texas A&M University - Kingsville, Ann Arbor, 2013, p. 47.
- [36] A. Lambora, K. Gupta, and K. Chopra, "Genetic algorithm-a literature review," in *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)*, IEEE, 2019, pp. 380–384.
- [37] J. H. Holland, "Genetic algorithms," *Scientific american*, vol. 267, no. 1, pp. 66–73, 1992.
- [38] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989, ISBN: 978-0201157673.
- [39] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *Journal of Business Research*, vol. 133, pp. 285–296, 2021. DOI: 10.1016/j.jbusres.2021.04.070. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0148296321003155>.

- [40] N. V. Eck and L. Waltman, "Software survey: Vosviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, 2010. DOI: 10.1007/s11192-009-0146-3. [Online]. Available: <https://doi.org/10.1007/s11192-009-0146-3>.
- [41] T. Brahim, "Mapping the landscape of ai research in the gcc region: A bibliometric and thematic analysis," *Cogent Engineering*, vol. 11, no. 1, p. 2428483, 2024. DOI: 10.1080/23311983.2024.2428483. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/23311983.2024.2428483>.
- [42] T. Brahim, "Similes and rhetorical structures: A bibliometric study of arabic literary research," *Journal of Arabic Studies*, 2022.
- [43] A. Sarirete and T. Brahim, "Scientific landscape of metaverse research: A bibliometric overview," *Meta-verse Journal*, 2023.
- [44] S. I. Amalou and T. Brahim, "Mapping the research landscape of organizational climate and performance using bibliometric analysis," en, *Modern Digital Analytics*, vol. 13, no. 3, pp. 7–27, 2023. [Online]. Available: https://psyjournals.ru/en/journals/mda/archive/2023_n3/Amalou_Brahim.
- [45] R. Prancute, "Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World," *Publications*, vol. 9, no. 1, p. 12, 2021. DOI: 10.3390/publications9010012. [Online]. Available: <https://ideas.repec.org/a/gam/jpubli/v9y2021i1p12-d515799.html>.
- [46] R. Asfour, T. Brahim, and M. F. El-Amin, "Wind farm layout optimization using genetic algorithm and wake modeling," in *IOP Conference Series: Earth and Environmental Science*, The 6th International Conference on Energy and Environmental Science (ICEES 2022), Kuala Lumpur, Malaysia, 6–8 January 2022, vol. 1008, IOP Publishing, 2022, p. 012004. DOI: 10.1088/1755-1315/1008/1/012004.