JESTT

Vol. 2, Issue. 3, 2025

SCOPUA

Journal of Engineering, Science and Technological Trends

Journal of Engineering, Science and Technological Trends

ISSN(e): 2959-1937

Volume 2 (Issue 3) Published on 2025-11-15

https://doi.org/10.64060/jestt.v2i3

2025

Journal Homepage

https://journals.scopua.com/index.php/JeSTT

Publisher

Scientific Collaborative Online Publishing Universal Academy

Publisher Homepage

https://scopua.com/

Model Town - Ferozepur Road, Near Atfaq Hospital, Lahore, Pakistan

Journal of Engineering, Science and Technological Trends

Aims & Scope

Journal of Engineering, Science and Technological Trends (JESTT) aims to advance the understanding of engineering sciences by providing a platform for the publication of unique contributions in the field of engineering, science, and technology.

The scope of the Journal of Engineering, Science and Technological Trends was interdisciplinary. During the board meeting on February 28, 2025, the scope of JESST was redefined to explicitly focus on "Materials Science and Engineering", "Nanoscience and Nanotechnology", "Environmental and Applied Engineering", "Computational and Data-Driven Approaches", and "Cybersecurity and Smart Systems".

The updated scope reflects our commitment to publishing high-quality, original, and innovative research that addresses contemporary challenges and advances theoretical and applied knowledge within the journal's core disciplines. JESTT welcomes manuscripts in both fundamental and applied research areas and encourages submissions that contribute novel and innovative insights to engineering science and technology. The scope of JESTT includes:

1. **Materials Science and Engineering**

- Additive Manufacturing (3D Printing of Materials)
- Biomaterials Engineering
- Inorganic and Organic Semiconductors
- Material Characterization
- Material Science
- Material Statistics (Use of Modern Statistics like Neutrosophic Analysis for Decision-Making in Materials Field)
- Materials Engineering
- Materials for Flexible and Wearable Electronics

2. **Nanoscience and Nanotechnology**

- Nanofabrication
- Nanomaterial-Based Sensors
- Nanotechnology

3. **Environmental and Applied Engineering**

- Heavy Metal Detection
- Hydraulic Engineering
- Sensor Technology

4. **Computational and Data-Driven Approaches**

- Computational Materials Science

- Computer Vision Applications
- Deep Learning in Pattern Recognition
- Digital Twin Technology in Engineering

5. **Cybersecurity and Smart Systems**

- Artificial Intelligence in Cybersecurity
- Smart Grid and Critical Infrastructure Protection

JESTT Policy on the Use of AI

This policy applies to all manuscripts submitted to the Journal of Engineering, Science and Technological Trends (JESTT), ISSN(e): 2959-1937. It aims to ensure transparency, research integrity, and responsible use of generative AI and AI-assisted tools in the preparation of scholarly work. Generative AI and AI-assisted tools cannot be listed as authors or co-authors. All listed authors must ensure that the work is original, lawfully prepared, and does not infringe third-party rights. Generative AI or AI-assisted tools must not be used to create or modify figures, images, or artwork in submitted manuscripts.

Authors may use generative AI and AI-assisted tools to support tasks such as:

- Literature structuring and summarization
- Improving language clarity and readability
- Organizing content

However

- AI tools must not replace human critical thinking or expertise.
- Authors remain fully responsible and accountable for the accuracy, completeness, impartiality, and originality of the manuscript.

Authors must

- Check AI outputs for factual errors, bias, and fabricated references.
- Thoroughly edit and adapt any AI-generated text so it reflects their own analysis and ideas.
- Review the terms and conditions of the tool to ensure that:
    - The privacy and confidentiality of data (including unpublished manuscripts and any personal data) are protected.
    - Only limited rights are granted to the tool (for service provision), and not broader rights such as using the content to train models.
    - Tool terms do not restrict future publication of the work in JESTT.
- Avoid uploading content that may infringe their own, third-party, or JESTT's intellectual property rights.

Authors must disclose any significant use of generative AI or AI-assisted tools in writing the manuscript. Use of AI tools in data analysis or research methods should be described in the Methods section (or sub-section), with sufficient detail for reproducibility. No disclosure is needed for standard spelling and grammar checkers or reference managers used only for organizing and formatting references. A "***Declaration of AI Use***" must be added at the end of the manuscript; a sample is given below

The authors declare that the AI tool [name and version of tools] was used [for language editing and readability improvement / to assist with drafting and restructuring sections, summarizing *background literature, and improving the fluency of the text / both in the writing and methodological stages / preparation, writing, or analysis] in the preparation of this manuscript submitted to the Journal of Engineering, Science and Technological Trends (JESTT), ISSN: 2959-1937. The tool did not generate original research ideas, research data, analyses, or conclusions. All content was critically reviewed, verified, and edited by the authors, who take full responsibility for the accuracy and integrity of the work.*

Editorial Board Members

Contents (Volume 2 & Issue 3 – 2025)

Journal of Engineering, Science and Technological Trends (JESTT)

Vol. 2, Issue 3, November 2025
DOI: https://doi.org/10.64060/jestt.v2i3.1

Review

Density Functional Theory: A Quantum Mechanical Framework for Novel Materials Design

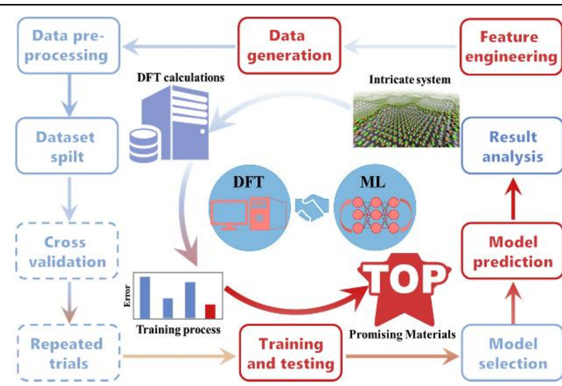
Mubashar Ali[1]*

[1] School of Materials Science and Engineering, Beijing Institute of Technology, Beijing, China.
* Corresponding Email: mubasharali.edu@outlook.com (M. Ali)

Received: 31 August 2025 / Revised: 08 October 2025 / Accepted: 19 October 2025 / Published online: 15 November 2025

ABSTRACT

Density Functional Theory (DFT) has become a fundamental principle of contemporary materials research, providing a quantum mechanical framework for the examination of matter at the electronic level. By changing the many-body problem into electron density, DFT makes it possible to make precise predictions of structural, electronic, and catalytic properties based on basic principles. Because it can make predictions, it has sped up the discovery of semiconductors, catalysts, and energy storage materials, which means we don't have to rely on expensive experiments as much. At the same time, projects like the Materials Project show how important it is for high-throughput computational design. Even if there are problems with the cost of processing and the accuracy of the results, new developments like hybrid methods, machine learning integration, and new quantum computing technologies keep making it more useful. So, DFT is not only a basic theoretical tool, but it is also a real driver of innovation in the creation of new materials.



Keywords: DFT; Kohn-Sham equation; Application of DFT; DFT Methods; Exchange-Correlation function; GGA and LDA; Hybrid Functional

1. Introduction

The discovery and development of new materials have powered technological revolutions throughout human history. Each period of civilisation has been marked by the materials that it controlled, from stone and bronze to steel, plastics, and semiconductors [1]. Today, the development of functional materials is crucial for new revolutions in renewable energy, microelectronics, biotechnology, and quantum information systems [2]. As society is fighting to combat serious international issues like climate change, energy security, and sustainable development, the demand for materials possessing specific properties has increased exponentially. Traditionally, scientists discovered materials by trial and error, intuition, or serendipity through experimental approaches [3]. While this method was good enough in the past, it is progressively inadequate for dealing with the complexity of modern technology. Experimental search is slow, expensive, and limited, especially in light of the enormous range of possible material composition, structure, and chemical environment. In response, materials science is undergoing a fundamental change, which involves the incorporation of computational modelling and quantum mechanical tools in the process of discovery and design [4],[5],[6].

At the helm of this computational era is Density Functional Theory (DFT), a quantum mechanical theory that allows the exploration of materials at the electronic level. In contrast to conventional theories of macroscopic approximation, DFT offers insight into the governing interactions behind material behaviour, e.g., bonding, magnetism, and charge transport. This makes DFT a powerful material property prediction, experimental design, and innovation-acceleration tool [6]. The importance of DFT is not limited to the scientific community. It is the computational foundation of the "Materials Genome Initiative," an international effort to accelerate materials discovery and development by applying computational screening and high-throughput design. By reducing the need for expensive and time-consuming experiments, DFT is now essential to industries such as semiconductors, catalysis, and energy storage. The ability to predict the electronic structure and material properties before synthesis in the lab is a revolutionary step away from the conventional way science and engineering go about discovering materials [7],[8].

The article is concerned with the use of DFT as a quantum mechanical approach to designing new materials. The article begins by presenting the theoretical framework and development history of DFT before proceeding to examine its use in materials science. It then goes on to discuss applications of DFT in the prediction and design of new materials, as well as challenges and limitations. It then looks at where research with DFT is headed in the future, including its combination with machine learning and hybrid computational strategies. Through this journey, it is apparent that DFT not only enhances our knowledge of materials at the atomic scale but also helps in forming technologies of the future.

2. Background on Density Functional Theory (DFT)

Density Functional Theory finds its origin in the broader branch of quantum mechanics, which emerged in the early twentieth century to explain matter's behaviour at atomic and subatomic levels [9],[10]. Quantum mechanics' greatest challenge is to solve the Schrödinger equation for many-electron systems. Although there are exact solutions for small atoms such as hydrogen, the intricacy of real materials with more than one interacting electron makes direct solutions impossible. In the middle of the twentieth century, profound theoretical breakthroughs made DFT applicable in practice. In 1964, Pierre Hohenberg and Walter Kohn proved that a many-electron system's ground-state behaviour depends only on its electron density, not upon the many-body wavefunction [11]. This insight, formalised in the Hohenberg–Kohn theorems, had greatly reduced the computational intensity of quantum mechanical calculations because electron density depends on only three spatial coordinates, compared to numerous coordinates needed for a complete wavefunction. Utilising the ground state density, we can precisely compute every observable, including the energy of a stationary quantum mechanical system. The observable characteristics of the quantum mechanical categorisation of stationary states can be elucidated as a function of density in the ground state. The ground density can be computed using the variation technique that solely incorporates density. The Hohenberg-Kohn theorem pertains to time-independent ground states; however, it has recently been extended to encompass time-dependent potentials and excited states [12].

By expanding on this idea, Kohn and Sham in 1965 introduced the Kohn–Sham equations, rendering DFT a matter of computational feasibility. Their approach replaced the complex many-body problem with an idealised system of non-interacting electrons with the same electron density as the real system. The electron–electron interaction problem was embodied in the exchange-correlation functional, still one of the most crucial DFT approximations. Choice and quality of this function are critical to the precision of DFT calculations. Years of uninterrupted advancement witnessed the emergence of improved exchange-correlation functionals like the Local Density Approximation (LDA) and the Generalised Gradient Approximation (GGA), and much more refined hybrid functionals that combine the virtues of DFT with the strengths of Hartree–Fock theory. Alongside the accelerating growth in computing facilities, DFT was quickly no longer a theoretical entity but had evolved into an everyday tool in physics, chemistry, and materials science. DFT now forms the basis of much computational materials science. Its flexibility covers a range of operations, from the calculation of crystal structure stability to catalysis on surfaces and interfaces. Crucially, DFT bridges basic quantum mechanics to material design, a level of prediction that is both rigorous and computationally tractable. The development of DFT provides an example of how profound theoretical understanding can lead to revolutionary technological instruments that transform whole fields.

3. Importance of DFT in Materials Science

The use of DFT in the field of materials science has significantly changed the research and innovation landscape. Its importance is based on its ability to link microscopic quantum mechanical descriptions to macroscopic material properties [13]. Unlike empirical models that depend on the fitting of experimental data, DFT gives first-principles calculations capable of predicting the material's behaviour without adjustable parameters. This predictive capability enables researchers to study new materials before they are synthesised, such that the most promising candidates can be identified and fewer costly trial-and-error experiments are required [14]. One of DFT's most important contributions lies in the prediction of electronic properties. Designing semiconductor, superconducting, and photovoltaic materials relies heavily on knowledge of electronic band structure and band gaps [15],[16],[17]. DFT offers a method of calculating these properties precisely and thereby informs experimental synthesis. For example, for renewable energy, DFT has played a key role in finding novel perovskite materials with tunable band gaps for high-efficiency solar cells. DFT also plays an important role in catalysis. The functioning of a catalyst depends on its electronic structure and the adsorption energies of reactants at active sites. DFT allows scientists to simulate reaction paths, find transition states, and predict catalytic efficiency. This has played a crucial role in designing catalysts for applications ranging from fuel cells to carbon dioxide reduction. The Nobel Prize-winning work of the surface scientists in heterogeneous catalysis has predominantly depended upon DFT studies, reflecting its usefulness [18]. In energy storage and conversion, DFT facilitates the exploration of new battery, supercapacitor, and hydrogen storage materials [19],[20],[21],[22]. In lithium-ion batteries, DFT is employed to assess ion diffusion barriers, voltage profiles, and electrode material stability. For example, Xiaojun Zhao et al. studied the various MXenes as anode material. They used only the DFT simulations to predict all the important characteristics. This ability to predict has hastened the search for next-generation solid electrolytes and high-capacity electrode materials, which are key to sustainable energy technology.

Besides, the relevance of DFT to structural and mechanical properties is also important [23],[24]. Calculations of phonon spectra, elastic constants, and defect energetics are used to determine how materials respond under stress, temperature, and impurities [25],[26],[27],[28],[29]. Such studies are important for the development of high-strength, long-lasting, and thermal-stability materials for use in aerospace, automobiles, and construction industries. Perhaps most importantly, DFT has enabled high-throughput materials discovery, where researchers are able to computationally screen vast libraries of theoretical compounds for desired properties. Initiatives like the Materials Project and Open Quantum Materials Database (OQMD) rely heavily on DFT calculations to construct predictive databases, enabling researchers worldwide to find candidates of interest for experimental investigation. Essentially, the contribution of DFT to materials science transcends its explanatory value and is still more significant as a predictive and guiding tool. By linking quantum mechanics with material applications in the real world, DFT has become a cornerstone of modern science, accelerating breakthroughs and formulating emerging technologies.

4. Theoretical Framework of DFT

The strength of Density Functional Theory (DFT) is its solid basis on the principles of quantum mechanics. Whereas standard quantum methods attempt to solve the many-body wavefunction directly, DFT reformulates the problem in terms of electron density. The problem, so simplified, is computationally manageable while maintaining essential physics. To understand this construct, it is necessary to start with the basic principles of quantum mechanics before venturing into the technical theory of DFT.

4.1. Entanglement

Quantum mechanics, developed in the early twentieth century, is the basic theoretical framework employed to describe matter behaviour at the atomic and subatomic levels. The Schrödinger equation forms the core of quantum mechanics and dictates the evolution of a system's wavefunction. For a single particle, the time-independent Schrödinger equation can be written as in Equation 1 [30].

$$H\psi(r) = E\psi(r) \qquad (1)$$

Here, **H** is the Hamiltonian operator, $\boldsymbol{\Psi(r)}$ is the wavefunction, and E is its corresponding energy eigenvalue. All information concerning a quantum system, including the spatial probability distribution of particles, is housed in the wavefunction. Applied to many-electron systems, i.e., atoms, molecules, or solids, the Schrödinger equation is extremely complex as in Equation 2. Each electron feels not only the positively charged nuclei but also any other electron due to Coulombic repulsion as represented in Equation 3. This leads to the many-body problem, where complexity increases exponentially with particle number. Direct solutions are only feasible for highly idealised systems, e.g., hydrogen or helium atoms

$$\left[\left[\sum_{i}^{N}\left(-\frac{\hbar^2}{2m}\Delta^2+V(r_i)\right)\right]+\left(\sum_{i<j}^{N}U(r_i),(r_j)\right)\right]\Psi_{(r_1,r_2,\dots\dots,r_N)} = E\Psi\left(r_{1,r_2,\dots\dots,r_N}\right)$$

(2)

$$\hat{U} = \Sigma_{i<j}U(i,j) = \Sigma i,j\,\frac{q^2}{|r_{i,r_j}|} \qquad (3)$$

$$\hat{V} = \Sigma_i V(i) = -\Sigma\frac{Z_k}{|r_{i-}R|} \qquad (4)$$

Equation 4 represents electron–nucleus attraction (external) potential in a many-electron Hamiltonian. To overcome this, various approximations and reformulations have been proposed. Methods like Hartree–Fock theory provided initial frameworks for approximating many-electron systems but generally failed to provide an accurate description of electron correlation effects. DFT has emerged as a different methodology that avoids the explicit many-body wavefunction dependence in favour of thinking about a simpler yet equally complete quantity: the electron density.

4.2. Theoretical Foundations of DFT

Theoretical breakthrough of DFT is founded on Hohenberg–Kohn theorems (1964), which established a solid ground for the use of electron density as the central variable [31]. The first theorem states that the ground-state properties of a many-electron system are uniquely determined by its electron density. This means there is a one-to-one correspondence between the external potential felt by the electrons (and thus the entire Hamiltonian) and electron density. In brief, knowledge of electron density is sufficient to explain all physical properties of a system. The second theorem says that the ground-state energy can be obtained by a variational principle in the electron density. Specifically, the exact ground-state density is obtained when the total energy functional is minimised. These consist of the kinetic energy functional, **T[ρ]**, electron–nucleus interaction energy, **Vne[ρ]**, classical electron–electron repulsion, **Vee[ρ]**, and the exchange-correlation functional, which accounts for all the remaining quantum mechanical effects, including the exchange and correlation interactions. In 1965, Kohn and Sham discovered an implementable scheme to use these ideas by developing a system of auxiliary non-interacting electrons whose density is identical with that of the true interacting system. They then derived the Kohn–Sham equations, a set of self-consistent equations which are solved iteratively until convergence as shown in Figure 1. The computational efficiency of this scheme and developments in exchange-correlation functionals made DFT a pillar in computational material science. The Kohn-Sham equation is formulated as an eigenvalue equation analogous to the Schrödinger equation as in Equation 5:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{eff}(r)\right)\varphi_i(r) = \varepsilon_i\varphi_i(r) \qquad (5)$$

The term $\boldsymbol{v_{eff}(r)}$ in equation (5) is Kohn–Sham potential, which can be represented as in Equation 6:

$$v_{eff}(r) = v_{ext}(r) + e^2\int\frac{\rho(r')}{|r-r'|}dr' + v_{XC}(r) \qquad (6)$$

$\boldsymbol{v_{ext}(r)}$ and $\boldsymbol{v_{XC}(r)}$ are the external and exchange-correlation potentials for a many-body system.

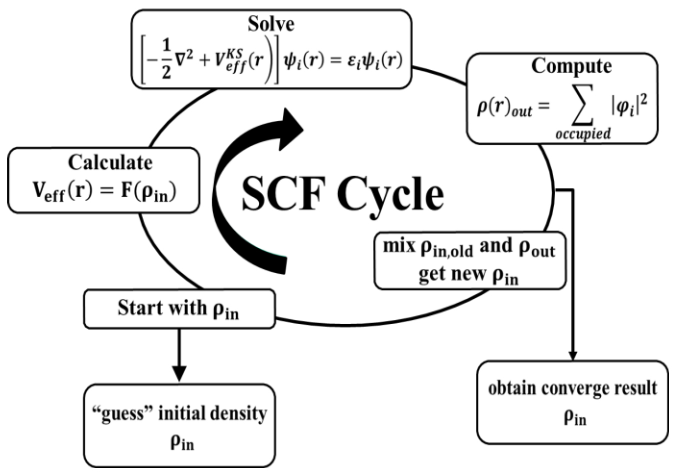


Figure 1: *Schematic representation of self-consistent solution of the Kohn-Sham equations and optimised system configuration.*

Over the years, numerous approximations to the exchange-correlation functional have been developed. This advancement is a testament to DFT's flexibility in balancing accuracy and cost. By expressing quantum mechanics in terms of electron density, DFT brings together theoretical underpinnings and application and is thus essential to material design.

5. Theoretical Framework of DFT

The utility of DFT in practice is best demonstrated through its myriad applications in the development of materials [32],[33],[34]. DFT not only explains experimental data by predicting chemical, electronic, and physical properties (See Figure 2), but it also facilitates the design of novel materials with specified functionalities. Two of the most important applications include material property prediction and the design of novel materials.

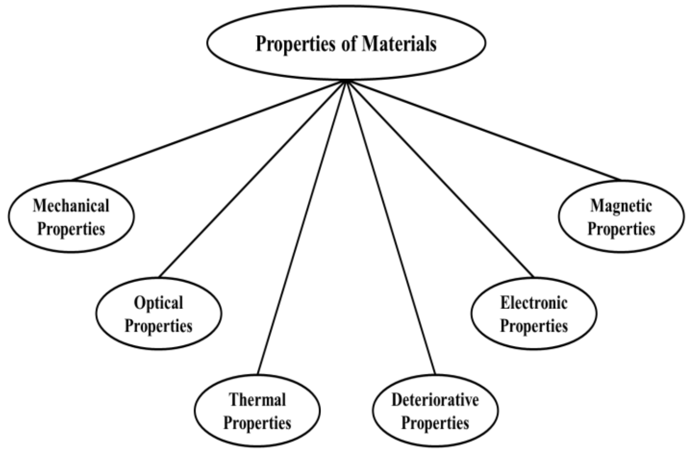


Figure 2: *A range of properties predicted with DFT simulations*.

5.1. Theoretical Foundations of DFT

One of the principal applications of DFT in the near term is that it is capable of predicting intrinsic material properties from first principles [35],[36]. These include a broad array of useful electronic properties relevant to condensed matter physics, chemistry, and engineering in general: DFT enables one to compute band structures, density of states, and effective masses of charge carriers. They are of extreme significance in the context of photovoltaics, semiconductors, and optoelectronic devices. DFT calculations, for example, have been used to predict perovskite solar cell material band gaps and the possibility of their tuning, along with their efficiency. Solid-state electronic spin structure of ferromagnets, antiferromagnets, and spintronic materials can be explored by spin-polarised DFT. This has been particularly helpful in magnetic storage devices and spin-based quantum computing devices. Furthermore, the elastic behaviour, phonon dispersions, and defect energies can be determined in order to comprehend material stability and the way in which it responds to external stress. To illustrate, DFT has been used to investigate the strength and flexibility of two-dimensional materials like graphene and transition-metal dichalcogenides. Phase diagrams, formation energies of defects, and defect chemistry can be predicted with the use of DFT. It has proved useful in materials screening for high-temperature or corrosive environments, e.g., applications in nuclear reactors. In energy storage, DFT has been used to predict the barriers to diffusion of hydrogen, sodium, and lithium ions in solid electrolytes. These predictions are useful towards the goal of creating high-performance future fuel cells and batteries. With such predictive capabilities, DFT creates a close connection between microscopic interactions and macroscopic material behaviour. It allows researchers to interpret experimental observations and even predict material performance under circumstances that could be very challenging, if not impossible, to reproduce in the laboratory.

5.2. Design of New Materials

Besides prediction, DFT is a design platform for synthesising materials with known functionalities. Supported by computational screening and high-throughput pipelines, researchers are able to scan broad compositional and structural spaces, excluding candidates of interest before synthesis. DFT has been the primary tool in discovering novel semiconductor alloys and 2D materials with engineered band gaps for applications in solar cells, light-emitting diodes, and photodetectors. Calculations of electronic and optical properties of layer materials like MoS_2 and WS_2 by DFT motivated the discovery of such materials. DFT has been applied to heterogeneous catalysis for active site identification, reaction intermediate prediction, and the proposal of novel catalytic mechanisms. DFT computations, for instance, have influenced the development of CO_2 reduction and hydrogen evolution reaction catalysts, which are both critical to renewable energy technologies. DFT calculations facilitated the optimisation of electrolytes and electrodes for lithium-ion as well as for future batteries. Redox potential, barrier to diffusion, and stability can be computed and utilised for designing materials that have greater energy density and safety. DFT has helped in the quest for novel families of superconducting materials, such as iron-based superconductors and high-pressure hydride superconductors. While it is challenging to predict critical temperatures, DFT helps in narrowing down possible candidates. The discovery of graphene spurred the research into other 2D materials, most of which were first researched using DFT. Theoretical band structure predictions, topological properties, and excitonic properties paved the way for experimental verification.

With the inclusion of DFT in the design, researchers can move away from serendipitous to rational design and accelerate innovation at a lower cost. DFT-based design has already achieved improvements in photovoltaics, catalysis, and energy storage and is entering new fields such as quantum computing and nanomedicine.

6. Challenges and Limitations of DFT

While Density Functional Theory has emerged as one of the most universal and widely applicable techniques in materials science, it is not without fault [37]. Its success is undeniable, yet as with any theory, it has limitations caused by approximations, computational requirements, and methodological concerns. It is of value to recognise these limitations both to interpret DFT results and to guide ongoing efforts to make it more accurate and useful.

6.1. Computational Limitations

One of the primary challenges of DFT is that it is computationally costly, particularly for large and complex systems. Although DFT is considerably less computationally costly than wavefunction-based methods like Coupled Cluster (CC) theory or Configuration Interaction (CI), it still scales approximately with the cube of the number of electrons in a system [38],[39]. This renders it computationally intensive for systems that consist of thousands of atoms, such as polymers, biomolecules, or realistic material interfaces. Besides, simulations of big systems like crystal defects, surfaces, and interfaces normally involve very big simulation cells and reciprocal space dense sampling. Such demands can result in computational requirements beyond the capability of typical desktop or mid-scale cluster capacity. Even on supercomputers, compromises have to be made between accuracy, computational feasibility, and simulation size. An additional computational challenge arises if time-dependent properties are needed. Standard DFT is a ground-state theory, and while there are extensions such as Time-Dependent DFT (TD-DFT), these are considerably more computationally intensive. Similarly, the description of excited states, strong correlations, and non-equilibrium processes remains challenging, especially for systems with strongly localised d or f electrons, such as transition metal oxides or rare earth compounds. Accordingly, although DFT is a very effective predictive method, researchers will always be aware of its associated computational problems and choose wisely the order of approximation and system size they can realistically compute.

6.2. Accuracy and Reliability Problems

Probably the most debated limitation of DFT is the accuracy of exchange-correlation (XC) functionals, the base of the theory. The exact function is not known, and all practical DFT calculations are based on approximations. The LDA works surprisingly well in small metal systems and gives good structural properties; it yields too narrow band gaps and overbinds molecules. However, the functionals like GGA with PBE improve the prediction of structures and energetics, but still underestimates electronic band gaps, in certain cases up to 50%. Incorporating a percentage of Hartree–Fock exchange improves band gap prediction and reaction energetics at a much higher computational expense. Furthermore, the materials containing localised electrons (e.g., transition metal oxides, Mott insulators), standard DFT completely breaks down, predicting metallic instead of insulating behaviours. DFT+U or Dynamical Mean Field Theory (DMFT) are some of the tools required to correct such breakdowns. Yet another problem is the absence of universal correctness for various property predictions. A good functional for a structural property can be a failure for electronic excitations and vice versa. Such a lack of consistency complicates the use of one functional for all tasks, and scientists must choose and validate the functional suitable for their system of interest. Extremely important, but not to account for van der Waals interactions, which reign supreme in layered compounds, molecular crystals, and biological compounds. To correct for this, special vdW corrected functionals or dispersion corrections must be used, which add complexity with a degree of uncertainty. Collectively, these lay out that whilst DFT is a very useful tool, it should never be used as a black box. Results should be subjected to critical scrutiny, referenced against experimental data whenever possible, and supplemented by higher-level techniques where necessary.

7. Future Directions in DFT Research

Notwithstanding its limitations, the future holds much promise for DFT studies with ongoing progress in algorithms in computation, function development, and hybridisation with emerging technologies. These will increase DFT scope, accuracy, and lower the computational intensity, making it an even more vital tool for discovering new materials. Time-dependent density functional theory (TDDFT) extends ground-state DFT to study the time evolution of electron density in many-electron systems. Based on the Runge-Gross theorem, TDDFT maps the complex interacting system to a simpler non-interacting Kohn-Sham system with the same time-dependent density. It is widely used to calculate electronic excitations and optical properties through linear response methods. The exact form of the time-dependent Kohn-Sham potential is unknown and typically approximated. TDDFT offers a computationally efficient way to analyse dynamic quantum phenomena beyond static ground states.

7.1. Integration with Machine Learning

It is also the leading edge of advanced studies in the interfacing of machine learning (ML) and DFT through training ML algorithms on massive collections of DFT calculations. Surrogate models predicting material properties at a fraction of a percentage of the computational effort can be built [40],[41]. This opens the way:

- **High-Throughput Screening:** With ML-enhanced predictions, virtual millions of compounds are screened for worthwhile properties, a number many orders beyond the capability of conventional DFT cycles.
- **Improved Functionals:** Machine learning is utilised to create data-driven exchange-correlation functionals that better represent electron correlation effects than current approximations.
- **Multiscale Modelling:** ML models can bridge the gap between atomistic DFT simulations and continuum-level material models, enabling predictions of material performance under realistic device and environmental conditions.

All these endeavours, like the Materials Project, AFLOWLIB, and Open Quantum Materials Database (OQMD) are already employing ML together with DFT datasets to accelerate materials discovery, from new catalysts to superconductors. The hybrid combination of DFT and ML is a paradigm shift toward data-driven materials design.

8. New Horizons: Quantum Computing and Beyond

Looking to the future, quantum computing can revolutionise computational materials science, starting with DFT [42]. Quantum algorithms can perhaps, in the future, directly address many-body problems in a manner that the need for approximate functionals is significantly reduced. While this is a vision still in its early stages, work on quantum algorithms for electronic structure has already started, looking forward to a time when quantum-enhanced DFT is possible. Yet another emerging trend is incorporating time-dependent and non-equilibrium effects. Extensions like Time-Dependent DFT (TD-DFT) already permit calculations of excited states and optical properties, but more robust models need to be formulated to compute ultrafast dynamics, electron transport, and photo-induced effects. Finally, the synergy between experimental input and computation is gaining speed. Automated labs, coupled with DFT and ML predictions, can be the closed-loop systems for autonomous materials discovery as illustrated in Figure a. DFT adds the quantum mechanical accuracy needed to ensure that there will be credible predictions, while experiments validate as well as correct computational results.

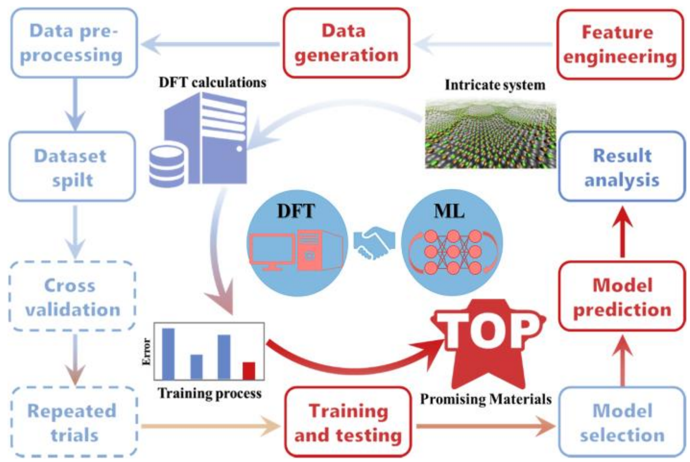


Figure 3: *A schematic illustration of the DFT-based and ML-accelerated method for the discovery of promising materials in intricate systems. The blue-dotted box means the procedure is optional.*

9. Conclusion

Our results demonstrate that neutrosophic logic provides a fundamentally new language for describing quantum systems, one that directly encodes ontological indeterminacy rather than reducing it to ignorance or vagueness. By applying neutrosophy to superposition, entanglement, and wavefunction collapse, we establish a framework that complements Hilbert space formalism while extending its interpretive power. This perspective has implications not only for the foundations of quantum theory but also for the design of algorithms, error correction, and decision-making in quantum information science. Future integration of neutrosophic models with experimental platforms such as weak measurements and quantum tomography could yield testable predictions that distinguish neutrosophic representations from classical probabilistic approaches. More broadly, embracing indeterminacy as an explicit dimension may reshape how uncertainty is modelled across physics, mathematics, and computation, offering a step toward a more complete understanding of quantum reality.

Declaration

Competing Interests: The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Issues: There are no ethical issues. All data in this paper is publicly available.

Author Contribution Statement: M.A. conceived idea and designed the research; Analyzed interpreted the data and wrote the paper.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Consent to Publish: The author is agreed to publish version of the manuscript in this journal.

References

1. Conway, E., Material world: The six raw materials that shape modern civilization. 2023: Vintage.
2. Maity, S., Materials Science and. Innovations in Energy Efficient Construction Through Sustainable Materials, 2024. 175.
3. Hallsworth, J.E., et al., Scientific novelty beyond the experiment. Microbial Biotechnology, 2023. 16(6): p. 1131-1173.
4. Guarnieri, M., Innovative materials, computational methods and their disruptive effects, in Optimal Design Exploiting 3D Printing and Metamaterials. 2022, Institution of Engineering and Technology. p. 5-55.
5. Singh, V., et al., Recent trends in computational tools and data-driven modeling for advanced materials. Materials Advances, 2022. 3(10): p. 4069-4087.
6. Shahzad, K., A.I. Mardare, and A.W. Hassel, Accelerating materials discovery: combinatorial synthesis, high-throughput characterization, and computational advances. Science and Technology of Advanced Materials: Methods, 2024. 4(1): p. 2292486.
7. Shang, Y., et al., Materials genome engineering accelerates the research and development of organic and perovskite photovoltaics. Materials Genome Engineering Advances, 2024. 2(1): p. e28.
8. Nwabara, U., et al., High throughput computational and experimental methods for accelerated electrochemical materials discovery. Journal of Materials Chemistry A, 2025.
9. Saluja, B., Quantum Mechanics: Fundamental Theories. 2025: Educohack Press.
10. Giliberti, M. and L. Lovisetti, Old Quantum Theory and Early Quantum Mechanics. 2024: Springer.
11. Masrour, R., Computational Methods: Ab Initio Calculations and Monte Carlo Simulations, in Electronic, Magnetic, and Thermoelectric Properties of Spinel Ferrite Systems: A Monte Carlo Study, Mean-Field Theory, High-Temperature Series Expansions, and Ab-Initio Calculations. 2023, Springer. p. 19-47.
12. Sahni, V., Modern Density Functional Theory, in Schrödinger Theory of Electrons: Complementary Perspectives. 2022, Springer. p. 215-279.
13. Heryanto, H., et al., Science mapping analysis of density functional theory (DFT) for material design: A review. Jom, 2024. 76(8): p. 4629-4642.
14. Misiurev, D. and V. Holcman, Modeling of Magnetic Films: A Scientific Perspective. Materials, 2024. 17(6): p. 1436.
15. Khan, M.J.I., et al., Investigating structural, electronic, magnetic, and optical properties of Zr doped and Ti-Zr co-doped GaN for optoelectronic applications. Physica Scripta, 2022. 98(1): p. 015821.
16. Taj, I., et al., Theoretical investigations of structural, electronic, magnetic, and optical properties of group V (X= V, Nb, Ta) added CeO2-X materials for optoelectronic applications. Journal of Molecular Modeling, 2024. 30(5): p. 159.
17. Khan, M.J.I., et al., Unveiling the structural, electronic, magnetic, and optical response of BaTiO3-X (Sm/Au/Hf) materials for advanced energy applications. The European Physical Journal B, 2025. 98(5): p. 107.
18. Van Santen, R.A., Mechanisms in heterogeneous catalysis. Vol. 22. 2023: World Scientific.
19. Ali, M., et al., Effective hydrogen storage in Na2 (Be/Mg) H4 hydrides: Perspective from density functional theory. International Journal of Hydrogen Energy, 2024. 64: p. 329-338.
20. Ali, M., et al., Atomistic simulations for electronic structure, mechanical stability and optical responses of sodium-based NaAH3 (A= Sc, Ti and V) metal hydride perovskites for hydrogen storage applications. International Journal of Hydrogen Energy, 2025. 128: p. 749-759.
21. Bibi, Z., et al., Atomistic DFT Simulations are Promising Techniques for the Discovery of Hydrogen Storage Materials. Journal of Engineering, Science and Technological Trends, 2024. 1(2): p. 100-106.
22. Ali, M., et al., DFT study of hydrogen-substituted KMgF3 fluoroperovskite: Structural, electronic, and storage implications. Inorganic Chemistry Communications, 2025: p. 115186.
23. Ali, M., et al., A comprehensive investigation of structural, mechanical and optoelectronics attributes of M2AsC (M= Zr, Hf, Ta, W) MAX phase carbides: A DFT investigation. Journal of Molecular Graphics and Modelling, 2025: p. 109102.
24. Ali, M., et al., Exploring the electronic structure, mechanical stability and optoelectronic responses of arsenic-based M2AsX (M= Nb, Mo and X= C, N) MAX phase ceramics. Journal of Molecular Graphics and Modelling, 2025. 136: p. 108965.
25. Ali, M., et al., Structural, optoelectronic and thermodynamical insights into 2H-ZrO2: A DFT investigation. Inorganic Chemistry Communications, 2024. 160: p. 111891.
26. Ali, M., et al., Enhancement of hydrogen storage characteristics of Na2CaH4 hydrides by introducing the Mg and Be dopant: A first-principles study. International Journal of Hydrogen Energy, 2024. 70: p. 579-590.
27. Ali, M., et al., First-principles investigation of structural, mechanical, and optoelectronic properties of Hf2AX (A═ Al, Si and X═ C, N) MAX phases. Journal of the American Ceramic Society, 2024. 107(4): p. 2679-2692.
28. Ali, M., et al., Physical properties and defect processes of Zr3AC2 (A= Al and Si) MAX phases: Implication for radiation tolerance. Materials Science and Engineering: B, 2024. 308: p. 117619.
29. Ali, M., et al., Layer-sliding-mediated controllable synthetic strategy for the preparation of multifunctional materials. Materials Today Communications, 2023. 37: p. 107022.
30. Jamshir, N., B. Lari, and H. Hassanabadi, The time independent fractional Schrödinger equation with position-dependent mass. Physica A: Statistical Mechanics and its Applications, 2021. 565: p. 125616.
31. Geerlings, P., Historic Overview. Conceptual density functional theory: towards a new chemical reactivity theory, 2022. 1: p. 1-16.
32. Ain, Q., et al., A precise prediction of structure stability and hydrogen storage capability of KCdH3 perovskite hydride using density functional theory calculations. Journal of Energy Storage, 2024. 100: p. 113734.
33. Ali, M., et al., First-principles evaluation of LiCaF3-αHα as an effective material for solid-state hydrogen storage. Journal of Energy Storage, 2024. 83: p. 110731.
34. Ali, M., et al., An accurate prediction of electronic structure, mechanical stability and optical response of BaCuF3 fluoroperovskite for solar cell application. Solar Energy, 2024. 267: p. 112199.
35. Khan, U.A., Computational Approaches for Predicting and Optimizing the Properties of Composite Materials. Advances in Material Technology, 2024: p. 58.
36. Shen, S.C., et al., Computational design and manufacturing of sustainable materials through first-principles and materiomics. Chemical Reviews, 2023. 123(5): p. 2242-2275.
37. Sholl, D.S. and J.A. Steckel, Density functional theory: a practical introduction. 2022: John Wiley & Sons.
38. Salihbegovic, F., Coupled cluster theory applied to zero-dimensional systems. 2023, Technische Universität Wien.
39. Bortone, M., Addressing Electronic Complexity. 2025, King's College London.
40. Liu, B. and W. Lu, Surrogate models in machine learning for computational stochastic multi-scale modelling in composite materials design. International Journal of Hydromechatronics, 2022. 5(4): p. 336-365.

41. Wang, T., et al., Surrogate model via artificial intelligence method for accelerating screening materials and performance prediction. Advanced Functional Materials, 2021. 31(8): p. 2006245.
42. Panda, S.S., et al., Quantum Computing in Materials Science and Discovery, in The Quantum Evolution. 2024, CRC Press. p. 184-210.

Journal of Engineering, Science and Technological Trends (JESTT)
Vol. 2, Issue 3, November 2025
DOI: https://doi.org/10.64060/jestt.v2i3.2

Review

Natural Disaster Prediction and Mitigation through Machine Learning

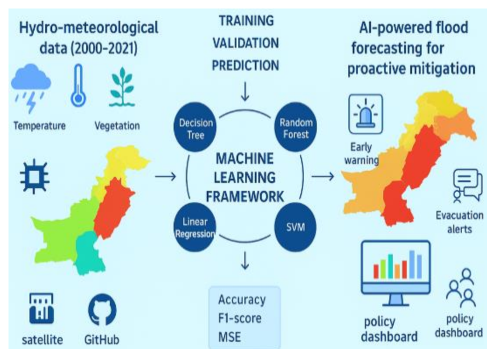
Fatima Abbas[1], Jamil Afzal[2*] and Sidra Shahid[1]

[1]University of Agriculture Faisalabad, Pakistan
[2]International Islamic University, Kuala Lumpur, Malaysia
* Corresponding Email: jamilafzal@iium.edu.my (J. Afzal)

Received: 06 September 2025 / Revised: 13 October 2025 / Accepted: 09 October 2025 / Published online: 15 November 2025

ABSTRACT

Flooding remains a major natural disaster affecting Pakistan's provinces of Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan, with increasing severity due to climate change and human activities. This research explores the application of machine learning techniques to enhance flood prediction accuracy for the years 2025 to 2030. The study utilises historical hydro-meteorological data, including rainfall, temperature, and vegetation indices, to train four machine learning models: Decision Tree, Random Forest, Linear Regression, and Support Vector Machine (SVM). Standard evaluation metrics such as precision, recall, F1-score, and mean squared error (MSE) are used to assess model performance. Results show that Random Forest and SVM outperform the other models in terms of both accuracy and generalizability. These models effectively identify high-risk flood zones across the studied provinces. The findings demonstrate the potential of data-driven approaches to support early warning systems, enabling better disaster preparedness, resource allocation, and mitigation planning. This research highlights how machine learning can play a critical role in reducing flood-related risks and enhancing resilience against future natural disasters in Pakistan.



Keywords*:* Climate Change; Disaster management; Flood Forecasting; Machine learning

1. Introduction

Natural disasters, particularly floods, continue to pose significant threats to human lives, infrastructure, and economies across the globe. The increasing frequency and intensity of such disasters over the past few decades can largely be attributed to climate change, rapid urbanisation, and ongoing environmental degradation [1]. Among these disasters, floods rank as one of the most common and devastating, affecting millions of people annually. They disrupt transportation, communication, and supply chains, while also displacing communities and causing long-term financial losses [2].In response to this growing threat, accurate and timely flood prediction has become critical in minimising the risk to both lives and livelihoods. The evolution of modern technologies, particularly in the domains of data science and machine learning (ML), has opened promising avenues for developing more accurate and adaptive flood forecasting systems. These technologies can process complex and large-scale environmental datasets to identify hidden patterns, offering faster and more informed decision-making for disaster management [3].

Previous research efforts have primarily focused on traditional hydrological models and early-stage statistical techniques for flood prediction. Commonly used models include Rainfall-Runoff models, GIS-based simulation tools, and various statistical forecasting methods, which have provided important baseline insights into flood [4],[5]. However, these approaches often struggle to adapt to real-time conditions and lack the flexibility required to integrate changing environmental variables dynamically. In recent years, the focus has shifted towards machine learning techniques due to their capacity to analyse non-linear relationships and improve predictive performance. Researchers such as Brunner et al. (2021) have demonstrated how models like Random Forests, Decision Trees, Artificial Neural Networks, and Support Vector Machines can offer improved forecasting capabilities [6]. However, many of these efforts have been limited by static or incomplete datasets, insufficient geographical coverage, or a lack of model validation in real-world disaster scenarios. Data scarcity, limited sensor availability, and inconsistent environmental monitoring especially challenge flood prediction efforts in developing countries like Pakistan [7],[8].

Addressing these limitations, the current study proposes a robust machine learning-based flood prediction framework specifically designed for Pakistan’s four most vulnerable provinces: Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan. Historical flood-related data from 2000 to 2021 have been utilised, focusing on essential environmental variables such as rainfall, temperature, vegetation index, and ice melt levels. The selected machine learning models, Decision Tree, Random Forest, Linear Regression, and Support Vector Machine (SVM), were trained and validated using standard benchmarks to ensure reliable and generalised performance. Monthly flood risk forecasts were generated for the years 2025 to 2030, providing an extended window for future disaster planning. The results are intended to support government agencies, policymakers, and emergency response teams in developing region-specific early warning systems and mitigation strategies [9].

The novelty of this study lies in the following contributions:

1. To focus on region-specific flood forecasting using machine learning, tailored to Pakistan's diverse climate zones and historical flood data.
2. To find the most accurate and generalizable ML model by comparing multiple algorithms under consistent validation benchmarks.
3. To explore the impact of hydro-meteorological variables - including temperature, rainfall, vegetation index, and ice - in enhancing flood prediction accuracy.
4. To use ML-generated monthly predictions (2025–2030) for aiding disaster management authorities in early warning systems and risk reduction planning.

2. Prior Investigation

2.1. Machine Learning in Flood Prediction: An Overview

The implementation of Machine learning models alongside classic statistical techniques has greatly increased the speed and reliability of flood predictions. Merely a few years ago, workers in meteorology and environmental science were incorporating machine learning in their forecasts. Earlier, hydrologists would always try to anticipate flood periods with the help of statistical models based on certain historic events. While these models have been useful in certain scenarios, they fail to account for multi-dimensional non-linear outcomes, rendering them ineffective in rapidly changing climatic conditions. The application of ML in these cases has remedied the deficiencies of earlier models. Rather than having to make guesses, ML can now fully extract complex relationships from large datasets and make precise predictions [10]. Originally, hydrological prediction models were expected to be a lot more accurate with the integration of other ML techniques. The reasoning behind the incorporation of physical parameters in earlier models is to strive for the maximum attainable results, which is not the case in machine learning [11]. Among the many methods in ML, the most widely utilised ones are Decision Trees and Random Forests due to their efficiency in high-dimensional datasets. These models also have some modifications which allow the analysis of both categorical and continuous variables at the same time, deal with the missing data optimally, and capture the model parameters by post-estimation processing. This helps in the assessment and control of risks related to potential flood damage [12].

Alongside decision trees and random forests, they have examined the application of more advanced ML methods like Support Vector Machines (SVM) and Artificial Neural Networks (ANN) in the area of flood prediction. SVM is particularly strong in performing classifications of meteorological data because it can simultaneously cluster a multitude of interactions from different flood indicators and also overfit them. At the same time, ANN mimics the functions of a human brain, which allows the model to capture complex dependencies and non-linear relationships among different sets of environmental data. These particular features of SVM and ANN models enable the construction of predictive models of flood events in areas with rapid climatic changes and where most models fail [13]. Implementing machine learning models to flood forecasting has greatly enhanced predicting and making provisions for disasters in numerous flood-prone areas, which is remarkable. These areas have benefited from these models as it helped improve disaster preparedness. With each new feature, a significant amount of information is received; thus, the possible enhancement of machine learning models to assist in flood risk management for decision makers, responders, communities, and others to anticipate and minimise flood impacts is immense. The study of the area requires an urgent application of sophisticated deep learning techniques and hybrid models that will aid in more accurate and effective flood forecasting owing to climate change and other environmental factors.

2.2. Effectiveness of Ensemble Learning Techniques

An ML technology for flood prediction has one of the greatest benefits in the application of ensemble learning strategies, which, through the use of multiple algorithms, have fundamentally transformed predictive models by enhancing their accuracy and effectiveness. Single-model approaches, in some cases, may be helpful, but they are often associated with overfitting, high variance, and sensitivity to noise in the data. On the other hand, ensemble approaches combining the predictive capabilities of numerous models mitigate the mentioned constraints and provide improved generalisation as well as robustness. Empirical evidence has shown that ensemble learning methods, specifically Random Forest, consistently outperform the use of conventional single-model methods by lowering their variance and overfitting, and in turn improve the stability and reliability of flood predictions [14]. This stability in particular, is greatly important in flood forecasting because the accuracy of the prediction has potentially severe consequences within disaster mitigation and preparedness regions. Random Forest is one of the most popular methods of ensemble learning. It is an extension of Decision Trees that improves prediction by building several decision trees and combining their predictions. Unlike a Decision Tree, which can easily overfit and become overly responsive to variations within training data, Random Forest builds a super set of trees by deliberately adding randomness within feature and training sample selection. This helps ensure that the model does not depend too much on a single predictor, hence improving generalisation and the chances of avoiding errors due to noisy or incomplete data [15]. As a flood prediction tool, Random Forest is a good candidate due to its strong performance in high-dimensional datasets, because so many variables comprising the precipitation, river discharge, temperature, and land surface change intimately monitor each other. The use of a collection of individual decision trees makes Random Forest easier to use without compromising prediction accuracy, which is ideal for hydrologists and disaster management officials.

Different ensemble learning methods, like boosting algorithms, have also become popular because of their ability to further improve the predictive accuracy of ML-based flood forecasting models. Some of the boosting methods are: Adaptive Boosting (AdaBoost), Gradient Boosting Machines (GBM). These methods assign new weights to weak learners, in turn making them better. For example, AdaBoost gives higher weights to misclassified instances, which forces the following models to focus on correcting previous mistakes, thus improving the accuracy of flood predictions [16]. Likewise, GBM builds trees in a forward-stage manner, optimising loss functions for each stage to improve the precision of flood-prone area identification. These boosting methods have been able to achieve the identification of the intricate features and the nonlinear dependences of hydrological variables, which is important for estimating the floods. The comparison done between the ensemble learning models and the traditional hydrologic models has shown that the effectiveness of ensemble approaches is better in predicting floods. Some studies show that some classical models, like the popular rainfall-runoff models, for instance, tend to underestimate the variability of meteorological measurements because they operate on predefined relationships between their input factors. However, Rahman et al. argue that ensemble-based ML techniques are more reliable in predicting uncertain environmental conditions [17]. Their ability to utilise multiple predictors and adapt to new data enables them to overcome uncertainties in their environment.

Maintaining flexibility is crucial for areas affected by climate change since the rainfall, extreme weather, and river flow are highly variable. These areas require better models for predicting and controlling flood disasters. The application of ensemble learning approaches in the automation of the processes connected with disaster forecasting has been described in a number of examples, particularly in regions classified as flood regions, where precise prediction is necessary for quick decision-making during a disaster. For instance, Ruichen et al. reported that Pakistan is susceptible and studies proved that ensemble learning models are performing better than the conventional techniques of forecasting, which rely on excessive monsoon rains, glacial thawing, and river flooding [18]. These models have significantly enhanced shallow disaster management in the area by providing timely and accurate predictive warnings, which enable the responsible authorities to evacuate and reinforce the required civic infrastructures. Furthermore, the application of ensemble learning techniques, together with real-time data collection systems such as satellite remote sensing and the Internet of Things (IoT), has made it possible to respond faster and more accurately to flood risks.

2.3. The Need for Province-Specific Model Selection

Although machine learning (ML) models for flood predictions have greatly advanced the accuracy of forecasting floods, the accuracy of these models differs greatly across regions due to differences in climate and hydrology. In Pakistan, which has an extremely diverse topography including the towering mountains of Khyber Pakhtunkhwa and Gilgit-Baltistan, fertile plains of Punjab and arid deserts of Balochistan, a single flood prediction model will most likely not provide accurate predictions within all provinces. These interprovincial differences require that ML model choice be tailored to specific provinces in order to achieve favourable results that consider the geophysical and climatic realities of the region [19]. The need for an appropriate selection of ML models in Pakistan is illustrated through the different flood mechanisms across the country. For example, in Punjab and Sindh Phases, floods are mainly attributed to the monsoons and river flooding from major rivers like the Indus, Jhelum, Chenab, Ravi, and Sutlej. With respect to river flooding, it is found that ensemble methods, including Random Forests and Gradient Boosting Machines (GBM), outperform others because they are simpler to use with large, structured datasets and are powerful in merging multiple meteorological and hydrological parameters [20]. Furthermore, these models capture data about rainfall, river water flow, soil moisture, and previous floods to help estimate possible floods. This helps the policymakers to devise plans that could mitigate potential risks and damages. Moreover, these regions are estimated for large magnitude floods because there are available methods for missing value reconstruction, multifactor examination, and large-scale hydrologic database application.

The dynamics of flooding in Khyber Pakhtunkhwa and Balochistan are predominantly governed by the thawing of glaciers, excessive rains from the mountains, steep slopes, and the region's low soil absorption rates. In comparison to Punjab and Sindh, these regions experience much more irregular, exotic, and violent flash floods instead of the more systematic river flooding. Recent investigations have shown how these provinces perform with Support Vector Machines (SVM) and Artificial Neural Networks (ANN) and most have astonishingly good results owing to the ability of these methods to model complex nonlinear interactions of the most important meteorological factors, which include the temperature, local rainfall, and snowmelt [21]. That is crucial for SVM in the situation when a vast number of environmental drivers in these regions strongly interact to form a highly complex classification. Much like ANNs, SVMs have been capable of capturing the nonlinear relationships in the occurrence of flash floods, thus enabling the activation and dynamic simulation of hydrological responses to climatic variations. In light of these differences on a provincial scale, the specialists highlight the need for developing training sets that correspond to the meteorological and hydrological characteristics of every region. Every flood prediction model associated with the implementation of ML techniques performs effectively with the right input data; thus, local climatic conditions, historical flooding, and sensor data are essential for the right predictions [22]. For example, Punjab and Sindh possess rivers whose data must be included in the databases of river inflow, precipitation, and groundwater level changes. Conversely, Khyber Pakhtunkhwa and Balochistan need to focus more on glacier movement, the land surface configuration, and the occurrence of intense storms. Ignoring such parameters would most likely lead to forecasting errors which would defeat the whole purpose of the ML flood prediction systems that disaster management organisations have to work with.

The use of IoT sensors together with remote sensing devices enables better integration of data to enhance the accuracy of province-level ML models in real time [23]. Active monitoring reveals a number of environmental factors, for example, remote sensing of surface water, precipitation, and changes in the land surface, which have the potential to increase the advantage that machine learning models could have if these parameters are provided as real-time inputs [24]. Punjab and Sindh's regions of Punjab and Sindh are served by the Sentinel 1 and Landsat satellites, which help in observing river basin borders and help in advancing flood models by detecting conditions that may cause river overbanking. Enabling real-time flood prediction from glacier melting, local rainfall, and soil moisture IoT-based sensor networks in the mountainous regions of Khyber Pakhtunkhwa and Balochistan greatly assists [25]. These data sources make it possible for ML algorithms to design sophisticated dynamic forecasting systems that increase intervention and disaster management system functionality.

In addition, further work should be devoted to the design of optimal hybrid models that combine several machine learning algorithms for alleviating regional discrepancies in flood prediction precision. It is true that deep learning approaches, or any single machine learning algorithm, will outperform the rest in one environment while underperforming in others, which is why there is a growing interest in hybrid models that combine different approaches. For example, the combination of ensemble learning algorithms with deep learning Convolutional Neural Networks (CNNs) is known to improve the accuracy of flood prediction by more fully exploiting structured and unstructured data patterns [26]. Similarly, other hybrid models that combine physical hydrological models and machine learning prediction models may improve flood forecasting and permit broader usage in different regions. The adoption of a provincial-level machine learning model selection is particularly important considering the climate change-induced irregularity of floods in the region of Pakistan. Over the past ten years, there has been an increased incidence of extreme weather, with the addition of monsoon rains being increasingly warmer, warmer temperatures, melting of glaciers, and other more recent changes in rainfall distribution further worsening the flooding's forecasting problem. In order to tackle these problems, there is a need to develop real-time climate impact adaptive ML frameworks that would strengthen resilience to disastrous flooding. It is beyond doubt that there is a requirement to fundamentally rethink the floods' prediction repeating periodicity and also the algorithmic models which should be implemented in the provinces of Pakistan to enhance the preparedness and response to the changing, complex and adverse hydrological threats.

3. Methodology

3.1. Data Collection

In order to create a credible and precise machine learning (ML)-based flood prediction model, a dataset with all relevant features concerning floods and their contributory aspects was collected from all possible sources, with a focus on publicly available open source datasets from GitHub. The dataset is compiled for the years 2000 to 2021 with the objective of having a sufficient time history of actual floods and the causative factors for floods, like the environment during that time period and to understand the outer phenomena over a longer term, which helps to better capture accurate results to enhance the ML models' predictive scope over time. The dataset was organised from 2000 to 2021 so as to capture insights over extensive durations, which helps in capturing cyclical tendencies. The criteria for research materials were determined after a thorough examination of available study materials and discussions with specialists in hydrology and climate sciences. The important variables selected in the dataset are precipitation, temperature, ice, and vegetation index data. These features were incorporated due to their significant influence on the incidence of floods, which is testified in past records [27]. Floods are caused primarily due to excessive downpour and supporting data explains rainfall as one of the strongest factors, alongside temperature, on the chances of floods and urban flooding [28]. Temperature, along with rainfall, exerts strong impacts, but understanding seasonal temperature variations along with snow, melting and glaciers' retreat provides deep insights into regions of Khyber Pakhtunkhwa and Gilgit-Baltistan, where glacial melting floods are prevalent. Gathering ice level data assisted in tracking the seasonal and long-term shifts in glacier masses since GLOF events have become more common lately because of climate change. The dataset consisted of satellite iceberg level measurements, which help evaluate how much glaciers are melting and the flooding that occurs downstream as a result [29],[30]. Also, deriving from the flood satellite imagery and remote sources, the vegetation index was added to track changes in land cover, soil moisture retention, deforestation, and vegetation growth that impact the frequency and severity of floods. Regions with greater vegetation cover tend to absorb more water and lessen the potential for floods, while deforested areas are prone to increased runoff and flooding because of less vegetation. The dataset was collected from a number of public GitHub repos which collect and process climate datasets from reputable sources such as NASA, NOAA, ESA, PMD, and many others. These repositories provided raw climate data, remote data, and precipitation records and then sanitized the data to make it appropriate for machine learning model training and testing. In the interest of assuring reliable data, a series of preprocessing steps were executed, including data normalization, missing value imputation, and outlier treatment. Historical discrepancies and gaps within the records were corrected by checking them against available data and verifying its correctness. Moreover, archival flood event information was integrated with other datasets to enable the model to better adapt to climate changes over time. The cleaned dataset was also split into training and testing datasets to ease model development and evaluation. Later stages of the study could incorporate more real-time data from IoT devices, which would increase the accuracy of the models.

3.2 Data Preprocessing

To ensure the quality and reliability of the dataset before training the machine learning (ML) models, several preprocessing techniques were applied. Proper data preprocessing is essential for minimising biases, handling inconsistencies, and optimising model performance. The primary steps included handling missing values, feature selection, and data normalisation.

Handling Missing Values: In order to maintain the dataset's accuracy and relevance prior to training the machine learning (ML) models, a set of preprocessing methods was conducted. Appropriate data categorisation is crucial for reducing biases, correcting errors, and improving the model's effectiveness. These procedures mostly comprised missing data treatment, relevant feature selection, and data standardisation.

Feature Selection: A critical stage in data cleansing revolved around pinpointing the parameters that had the greatest influence on flooding. A correlation study was undertaken by applying Pearson's correlation coefficient to determine the associations between independent variables (precipitation, temperature, ice, and vegetation index) and a dependent variable (flooding). Those features which showed a strong correlation with flood occurrence were preserved, while those with very little impact were eliminated to ensure noise reduction and improve model effectiveness. A Variance Inflation Factor analysis was also performed to test for multicollinearity among features to ensure that non-essential variables did not bias model estimate outcomes.

Data Normalisation: MinMaxScaler was implemented to scale features to the same range of values. All numerical features were scaled to lie within the range of zero and one. This approach helps preserve the original distributions while preventing features from skewing model training. Such normalisation was particularly effective in enhancing convergence and guaranteeing even weight distribution across predictors. These methods were particularly useful for SVM and Linear Regression, which are sensitive to feature magnitudes.

4. Model Selection and Implementation

In order to create a strong and reliable system for flood prediction, four separate machine learning models were implemented with Scikit-learn. These models were selected because they could separately tackle classification, regression, and ensemble learning, which are part of flood forecasting.

Decision Tree: Classification based on Gini impurity was structured using a Decision Tree model. This model was picked because of its flexibility and ease of interpretation for non-linear relationships between input variables. Decision Trees break down the dataset into successive layers of decision points, which is helpful in pinpointing the limits of flood-inducing parameters, aiding in flood.

Random Forest: In order to refine accuracy and mitigate overfitting, a Random Forest model was applied as an ensemble learning technique. Random Forest enhances the effectiveness and dependability of flood forecasting by constructing multiple Decision Trees and combining their outputs. This approach is especially beneficial for high-dimensional data and for cases with absent data points while reducing bias.

Linear Regression: A linear regression was incorporated to estimate flood severity by using numerical variables like rainfall and river water level. This model attempts to partition the flooding by quantifying it, thereby ascertaining the impending threat level of floods by applying a linear function to this dataset.

Support Vector Machine (SVM): An SVM model differentiated whether a flood occurred or not (binary outcome) by classifying it as a flood or no flood. With this, an SVM model can effectively classify data points based on any historical data. SVMs were selected for this project due to their efficient performance with high-dimensional data and their tendency to overfit smaller datasets. The following Figure 1 represents the average accuracy distribution of machine learning models for flood prediction.

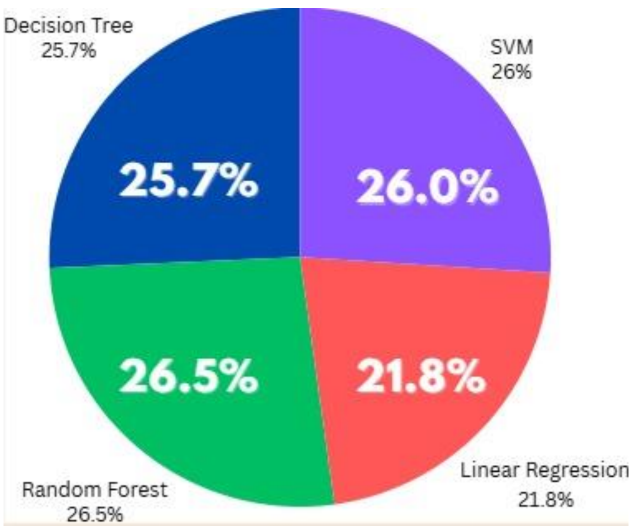


Figure 1: *Average accuracy distribution of machine learning models for flood prediction*

To assess the performance of the implemented models, several evaluation metrics were used, ensuring a comprehensive analysis of their predictive accuracy and reliability. The selected metrics included:

- **Accuracy:** Measures the proportion of correct predictions out of total predictions, providing an overall assessment of model effectiveness.
- **Precision:** Evaluates the proportion of true positive flood predictions out of all positive predictions, indicating the model's reliability in correctly identifying flood occurrences.
- **Recall:** Measures the ability of the model to detect actual flood events, crucial for minimising false negatives and ensuring timely warnings.
- **F1-Score:** A harmonic mean of precision and recall, balancing both metrics to provide a robust performance measure, particularly for imbalanced datasets.
- **Mean Squared Error (MSE):** Applied to regression models like Linear Regression, MSE quantifies the average squared difference between actual and predicted flood severity levels, helping assess prediction accuracy in continuous output variables.

Table 1 summarises the model accuracy across the four provinces.

As shown in Figure 2, Random Forest and SVM consistently outperform other models, demonstrating superior predictive capabilities. Figure 3 represents the Actual vs predicted occurrence of floods in Punjab, 2025-2030.

- **Punjab:** High flood risks in July due to peak monsoon rainfall.
- **Sindh:** Seasonal variations with increased risks from June to August.
- **KPK:** Notable risk fluctuations, with high risks in mountainous regions.
- **Balochistan:** Moderate flood risks, with peaks in July-August.

Table 1:
Model accuracy across the different provinces of Pakistan

Province	Decision Tree (%)	Random Forest (%)	Linear Regression (%)	SVM (%)
Punjab	88.13	99.15	42.23	99.15
Sindh	96.98	96.98	96.98	96.98
KPK	99.92	98.40	90.83	90.83
Balochistan	97.17	99.00	93.50	99.00

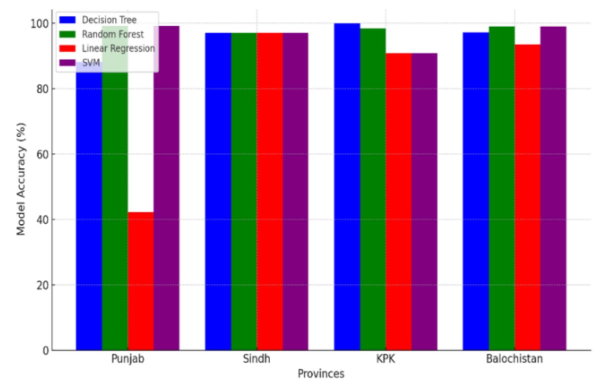


Figure 2: *Comparison of Model Accuracy Across Provinces*

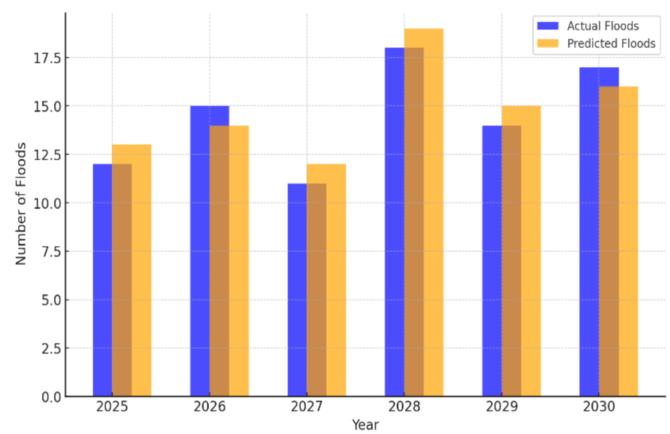


Figure 3: *Actual vs Predicted occurrence of flood in Punjab*

5. Discussion

This study's findings verify that ensemble models, especially Random Forest, are more effective than other machine learning methods in accurately predicting floods. The effectiveness of Random Forest can be attributed to its capability to handle large and complex multidimensional data sets without suffering from overfitting or data imbalance. Due to the aggregation of multiple decision trees, the model comprehensively captures the complex underlying relationships among environmental variables and is well-suited for regions prone to flooding that exhibit high variability of climatic and hydrologic phenomena. The model's strength in dealing with absent data, the ability to capture categorical and continuous data, and the diminishment of noise within predictions showcase the model's efficiency in flood forecasting. In this study, one of the major benefits of using Random Forest was the ability to obtain accurate flood forecasts in various regions without changes in rainfall, land surface conditions, and water flow changing accuracy. Unlike conventional hydrological models that depend on fixed equations and preconceived notions, Random Forest takes into consideration the real-time data, which makes it more flexible and responsive in predicting flood occurrences. This feature is specifically useful for nations such as Pakistan, which, owing to its varied topography from the glacial Khyber Pakhtunkhwa region to the flat valley of Sindh and Punjab, has to deal with environmental factors which need to be incorporated in the model.

The research also claimed that Linear Regression, which is very popular in predictive analysis, has profound limitations in areas with high climate variability. Using Linear Regression, it is assumed that the relationships of each pair of variables do not alter over time. It also does not estimate nonlinear dependencies in flood-associated data very accurately. In Balochistan and Khyber Pakhtunkhwa regions, where flash floods occur due to sudden and extreme precipitation or ice cap melting, Linear Regression fails to make the necessary predictions. The model does poorly because of its restrictive assumptions about complex interrelated variables, which deeply reduce its accuracy performance. Thus, it is not suitable to serve as an independent flood prediction model in these areas. The findings of this study reiterate the need to develop machine learning models specific to a region in order to enhance the precision of flood forecasting. Due to the geographical diversity of Pakistan, an all-encompassing model will not work for flood prediction. For example, Random Forest outperformed the other methods in Punjab and Sindh, where monsoonal precipitation and riverine flooding are the main issues, but SVM and ANN are better suited to Khyber Pakhtunkhwa and Balochistan, where flooding is associated with glacial lake outburst and sudden shifts in precipitation. This illustrates the need for predictive models which are focused towards the specific hydro-meteorological characteristics of each province as opposed to basing it on a generalised methodology.

The results also emphasise the importance of localised training datasets for improving model performance. The flood prediction models that used region-specific data, which included historical flood patterns, local weather, and topographical data, outperformed all other models trained on more general datasets by a wide margin. This indicates that model predictions can be further improved by the use of real-time data sources such as satellite and remote sensing images or IoT sensors that can modify environmental variables in real-time. Incorporating such real-time data streams into ML models enables dynamic forecasting, which allows the timely issue of flood alerts and better disaster management response.

6. Conclusion

Flood is being classified as an increasing risk to life, infrastructure and economic development, especially in areas like Pakistan, where there is climatological and geographical variation which leads to timely and severe flooding. As the decades have passed, flooding prediction has relied more and more on the standard hydrological models, which were useful but limited as they are not capable of describing the complex, non-linear relationships between different environmental factors. However, with the availability of machine learning (ML), the scope for more advanced tools that can handle enormous data sets, recognise concealed patterns and enhance predictions of flood forecasting has become available to researchers and policymakers. The goals of this study are to demonstrate how different ML algorithms can accurately predict floods in Pakistan, arguing the most pronounced results for prediction using Random Forest and Support Vector Machines (SVM). These Models have proven their ability to manage big data or high-dimensional data sets along with uncertainty in meteorological data. Through these advanced models, the study sponsors AI-enabled flood disaster management programs that may substantially improve early warning systems for disaster prevention and minimise losses associated with floods. However, to provide more accurate and relevant adaptation for these kinds of models, further development is needed regarding real-time weather data integration. The integration of machine learning algorithms has advanced flooding prediction as it shifted research from the use of rigid static methods to more adaptive and multifaceted methods that can model the interrelationships between the variables in question. Among the tested ML models, Random Forest and SVM algorithms outperformed other techniques, such as Decision Trees and Linear Regression, on metrics suited for tele monitoring systems. Predictive flooding forecasting, as well as anticipating the extent of flooding, was highly accurate from using Random Forest, a multi-decision tree ensemble learning method. One of the most important features of Random Forest is, however, mediating missing values and the ability to classify and quantify variables without the need for extensive time consumptive preprocessing. Additionally, Random Forest averts single decision tree model overfitting, which compounds error by averaging predictions, which is another strength that augments its effectiveness in regions such as Punjab and Sindh, where monsoonal flooding and overbank flooding due to the Indus River system are common.

On the contrary, SVM was extremely efficient at classifying flood events given the complex meteorological and hydrological factors. Unlike other modelling techniques that rely on regression analysis and use the linear model as a first assumption, SVM deals with non-linear distributions of data, making it appropriate for Khyber Pakhtunkhwa and Balochistan, which are known to be erratic. These provinces are characterised by sudden flash floods caused by heavy rainfall and rapid glacial melt. SVM performed well with some sets of data due to its ability to separate features corresponding to flood conditions from those related to non-flood situations by optimising the decision boundary. While SVM performed best, other models like Decision Trees and Linear Regression did offer some value, but they weren't as reliable or flexible, which is essential for predicting floods with high accuracy. Decision Trees are computationally simple and cheap when it comes to classifying data, but they do suffer from overfitting when applied to datasets with high variability, when it comes as climate. Regression analysis, especially in the form of Linear Regression, which falls under the umbrella of quantitative analysis where equations are used to depict the relationship between the relevant factors and the occurrence of floods, proved ineffective in places where multiple, highly interacting flood drivers exist, such as terrain slope, soil moisture, and wind direction. In terms of prediction, the earlier floods can be projected, the better prepared society can be for responding to the disaster. Predictive alert systems based on Machine learning models, such as Random Forest and SVM, have the potential to assist governmental agencies as well as citizens in protecting themselves from various forms of floods. By warning users days before actual floods take place, these warning systems can enhance resource deployment for aid services, reduce casualties, and minimise infrastructural damage. In regions of Pakistan which are prone to floods, employing SVM and Random Forest models in alert systems may prove beneficial. Signals for late evacuation due to poorly forecasted floods resulted in large-scale financial losses during the 2022 floods that affected Sindh and Balochistan. These models would enable authorities to provide timely alerts if proper forecasting tools were in place, resulting in proactive measures being employed. There is a pressing need for AI-centred technology in flood forecasting and management, as well as for investment in holistic approaches that incorporate using such models to enhance national disaster management strategies.

Declaration

Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contribution Statement: FA, JA, and SS conceived the idea and designed the research; Analyzed and interpreted the data and wrote the paper.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Consent to Publish: All authors agree to publish this version of the manuscript in JESTT.

Data Available Statement: Data will be made available on request by Fatima Abbas (fatimaabbas3815@gmail.com)

Clinical trial number: Not Applicable

Ethical Approval: N/A

Consent to Participate: N/A.

References

1. El-Magd, S.A.A., B. Pradhan, and A. Alamri, *Machine learning algorithm for flash flood prediction mapping in Wadi El-Laqeita and surroundings, Central Eastern Desert, Egypt.* Arabian Journal of Geosciences, 2021. **14**(4): p. 323.
2. Ahmad, A., et al., *Flood risk modelling by the synergistic approach of machine learning and best-worst method in Indus Kohistan, Western Himalaya.* Geomatics, Natural Hazards and Risk, 2025. **16**(1): p. 2469766.
3. Akhyar, A., et al., *Deep artificial intelligence applications for natural disaster management systems: A methodological review.* Ecological Indicators, 2024. **163**: p. 112067.
4. Efraimidou, E. and M. Spiliotis, *A gis-based flood risk assessment using the decision-making trial and evaluation laboratory approach at a regional scale.* Environmental Processes, 2024. **11**(1): p. 9.
5. Alshammary, M.J., I.R. Karim, and M.Y. Fattah, *Monitoring Flood Waves Due to Overtopping: Case Study of Mosul Dam from Iraq.* Journal of Engineering, Science and Technological Trends, 2025. **2**(1).
6. Brunner, M.I., et al., *Challenges in modeling and predicting floods and droughts: A review.* Wiley Interdisciplinary Reviews: Water, 2021. **8**(3): p. e1520.
7. Islam, A.R.M.T., et al., *Predicting flood risks using advanced machine learning algorithms with a focus on Bangladesh: influencing factors, gaps and future challenges.* Earth Science Informatics, 2025. **18**(3): p. 300.
8. Afzal, J., et al., *Effects of dam on temperature, humidity and precipitation of surrounding area: a case study of Gomal Zam Dam in Pakistan.* Environmental Science and Pollution Research, 2023. **30**(6): p. 14592-14603.
9. Jain, H., et al., *Leveraging machine learning algorithms for improved disaster preparedness and response through accurate weather pattern and natural disaster prediction.* Frontiers in Environmental Science, 2023. **11**: p. 1194918.
10. Khalaf, M., et al., *IoT-enabled flood severity prediction via ensemble machine learning models.* IEEE Access, 2020. **8**: p. 70375-70386.
11. Khan, I., et al., *Climate change impact assessment, flood management, and mitigation strategies in Pakistan for sustainable future.* Environmental Science and Pollution Research, 2021. **28**(23): p. 29720-29731.
12. Khan, M., et al., *Developing a machine learning-based flood risk prediction model for the Indus Basin in Pakistan.* Water Practice & Technology, 2024. **19**(6): p. 2213-2225.
13. Kumar, V., et al., *Machine learning applications in flood forecasting and predictions, challenges, and way-out in the perspective of changing environment.* AIMS Environmental Science, 2025. **12**(1): p. 72-105.
14. Shaikh, T.A., et al., *A fundamental overview of ensemble deep learning models and applications: systematic literature and state of the art.* Annals of Operations Research, 2024: p. 1-77.
15. Lawal, Z.K., H. Yassin, and R.Y. Zakari. *Flood prediction using machine learning models: a case study of Kebbi state Nigeria*. in *2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*. 2021. IEEE.
16. Mahmood, S., A. Sajjad, and A.-u. Rahman, *Cause and damage analysis of 2010 flood disaster in district Muzaffar Garh, Pakistan.* Natural hazards, 2021. **107**(2): p. 1681-1692.
17. Rahman, Z.U., et al., *GIS-based flood susceptibility mapping using bivariate statistical model in Swat River Basin, Eastern Hindukush region, Pakistan.* Frontiers in Environmental Science, 2023. **11**: p. 1178540.
18. Ruichen, M., et al., *Vegetation variation regulates soil moisture sensitivity to climate change on the Loess Plateau.* Journal of Hydrology, 2023. **617**: p. 128763.
19. Shehzadi, M., et al., *Enhancing flood resilience: Streamflow forecasting and inundation modeling in pakistan.* Engineering Proceedings, 2023. **56**(1): p. 315.
20. Siddiqui, M.A., et al., *Advance Ensemble Flood Warning System: A Case Study for Nullah Lai.* Environmental Sciences Proceedings, 2023. **25**(1): p. 96.
21. Tehrany, M.S., B. Pradhan, and M.N. Jebur, *Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS.* Journal of hydrology, 2014. **512**: p. 332-343.
22. Ahmed, R., *Essays on Infrastructure, Firm Productivity, Natural Disaster and Life Course Transition in South Asia*. 2017.
23. Nishtar, Z. and J. Afzal, *A review of real-time monitoring of hybrid energy systems by using artificial intelligence and IoT.* Pakistan Journal of Engineering and Technology, 2023. **6**(3): p. 8-15.
24. Waleed, M. and M. Sajjad, *Advancing flood susceptibility prediction: A comparative assessment and scalability analysis of machine learning algorithms via artificial intelligence in high-risk regions of Pakistan.* Journal of Flood Risk Management, 2025. **18**(1): p. e13047.
25. Afzal, J., et al., *A complex wireless sensors model (CWSM) for real time monitoring of dam temperature.* Heliyon, 2023. **9**(2).
26. Rajab, M.A., F.A. Abdullatif, and T. Sutikno, *Classification of grapevine leaves images using VGG-16 and VGG-19 deep learning nets.* TELKOMNIKA (Telecommunication Computing Electronics and Control), 2024. **22**(2): p. 445-453.
27. Khan, M., et al., *Assessing Agrometeorological Damage after the 2022 Pakistan Floods: Insights from Multi-Sensor Satellite Data.* Earth Systems and Environment, 2025: p. 1-21.
28. Şen, Z., *Flood modeling, prediction and mitigation*. 2018: Springer.
29. Goldberg, M.D., et al., *Mapping, monitoring, and prediction of floods due to ice jam and snowmelt with operational weather satellites.* Remote Sensing, 2020. **12**(11): p. 1865.
30. Benn, D.I., et al., *Response of debris-covered glaciers in the Mount Everest region to recent warming, and implications for outburst flood hazards.* Earth-Science Reviews, 2012. **114**(1-2): p. 156-174.

Journal of Engineering, Science and Technological Trends (JESTT)

Vol. 2, Issue 3, November 2025
DOI: https://doi.org/10.64060/jestt.v2i3.3

Article

Neutrosophic Logic Reveals Ontological Indeterminacy in Quantum Systems

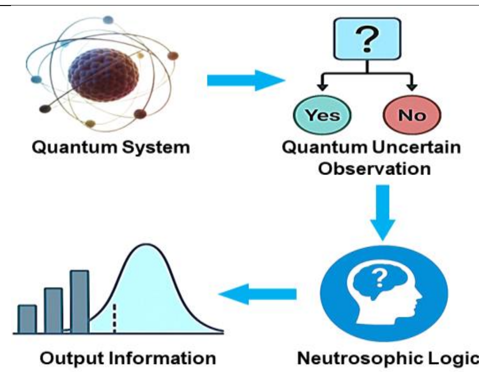
Ansa Firdous[1]* and Hafiza Kanza Maryam[2]

[1]School of Electrical Engineering and Information, Tianjin University, Tianjin 300072, China
[2]School of Chemical Engineering and Technology, Herbin Institute of Technology Shenzhen, Shenzhen, China
* Corresponding Email: ansaeducation7@gmail.com (A. Firdous)

Received: 18 July 2025 / Revised: 03 October 2025 / Accepted: 24 October 2025 / Published online: 15 November 2025

ABSTRACT

Quantum mechanics reveals a reality where uncertainty is not merely epistemic but ontological, challenging classical logics that rely on determinacy. We introduce neutrosophic logic as a novel framework to formalise this intrinsic indeterminacy. Unlike probability or fuzzy systems, neutrosophy treats truth (T), falsity (F), and indeterminacy (I) as independent dimensions, enabling a richer representation of quantum superposition, entanglement, and measurement collapse. We show that neutrosophic triplets capture the undefined yet real state of qubits before observation and provide a coherent description of entangled correlations beyond probabilistic models. This approach offers a unifying formalism that accommodates both structural indeterminacy and outcome probabilities, bridging foundational debates in quantum theory with practical advances in quantum computation and information science. By explicitly integrating indeterminacy into mathematical modelling, neutrosophy advances our understanding of quantum reality and suggests new avenues for experimental validation and quantum technology design.



Keywords*:* Entanglement; Indeterminacy; Neutrosophic logic; Superposition; Quantum information science; Quantum uncertainty

1. Introduction

Quantum uncertainty is one of the most profound and defining features of modern physics. Unlike the deterministic framework of classical mechanics, where the future trajectory of a system can be precisely predicted given initial conditions, quantum mechanics reveals a world where outcomes are inherently probabilistic and indeterminate [1],[2],[3],[4]. The uncertainty principle articulated by Heisenberg illustrates that certain pairs of physical quantities, such as position and momentum, cannot be simultaneously known with arbitrary precision. Beyond this well-known principle, quantum systems exhibit superposition, where a particle exists in multiple states simultaneously until measured and entanglement, in which particles share correlations that defy classical locality. These phenomena suggest that uncertainty in quantum mechanics is not merely a product of incomplete knowledge but a fundamental characteristic of reality itself [5],[6]. Traditional mathematical and logical tools, which presuppose determinacy and binary truth values, often struggle to adequately capture this intrinsic uncertainty. Thus, the challenge remains: how can we model quantum systems in a way that respects both their probabilistic outcomes and their indeterminate nature before measurement?

Neutrosophic theory, introduced by Florentin Smarandache in the late 1990s, offers a novel framework to address precisely this type of complexity [7],[8],[9]. At its core, neutrosophy generalises classical logic by introducing three independent components: truth (**T**), indeterminacy (**I**), and falsity **(F**) to describe the state of any proposition or system. Unlike classical or fuzzy logic, neutrosophy does not require these three values to be complementary or to sum to unity [10]. Instead, they are treated as independent, allowing a proposition to be simultaneously true, false, and indeterminate to varying degrees. This innovation is especially powerful in contexts where contradiction and incompleteness coexist, as it provides a richer language for describing uncertain and paradoxical states. Beyond its philosophical implications, neutrosophic logic has been formalised through neutrosophic sets and numbers, which allow independent specification of **T**, **I** and **F** values for each element in a system. These tools extend beyond fuzzy or probabilistic models by explicitly incorporating indeterminacy, making neutrosophy particularly suitable for systems characterised by openness, contradiction, and undefined states.

Applied to the quantum domain, neutrosophy provides a powerful means of conceptualising and quantifying quantum uncertainty [11],[12],[13]. A particle in a superposition of states, such as spin-up and spin-down, can be described neutrosophically as being partly true (up), partly false (not-up), and partly indeterminate (undefined until measurement). This differs fundamentally from probabilistic approaches, which only express likelihoods, and from fuzzy approaches, which only allow partial membership values. Neutrosophy instead acknowledges that a system may contain intrinsic indeterminacy, not reducible to ignorance, but built into the structure of quantum reality. Similarly, entangled systems can be modelled as coupled neutrosophic triplets, where the indeterminacy of one particle is directly linked to that of its partner. In this way, neutrosophy not only captures the probabilistic aspects of quantum systems but also provides a formal means to represent their ontological indeterminacy. Such a perspective has potential applications in quantum computation, where qubits inherently exploit indeterminate states; in quantum information theory, where measures of entropy and coherence depend on how uncertainty is defined; and in quantum interpretation, where debates about the meaning of the wavefunction hinge on the role of indeterminacy.

This study aims to explore how neutrosophic logic and neutrosophic sets can enhance our understanding of uncertainty in quantum systems. Specifically, the paper seeks to demonstrate that neutrosophy provides a conceptual and mathematical framework better aligned with the fundamental nature of quantum mechanics than classical probabilistic or binary logics. The objectives are fourfold: (1) to articulate the limitations of conventional approaches to quantum uncertainty and highlight the unique contributions of neutrosophy; (2) to apply neutrosophic principles to key quantum phenomena such as superposition, entanglement, and wavefunction collapse; (3) to discuss potential applications of neutrosophic models in quantum computation and information theory; and (4) to consider the broader philosophical and scientific implications of adopting a neutrosophic perspective in quantum studies. By achieving these objectives, the study aspires to contribute not only to the theoretical foundations of quantum mechanics but also to practical advances in quantum technologies and to ongoing debates about the interpretation of quantum reality.

2. Quantum Uncertainty: Challenges and Phenomena

Uncertainty lies at the heart of quantum mechanics, shaping its formalism and its interpretation. Unlike classical systems, where uncertainty often arises from incomplete information or practical limitations of measurement, quantum uncertainty is intrinsic and structural. The principles of superposition, wavefunction collapse, entanglement, and the Heisenberg uncertainty relation all reveal that quantum systems resist full determinacy in ways that defy classical logic. These phenomena pose challenges not only for physical theory but also for the logical and mathematical tools we use to model them.

2.1. Superposition

One of the most striking features of quantum mechanics is superposition [14],[15]. A quantum system can exist in a linear combination of basis states, such as a qubit described by the state Equation 1:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad (1)$$

Where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are complex amplitudes subject to the normalisation condition $|\alpha|^2 + |\beta|^2 = 1$. Unlike classical probabilities, which represent uncertainty about which definite state a system occupies, superposition represents a genuine co-existence of states before measurement. A qubit in superposition is not "sometimes $|\mathbf{0}\rangle$" and "sometimes $|\mathbf{1}\rangle$," but rather exists simultaneously in both states until a measurement is made.

These challenges classical notions of truth and falsity. A statement such as "the qubit is in state $|\mathbf{0}\rangle$" is neither wholly true nor wholly false while the system is in superposition. Instead, the statement has a truth-like weight associated with $|\boldsymbol{\alpha}|^2$ and a falsity-like weight associated with $|\boldsymbol{\beta}|^2$, but these do not exhaustively describe the system. There is also an additional element: the indeterminacy inherent in the superposed condition, which cannot be reduced to probabilities alone. The system's true condition remains undefined until interaction, highlighting a gap in classical logic that neutrosophic reasoning can potentially fill.

2.2. Measurement and Collapse

A central puzzle in quantum mechanics is the measurement problem, which arises from the apparent discontinuity between unitary evolution and wavefunction collapse [16],[17],[18]. According to the Schrödinger equation, quantum states evolve smoothly and deterministically over time. Yet, when an observation is made, the wavefunction seems to abruptly "collapse" into one of the possible eigenstates associated with the measurement.

For example, measuring the state of the qubit above will yield either $|\mathbf{0}\rangle$ with probability $|\boldsymbol{\alpha}|^2$ or $|\mathbf{1}\rangle$ with probability $|\boldsymbol{\beta}|^2$. After measurement, the superposition vanishes, and the system resides in a definite state. This transition raises profound questions: was the qubit truly in both states before observation, or was the superposition merely a representation of our knowledge? Why does the deterministic evolution suddenly give way to a probabilistic collapse upon measurement?

From a logical perspective, the pre-measurement state resists classification. Propositions such as "the qubit is $|\mathbf{0}\rangle$" cannot be neatly labelled as true or false before observation. The outcome only becomes definite post-measurement, illustrating that truth values in quantum mechanics can be contextual and emergent. Classical binary frameworks are insufficient to capture this discontinuity, and even probabilistic descriptions fail to represent the undefined nature of the pre-collapse state.

2.3. Entanglement

Perhaps no phenomenon illustrates the strangeness of quantum mechanics more vividly than entanglement [19]. When two or more particles interact in such a way that their states become correlated, they can no longer be described independently. Instead, the system is represented by a joint wavefunction that encodes their inseparability.

A simple example is the Bell state as shown in Equation 2:

$$|\Phi\rangle^+ = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \qquad (2)$$

Which describes two perfectly correlated qubits. Measuring the first qubit immediately determines the state of the second, no matter how far apart they are spatially. Einstein famously referred to this as "spooky action at a distance," as it challenges classical notions of locality and realism.

Entanglement complicates the representation of uncertainty in profound ways [20]. For each particle, the state appears maximally indeterminate: each qubit alone is in a mixed state, offering no definite outcome before measurement. Yet, taken together, the pair exhibits perfect correlation. This duality of individual indeterminacy combined with collective determinacy cannot be adequately expressed in conventional logical frameworks. It requires a richer language that can describe how uncertainty is distributed across systems, how indeterminacy can exist locally but be resolved non-locally, and how truth values emerge relationally.

2.4. Heisenberg's Uncertainty Principle

Finally, the uncertainty principle reveals a structural limitation in quantum systems that has no classical analog [21],[22]. Proposed by Werner Heisenberg in 1927, the principle asserts that certain pairs of conjugate variables, such as position (**x**) and momentum (**p**), cannot be simultaneously known with arbitrary precision as mention in Equation 3:

$$\Delta x . \Delta p \geq \frac{\hbar}{2} \tag{3}$$

This inequality does not arise from flaws in measurement apparatus or incomplete knowledge; rather, it reflects the mathematical structure of quantum mechanics. Position and momentum operators do not commute, meaning that the system itself does not possess simultaneous definite values of these properties. Thus, the uncertainty principle captures a structural indeterminacy, not an epistemic one.

The implications are far-reaching. A particle does not have a sharply defined trajectory in space-time; instead, it exists in a spread of possibilities constrained by the uncertainty relation. Classical reasoning, which assumes that an object must have well-defined properties at all times, is insufficient for describing such systems [23]. Logical frameworks that insist on binary truth values, "the particle has a precise position" or "the particle does not", cannot represent the reality of a particle that fundamentally lacks such precision.

3. Quantum Uncertainty: Challenges and Phenomena

Uncertainty has been a central topic of inquiry in both science and philosophy for centuries. Long before the advent of quantum mechanics, mathematicians and logicians developed tools to model unpredictability, vagueness, and incomplete knowledge. Classical probability theory, fuzzy logic, and intuitionistic fuzzy logic represent three of the most influential frameworks for capturing different forms of uncertainty. Each of these approaches has proven highly valuable in domains ranging from statistics and engineering to artificial intelligence and decision theory. However, when applied to quantum systems, they exhibit fundamental limitations. Quantum uncertainty is not simply a matter of incomplete information or linguistic vagueness; it is structural and ontological, embedded in the very fabric of quantum states. In this section, we examine the strengths and shortcomings of these traditional approaches, highlighting why they cannot fully capture the indeterminacy inherent in quantum mechanics.

3.1. Classical Probability

Classical probability theory provides a rigorous framework for quantifying uncertainty about outcomes [24]. Based on the axioms formulated by Kolmogorov, probability describes uncertainty as a measure over a sample space of possible outcomes [25]. Each outcome is assumed to exist as a definite state of the world, even if the observer does not know which state has occurred. For example, the probability of rolling a six on a fair die is **1/6**. The die will, in fact, land on one of the six faces; the probability distribution reflects the observer's ignorance of which particular outcome will occur before the roll.

In many areas of science and engineering, probability theory offers a powerful and reliable language for prediction and inference. In classical statistical mechanics, for instance, probabilities describe ensembles of particles, allowing predictions about macroscopic properties such as temperature and pressure even when the exact microstates are unknown.

However, in the context of quantum mechanics, classical probability fails to capture the depth of indeterminacy. A quantum particle in a superposed state is not in a definite, hidden configuration that the observer is ignorant of. Rather, before measurement, the system genuinely lacks a single determinate value for the observable. When we assign probabilities to measurement outcomes, such as $|\alpha|^2$ and $|\beta|^2$ in a qubit state, these do not represent ignorance about an underlying truth but reflect the structure of potentiality itself. This ontological indeterminacy cannot be fully represented within the Kolmogorov framework, which presumes the existence of well-defined states. Thus, while probability theory is indispensable for predicting quantum outcomes, it is conceptually inadequate for modelling the undefined pre-measurement reality of quantum systems.

3.2. Fuzzy Logic

To move beyond the rigid binaries of classical logic, fuzzy logic was introduced by Lotfi Zadeh in 1965 [26],[27]. Fuzzy logic allows propositions to hold partial truth values, ranging continuously between 0 and 1. For example, the statement "the temperature is hot" may be 0.7 true and 0.3 false. This framework has been particularly effective in modelling vagueness and imprecision, especially in fields like natural language processing, control systems, and approximate reasoning.

Fuzzy sets generalise classical sets by assigning each element a membership function that indicates the degree to which the element belongs to the set. This allows for smooth representation of boundaries, in contrast to the sharp distinctions of classical sets. In practical applications, fuzzy logic has proven extremely useful, offering robust tools for systems where data is imprecise or linguistic categories are subjective.

Yet, fuzzy logic remains insufficient for quantum uncertainty. The key limitation is that fuzzy logic presupposes a complementarity between truth and falsity: the more a statement is true, the less it is false, and vice versa. Even though the values need not be binary, they are still constrained within a single continuum. Quantum mechanics, however, presents situations where a proposition is not merely partially true and partially false but also fundamentally indeterminate, a state not captured by fuzzy membership functions. For example, in a quantum superposition, the proposition "the particle is in state A" cannot be described simply as **0.5** true and **0.5** false, because the system is not in any definite state until measurement. There is a dimension of undefinedness that fuzzy logic does not accommodate.

In short, fuzzy logic is well-suited for vagueness but not for indeterminacy. Vagueness arises when boundaries are unclear, as in linguistic or perceptual categories. Indeterminacy, by contrast, arises when the state itself does not exist until defined by interaction. Quantum mechanics belongs to the latter case, requiring a framework that explicitly incorporates undefinedness as an independent category.

3.3. Intuitionistic Fuzzy Logic

To address some of the shortcomings of fuzzy logic, intuitionistic fuzzy logic was introduced by Krassimir Atanassov in the 1980s [28]. Intuitionistic fuzzy sets extend fuzzy sets by including three components: a degree of membership (truth), a degree of non-membership (falsity), and a degree of hesitation (incompleteness). The hesitation component accounts for the fact that in many cases, the sum of membership and non-membership does not fully capture the situation; there may be uncertainty about the degree to which an element belongs to a set.

This framework represents a significant improvement over classical fuzzy logic, as it acknowledges that information can be incomplete or conflicting. Intuitionistic fuzzy logic has found applications in decision-making, risk analysis, and multi-criteria evaluation, where hesitation is a natural part of the process.

Nevertheless, intuitionistic fuzzy logic still interprets hesitation in an epistemic sense. That is, hesitation reflects the observer's lack of knowledge about the system, not the system's inherent indeterminacy. In other words, hesitation exists because the available information is incomplete, not because the state itself is undefined. In quantum mechanics, however, the indeterminacy is not epistemic but ontological. The system itself lacks a definite state until measured. No matter how much additional information one might acquire, the pre-measurement state remains fundamentally indeterminate. Intuitionistic fuzzy logic, therefore, remains inadequate for capturing the essence of quantum uncertainty.

4. Neutrosophy as a Novel Approach to Uncertainty

Neutrosophy was conceived as a philosophical and logical framework to deal with incomplete, contradictory, and indeterminate information**.** The term derives from *neuter* (neutral, neither one extreme nor the other) and *sophia* (wisdom), reflecting its goal of transcending binary oppositions such as truth/falsehood.

At its core, neutrosophy asserts that every proposition or state is characterised by three independent degrees**:**

Truth (T)**:** the extent to which the proposition is true.

Falsity (F)**:** the extent to which the proposition is false.

Indeterminacy (I)**:** the extent to which the proposition is indeterminate, undefined, or undecidable.

Each of these values is expressed within the real standard or non-standard interval **[0, 1]**, although Smarandache's generalisation even allows them to take values beyond conventional bounds (e.g., "overtruth" > **1**, or "underfalsity" < **0**) to model paradoxical or extreme situations. Unlike in probability or fuzzy logic, the sum of T, **I**, and **F** is not required to equal 1. This absence of a normalisation constraint reflects the recognition that truth, falsity, and indeterminacy can coexist in complex ways.

This independence of components is crucial. For instance, a statement may simultaneously be **0.7** true, **0.4** false and **0.2** indeterminate. In such a case, the values overlap, capturing contradictions and incompleteness that classical, fuzzy, or intuitionistic logics cannot adequately represent.

In philosophy, this move corresponds to the recognition that human reasoning and natural systems often exhibit paradoxes, uncertainties, and gaps. In physics, it provides a formal tool to model quantum superposition, collapse and entanglement**,** where truth values are neither fully determined nor mutually exclusive.

4.1. Mathematical Representation

Formally, neutrosophic sets generalise fuzzy and intuitionistic fuzzy sets by introducing three membership functions instead of one or two. Let **U** be a universe of discourse. A neutrosophic set '**A**' in '**U**' is defined as in Equation 4:

$$A = \{\langle x, T_A(x), I_A(x), F_A(x)\rangle : x \in U\} \quad (4)$$

Where:

$T_A(x)$ is the degree of truth of x belonging to **A**,

$I_A(x)$ is the degree of indeterminacy of x belonging to **A** and

$F_A(x)$ is the degree of falsity of x belonging to **A**.

Each of these functions' maps from U to the interval [0,1]. Importantly, there is no requirement that $T_A(x) + I_A(x) + F_A(x) = 1;$ instead, we only have the condition as in Equation 5:

$$0 \le T_A(x), I_A(x), F_A(x) \le 1 \quad (5)$$

This flexibility makes neutrosophic sets capable of representing contradictions and overlapping states.

Example:

Consider a qubit in the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $|\alpha|^2 + |\beta|^2 = 1$. If we ask whether the qubit is in state $|0\rangle$, a neutrosophic model might assign as in Equation 6:

$$T = |\alpha|^2 \ \ \& \ F = |\beta|^2 \quad (6)$$

I = a nonzero value representing the indeterminacy of the system before measurement.

Unlike probability, which forces all uncertainty into a normalised distribution between outcomes, neutrosophy captures the indeterminacy dimension that persists until measurement collapses the state.

5. Neutrosophic Representation of Quantum Systems

Building on the foundations of neutrosophy, we now illustrate how quantum systems can be represented using neutrosophic logic and sets**.** This approach treats **T, F** and **I** as independent components, providing a framework for modelling quantum states that are fundamentally uncertain before measurement. We discuss three central phenomena: superposition, entanglement and measurement collapse, demonstrating how neutrosophic logic captures their intrinsic indeterminacy.

5.1. Superposition States

Consider a single qubit in a superposition state in Equation 7:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (7)$$

Where $|\alpha|^2 + |\beta|^2 = 1$ ensures normalisation. Classical probability theory interprets $\boldsymbol{|\alpha|^2}$ and $\boldsymbol{|\beta|^2}$ as the likelihoods of measuring the qubit in states $|0\rangle$ or $|1\rangle$, respectively. While this probabilistic interpretation predicts measurement outcomes, it does not capture the ontological nature of the superposition: before measurement, the qubit is not in any definite state and its "truth value" relative to a particular basis is undefined.

Neutrosophically, the pre-measurement state can be represented as a triplet (**T**, **I**, **F**) for each basis state. For example, with respect to $\boldsymbol{|0\rangle}$ as mentioned in 8:

$$T(|0\rangle) = |\alpha|^2,\ F(|0\rangle) = |\beta|^2,\ I(|0\rangle) \ne 0 \quad (8)$$

Here, $\boldsymbol{T(|0\rangle)}$ represents the degree to which the qubit can be considered true in state $\boldsymbol{|0\rangle}$, $\boldsymbol{F(|0\rangle)}$ the degree to which it is false, and $\boldsymbol{I(|0\rangle)}$ represents the residual indeterminacy inherent to the superposition. This framework explicitly accounts for the fact that before measurement, the qubit is neither purely $|0\rangle$; it exists in a state that is fundamentally undefined with respect to classical truth assignments.

By modelling superposition in this way, neutrosophy captures both the probabilistic and indeterminate aspects of quantum states. Unlike probability distributions, which collapse uncertainty into likelihoods alone, the neutrosophic triplet preserves the ontological indeterminacy that is essential to quantum mechanics. This allows for a richer and more accurate representation of the system before observation.

5.2. Entangled Systems

Entanglement presents a particularly challenging case for classical and even probabilistic frameworks, as the states of individual particles cannot be described independently. Consider the Bell state as shown in Equation 9:

$$|\Phi^+\rangle = \frac{1}{2}(|00\rangle + |11\rangle) \quad (9)$$

In this entangled state, neither qubit possesses a definite individual state before measurement. The measurement of one qubit instantaneously determines the state of the other, regardless of spatial separation, highlighting the nonlocal correlations that are a hallmark of quantum mechanics.

A neutrosophic approach models entanglement by assigning correlated truth, falsity and indeterminacy values to each particle. For the first qubit, we might define as shown in Equation 10:

$$T(|0\rangle_1) = 0.5,\ F(|0\rangle_1) = 0.5,\ I(|0\rangle_1) \neq 0 \quad (10)$$

The second qubit exhibits a similar triplet. Crucially, the indeterminacy components $\mathbf{I(|0\rangle_1)}$ and $\mathbf{I(|0\rangle_2)}$ are linked, reflecting the shared unresolved state of the entangled pair. Only when a measurement is performed does the indeterminacy resolve, simultaneously collapsing the state of both qubits.

Neutrosophic representation is particularly powerful here because it accommodates the simultaneous existence of partial truth, falsity and indeterminacy while maintaining correlations between entangled subsystems. Classical probability can capture outcome correlations but cannot represent the pre-measurement indeterminacy; fuzzy or intuitionistic fuzzy logic captures partial truth but either conflates indeterminacy with lack of knowledge or restricts the independence of **T**, **F** and **I**. Neutrosophy, by contrast, naturally expresses the indeterminate yet correlated nature of entangled states.

5.3. Measurement Collapse

The process of measurement in quantum mechanics is often described as wavefunction collapse marks the transition from indeterminate superpositions to definite outcomes. In neutrosophic terms, this can be interpreted as a reduction of indeterminacy accompanied by a corresponding increase in truth or falsity as mentioned in Equation 11:

$$I \to 0,\ T \to 1 \text{ or } F \to 1 \quad (11)$$

For instance, measuring a qubit in the $|0\rangle/|1\rangle$ basis results in one outcome being fully realised ($\boldsymbol{T} = 1$) and the other fully falsified ($\boldsymbol{F} = 1$), with the indeterminacy $\boldsymbol{I}$ dropping to zero. This perspective allows collapse to be understood not as a sudden elimination of possibilities but as the resolution of intrinsic indeterminacy into a determinate state.

This interpretation offers several conceptual advantages:

- **Continuity of Modelling:** Unlike classical logic, which cannot accommodate the intermediate superposed state, neutrosophy provides a smooth transition from indeterminacy to determinacy.
- **Explicit Representation of Quantum Features:** Indeterminacy $\boldsymbol{I}$ is treated as a real, independent component rather than a surrogate for ignorance.
- **Compatibility with Probabilities:** The degrees of truth and falsity remain aligned with measurement probabilities, ensuring that the neutrosophic model is physically meaningful.

Applied to entangled systems, measurement collapse is represented as a simultaneous resolution of indeterminacy across correlated particles, preserving the nonlocal correlations characteristic of quantum mechanics. This unified framework thus integrates probabilistic outcomes, ontological indeterminacy, and correlation structure in a single, coherent model.

6. Broader Role of Neutrosophy in Quantum Mechanics

The application of neutrosophy extends beyond formal modelling of quantum states; it also provides a conceptual and mathematical framework that can inform interpretations of quantum mechanics, quantum decision theory and quantum information science. By explicitly incorporating indeterminacy as a distinct dimension, neutrosophic logic offers new ways of understanding, reasoning, and engineering in domains where uncertainty is fundamental.

6.1. Interpretations of Quantum Theory

Quantum mechanics has long been accompanied by philosophical debates regarding the nature of reality, measurement, and determinacy. Different interpretations propose varying mechanisms for the emergence of definite outcomes, each of which can be enriched by a neutrosophic perspective. In the Copenhagen interpretation, the act of measurement causes the wavefunction to collapse, producing a single realised state. Neutrosophy provides a natural formalism for this process: before measurement, the quantum system is represented with significant $\mathbf{I} \neq 0$, and the collapse corresponds to a reduction of **I** to zero, with **T** or **F** attaining a definitive value.

The Many-Worlds interpretation envisions all potential outcomes as actualised across branching universes. In neutrosophic terms, each branch can be associated with distinct truth and falsity values, while global indeterminacy persists across the multiverse. This allows for a quantitative representation of coexistence and uncertainty across parallel branches, reflecting the simultaneous realisation of multiple possibilities without collapsing indeterminacy in any absolute sense.

For objective collapse theories, which propose that indeterminacy resolves dynamically over time due to intrinsic physical processes, neutrosophic logic can model the gradual reduction of **I** while updating **T** and **F** values correspondingly. In each interpretation, neutrosophy offers a formal mechanism for tracking the evolution of indeterminacy, providing a bridge between abstract philosophical ideas and precise mathematical representation.

6.2. Quantum Decision Theory

Quantum systems are increasingly studied in the context of decision-making under uncertainty, particularly in quantum game theory, cognitive modelling, and adaptive algorithms. Traditional probabilistic models often suffice for calculating expected outcomes, but they do not capture the fundamental indeterminacy that may influence strategies or behaviours. Neutrosophic logic enriches this domain by representing not only the likelihood of different outcomes but also the intrinsic indeterminacy associated with each option.

For instance, in quantum games, a player's decision can be affected by superposed or entangled states. Modelling these states neutrosophically allows for a three-dimensional assessment of each strategy: the degree to which it is advantageous, disadvantageous, and indeterminate before measurement or outcome realisation. This richer representation can enhance the formulation of quantum decision rules, improve predictions of behaviour in quantum-influenced environments, and provide insights into cognitive processes that exploit quantum-like uncertainty. By incorporating indeterminacy as an explicit parameter, neutrosophy enables decision-theoretic models that are more aligned with the ontological nature of quantum uncertainty.

6.3. Quantum Information Science

Quantum information science relies fundamentally on managing uncertain and superposed states. Applications such as quantum error correction, cryptography, and algorithm design require precise handling of probabilistic outcomes while simultaneously accounting for indeterminate aspects of qubits or quantum registers. Neutrosophic logic offers a potentially valuable tool for these tasks by providing a structured way to quantify and manipulate truth, falsity and indeterminacy.

In quantum error correction, for example, neutrosophic sets can represent not only the probability of error occurrence but also the residual indeterminacy of a qubit's state after partial correction. In cryptographic protocols, such as quantum key distribution, neutrosophy can model the indeterminate state of eavesdropping attempts or uncertainties in entangled qubits, providing a more nuanced measure of security risk. Likewise, in algorithm design, incorporating indeterminacy explicitly can aid in optimising operations on superposed or entangled registers, enhancing both efficiency and fault tolerance. Across these applications, neutrosophy provides a formal and computationally tractable way to integrate fundamental quantum uncertainty into the design, analysis, and implementation of quantum technologies.

7. Challenges and Future Directions

While neutrosophic logic provides a powerful framework for representing quantum uncertainty, its integration into mainstream quantum mechanics and experimental practice faces several challenges. Addressing these issues is essential for advancing both the theoretical foundations and practical applications of neutrosophy in quantum science.

Formal Integration with Hilbert Spaces: Quantum mechanics is rigorously formulated within the mathematical structure of **Hilbert spaces**, where states are represented by vectors and observables by linear operators. To fully incorporate neutrosophic logic, it is necessary to develop a formal mapping between neutrosophic sets and the Hilbert space formalism. This involves defining how the neutrosophic components **T**, **F** and **I** correspond to vector amplitudes, density matrices, or probability distributions derived from quantum states. A successful integration would allow neutrosophy to complement the existing formalism without violating fundamental principles such as unitarity, superposition, and entanglement. Such a formalism could also enable computational implementations in quantum algorithms, simulations, and error-corrected systems, bridging the gap between abstract neutrosophic reasoning and standard quantum mechanics.

Operational Meaning of Indeterminacy: While neutrosophy provides a philosophically rich notion of indeterminacy, assigning it a precise operational meaning remains an open challenge. To be useful in physical models, indeterminacy must be connected to measurable or calculable quantities. Potential avenues include linking **I** to decoherence rates, quantum entropy or the contextuality of observables. For example, indeterminacy could be formalised as a function of the coherence length in a superposed state or the entropy associated with entangled subsystems. Establishing such links would not only clarify the physical interpretation of **I** but also allow neutrosophic models to produce experimentally testable predictions, thereby grounding philosophical constructs in quantitative science.

Experimental Relevance: A critical step for the broader adoption of neutrosophy in quantum mechanics is demonstrating its empirical relevance. Experimental protocols such as weak measurements, quantum tomography or delayed-choice experiments offer promising platforms for testing neutrosophic predictions. For instance, weak measurement techniques allow partial information to be extracted from quantum systems without full collapse, providing a natural context for measuring residual indeterminacy. Similarly, delayed-choice experiments could reveal how pre-measurement indeterminacy influences correlations in entangled systems. By designing experiments specifically to probe the independent components **T**, **F**, and **I**, researchers can validate the predictive power of neutrosophic models and potentially uncover new quantum phenomena that conventional probabilistic or fuzzy frameworks cannot fully capture.

7. Conclusion

Our results demonstrate that neutrosophic logic provides a fundamentally new language for describing quantum systems, one that directly encodes ontological indeterminacy rather than reducing it to ignorance or vagueness. By applying neutrosophy to superposition, entanglement, and wavefunction collapse, we establish a framework that complements Hilbert space formalism while extending its interpretive power. This perspective has implications not only for the foundations of quantum theory but also for the design of algorithms, error correction, and decision-making in quantum information science. Future integration of neutrosophic models with experimental platforms such as weak measurements and quantum tomography could yield testable predictions that distinguish neutrosophic representations from classical probabilistic approaches. More broadly, embracing indeterminacy as an explicit dimension may reshape how uncertainty is modeled across physics, mathematics, and computation, offering a step toward a more complete understanding of quantum reality.

Declaration

Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Issues: There are no ethical issues. All data in this paper is publicly available.

Author Contribution Statement: A.F and H.K.M conceived idea and designed the research; Analyzed interpreted the data and wrote the paper.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Consent to Publish: All authors are agreed to publish version of the manuscript in this journal.

References

1. Berta, Mario, et al. "The uncertainty principle in the presence of quantum memory." Nature Physics 6.9 (2010): 659-662.
2. Zhao, Yuan-Yuan, et al. "Experimental study of quantum uncertainty from lack of information." npj Quantum Information 8.1 (2022): 64.
3. Oppenheim, Jonathan, and Stephanie Wehner. "The uncertainty principle determines the nonlocality of quantum mechanics." Science 330.6007 (2010): 1072-1074.
4. Sen, Debashis. "The uncertainty relations in quantum mechanics." Current Science (2014): 203-218.
5. Ozawa, Masanao. "Quantum limits of measurements and uncertainty principle." Quantum Aspects of Optical Communications: Proceedings of a Workshop Held at the CNRS, Paris, France 26–28 November 1990. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.
6. Cho, Adrian. "Furtive approach rolls back the limits of quantum uncertainty." (2011): 690-693.
7. Afzal, Usama. "Estimate calculation of discrete energy levels in the hydrogen atom using an indeterministic method." High Energy Density Physics (2025): 101192.
8. Fatima, Adeena, et al. "A Comprehensive Review of Neutrosophic Statistics for Data Analysis in Applied Sciences." Journal of Reliability and Statistical Studies (2025): 25-40.
9. Smarandache, Florentin, and Maissam Jdid. "An overview of neutrosophic and plithogenic theories and applications." (2023).
10. Sarkar, Debasmita, and Pankaj Kumar Srivastava. "Recent development and applications of neutrosophic fuzzy optimization approach." International Journal of System Assurance Engineering and Management 15.6 (2024): 2042-2066.
11. Çevik, A., Topal, S., & Smarandache, F. (2018). Neutrosophic Logic Based Quantum Computing. Symmetry, 10(11), 656. https://doi.org/10.3390/sym10110656

12. Smarandache, Florentin. "Neutrosophic Quantum Theory: Partial Entanglement, Partial Effect of the Observer, and Teleportation." Neutrosophic Sets and Systems 86.1 (2025): 1.
13. Smarandache, Florentin. An introduction to the Neutrosophic probability applied in quantum physics. Infinite Study, 2000.
14. Baghaturia, Iuri, et al. "CRITICAL REVIEW OF FUNDAMENTAL CONCEPTS IN PHYSICS Part 5–“Quantum Superposition”." GSAR J Math Sci. 2025b 4.8: 83-90.
15. Lahiri, Anuradha, Prodyot Kumar Roy, and Bijan Bagchi. "Supersymmetry in quantum mechanics." International Journal of Modern Physics A 5.08 (1990): 1383-1456.
16. Hsu, Stephen DH. "The measure problem in no-collapse (many worlds) quantum mechanics." International Journal of Modern Physics D 26.03 (2017): 1730008.
17. Jordan, Andrew N., and Alexander N. Korotkov. "Uncollapsing the wavefunction by undoing quantum measurements." Contemporary Physics 51.2 (2010): 125-147.
18. Carlesso, Matteo, et al. "Present status and future challenges of non-interferometric tests of collapse models." Nature Physics 18.3 (2022): 243-250.
19. Zhang, Zheshen, et al. "Entanglement-based quantum information technology: a tutorial." Advances in Optics and Photonics 16.1 (2024): 60-162.
20. Yu, Yue. "Advancements in applications of quantum entanglement." Journal of Physics: Conference Series. Vol. 2012. No. 1. IOP Publishing, 2021.
21. Busch, Paul, Teiko Heinonen, and Pekka Lahti. "Heisenberg's uncertainty principle." Physics reports 452.6 (2007): 155-176.
22. Coles, Patrick J., et al. "Entropic uncertainty relations and their applications." Reviews of Modern Physics 89.1 (2017): 015002.
23. Aristarhov, Serj. "Heisenberg’s uncertainty principle and particle trajectories." Foundations of Physics 53.1 (2023): 7.
24. Rau, Jochen. "On quantum vs. classical probability." Annals of Physics 324.12 (2009): 2622-2637.
25. Halliwell, Joe, and Qiang Shen. "Linguistic probabilities: theory and application." Soft Computing 13.2 (2009): 169-183.
26. Kindo, Abdoul Azize, et al. "Fuzzy logic approach for knowledge modeling in an Ontology: A review." 2020 IEEE 2nd International Conference on Smart Cities and Communities (SCCIC). IEEE, 2020.
27. Ruspini, Enrique H., James C. Bezdek, and James M. Keller. "Fuzzy clustering: A historical perspective." IEEE Computational Intelligence Magazine 14.1 (2019): 45-55.
28. Melliani, Said, and Oscar Castillo, eds. "Recent advances in intuitionistic fuzzy logic systems and mathematics." (2021).

Journal of Engineering, Science and Technological Trends (JESTT)

Vol. 2, Issue 3, November 2025
DOI: https://doi.org/10.64060/jestt.v2i3.4

Perspective

Assessment and Critical Review of $PM_{0.1}$ Pollution in Pakistan

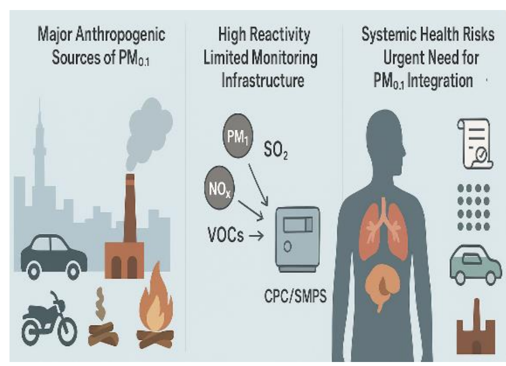
Faiza Irfan[1], Rida Munir[2], Sadique Akbar[1] and Ramna Zafar[1]*

[1]Department of Physics, University of Education, Township, Lahore 54570, Pakistan
[2]Canterbury Christ Church University, Canterbury, England
* Corresponding Email: ramnazafar96@gmail.com (R. Zafar)

Received: 31 July 2025 / Revised: 17 October 2025 / Accepted: 19 October 2025 / Published online: 15 November 2025

ABSTRACT

Particulate matter (PM) is among the most significant air pollutants globally, with severe implications for environmental integrity, human health, and climate stability. Among its various fractions, ultrafine particles ($PM_{0.1}$, particles with an aerodynamic diameter ≤ 0.1 μm) are gaining increasing attention due to their high surface area-to-mass ratio, deep pulmonary penetration, and potential to translocate into systemic circulation and vital organs. This paper presents a comprehensive critical review of $PM_{0.1}$ assessment in Pakistan, emphasising its sources, spatiotemporal distribution, measurement limitations, and health consequences. Despite the mounting evidence of air quality degradation in Pakistan, data on $PM_{0.1}$ remain scarce and fragmented. The few available studies indicate that urban centres such as Lahore, Karachi, Islamabad, and Faisalabad exhibit ultrafine particle concentrations substantially higher than international safety benchmarks. Anthropogenic activities, including vehicular emissions, industrial combustion, biomass burning, and construction dust, are dominant contributors. This review identifies key gaps in current research, highlighting the lack of long-term monitoring, standardised methodologies, and toxicological assessments specific to $PM_{0.1}$ exposure in local populations. It further stresses the urgent need for policy integration, investment in high-resolution monitoring technologies, and public health interventions. Overall, the assessment underscores that $PM_{0.1}$ pollution in Pakistan poses an emerging environmental health crisis that remains scientifically underexplored and administratively underprioritized.



Keywords*:* Air pollution; Air quality monitoring; Combustion emissions; Environmental policy; Health effects; Ultrafine particles

1. Introduction

Air pollution has emerged as a global environmental concern, and particulate matter (PM) is recognised as a leading cause of premature mortality, chronic respiratory disease, and reduced life expectancy [1],[2],[3],[4]. In Pakistan, where industrialisation, rapid urbanisation, and energy consumption patterns have intensified over the past two decades, particulate pollution has reached critical levels [5],[6],[7]. The country consistently ranks among the most polluted nations globally, with urban air quality indices (AQI) frequently exceeding the World Health Organization (WHO) guidelines several times. While coarse (PM_{10}) and fine ($PM_{2.5}$) particles have been the primary focus of air quality monitoring, the ultrafine fraction ($PM_{0.1}$) remains largely overlooked in Pakistan's environmental assessment frameworks [8].

$PM_{0.1}$ particles, owing to their nanoscale dimensions and physicochemical reactivity, differ fundamentally from larger PM fractions. They can penetrate deep into the alveolar region of the lungs, cross biological membranes, and even reach the bloodstream and brain through systemic circulation [9],[10],[11]. Consequently, their potential for causing oxidative stress, inflammation, and multi-organ toxicity is far greater than that of coarser particulates. Yet, Pakistan's monitoring networks and academic research have not adequately addressed this pollutant category, leaving significant knowledge gaps regarding its prevalence, sources, and health effects .

This study aims to critically evaluate existing assessments of $PM_{0.1}$ in Pakistan, synthesise findings from national and international research, and highlight the methodological, policy, and public health challenges associated with this pollutant. By analysing the limited but growing body of evidence, this review intends to (i) identify dominant emission sources and exposure hotspots, (ii) assess health implications and potential societal burdens, (iii) examine deficiencies in monitoring infrastructure, and (iv) propose strategies for future research and policy action. The broader purpose is to underscore the urgency of integrating ultrafine particulate pollution into Pakistan's air quality management and public health planning.

2. Critical Analysis

$PM_{0.1}$ particles, often referred to as ultrafine particles (UFPs), constitute a distinct class of aerosols characterised by their small size, large surface area, and high number concentration. These particles are primarily generated through combustion-related processes, both anthropogenic and natural, rather than mechanical disintegration [10],[12]. Their aerodynamic diameter (<0.1 μm) allows them to remain suspended for long periods and facilitates deep penetration into human respiratory systems. Because $PM_{0.1}$ particles can adsorb heavy metals, volatile organic compounds (VOCs), and polycyclic aromatic hydrocarbons (PAHs), their toxic potential is amplified compared to coarser PM fractions.

In the South Asian context, where population density and energy consumption are high, $PM_{0.1}$ represents a largely invisible yet potent public health threat [13]. The absence of $PM_{0.1}$ in most air quality standards, including Pakistan's National Environmental Quality Standards (NEQS), further complicates the issue, as policymakers and researchers rely primarily on $PM_{2.5}$ as a proxy indicator. However, recent studies demonstrate that $PM_{2.5}$ mass concentrations do not reliably reflect $PM_{0.1}$ number concentrations, making it essential to address UFPs independently.

2.1. Sources of $PM_{0.1}$ in Pakistan

$PM_{0.1}$ emissions in Pakistan predominantly arise from anthropogenic activities [6],[7]. Vehicular exhaust, particularly from diesel engines and two-stroke motorcycles, is the single largest contributor in urban areas. The lack of emission control technologies, poor fuel quality, and ageing vehicle fleets exacerbate ultrafine particle emissions. Karachi and Lahore, both megacities with dense traffic networks, exhibit consistently high ultrafine particle number concentrations near major roadways and intersections.

Industrial processes, including steel manufacturing, brick kilns, cement production, and power generation, represent another significant source. Many industries in Pakistan operate on outdated combustion systems, using low-grade fuels such as furnace oil and coal, which emit substantial quantities of nucleation-mode particles. Brick kilns, especially those employing traditional clamp or Bull's trench technologies, release ultrafine carbonaceous soot mixed with trace metals.

Biomass burning, both intentional (crop residue burning) and domestic (firewood, charcoal, and dung combustion), adds another layer of complexity. Rural communities depend heavily on biomass for cooking and heating, leading to elevated $PM_{0.1}$ exposures indoors. During post-harvest seasons, open burning of agricultural residues in Punjab and Sindh contributes significantly to regional haze events and cross-border pollution episodes.

Additionally, secondary formation of $PM_{0.1}$ through atmospheric reactions involving sulfur dioxide (SO_2), nitrogen oxides (NO_x), ammonia (NH_3), and volatile organics plays a considerable role. The combination of high solar radiation, stagnant air conditions, and urban heat islands enhances these photochemical processes, particularly in Lahore and Islamabad.

2.2. Assessment and Measurement Limitations

The scientific assessment of $PM_{0.1}$ in Pakistan remains highly limited due to the absence of dedicated monitoring networks and standardised measurement protocols [14]. While $PM_{2.5}$ is monitored intermittently by the Pakistan Environmental Protection Agency (Pak-EPA) and some universities, ultrafine particle number concentrations are rarely reported [15],[7]. Only a few research studies, mostly pilot investigations, have used condensation particle counters (CPCs) or scanning mobility particle sizers (SMPS) in Karachi, Lahore, and Islamabad to characterise $PM_{0.1}$ levels.

Results from these isolated studies indicate that $PM_{0.1}$ concentrations in Pakistan's urban areas often exceed those observed in developed countries by an order of magnitude. For example, mean UFP concentrations of $1–3×10^5$ particles/cm^3 have been reported in traffic-dominated zones of Lahore, while background levels remain above $5×10^4$ particles/cm^3 even during low-traffic hours [16]. However, data remain fragmented, short-term, and largely non-comparable due to inconsistencies in sampling duration, instrumentation, and meteorological adjustments.

Another critical limitation lies in the lack of chemical characterisation. Most studies measure total particle number or mass without identifying elemental composition or toxicity profiles. Consequently, the relative contributions of metals, organic carbon, and secondary aerosols remain uncertain. Without such data, it is challenging to develop source apportionment models or to evaluate specific mitigation strategies.

2.3. Health Effects of $PM_{0.1}$ Exposure

The health effects of $PM_{0.1}$ are increasingly recognised as severe, with both epidemiological and toxicological studies linking exposure to cardiovascular, respiratory, and neurological disorders. Due to their nanoscale size, $PM_{0.1}$ particles can bypass the mucociliary clearance system, penetrate alveolar membranes, and enter systemic circulation. Once in the bloodstream, they induce oxidative stress, inflammation, and endothelial dysfunction, leading to hypertension, atherosclerosis, and other cardiovascular complications [2],[6].

Inhalation of ultrafine particles has also been associated with reduced lung function, asthma exacerbation, and chronic obstructive pulmonary disease (COPD). Neurological studies suggest potential associations between $PM_{0.1}$ exposure and cognitive decline, neurodegenerative disorders, and altered neurotransmitter regulation. Maternal exposure may affect fetal development, contributing to low birth weight and developmental impairments.

In Pakistan, where baseline exposure to ambient air pollution is already high, these health effects are likely exacerbated. However, no large-scale epidemiological studies have specifically evaluated $PM_{0.1}$-related health outcomes. Public health assessments continue to rely on $PM_{2.5}$-based dose-response relationships, which may underestimate the actual risk. The absence of hospital-based exposure tracking, combined with inadequate medical data integration, hinders the quantification of health burdens attributable to ultrafine particles.

2.4. Policy, Research and Institutional Gaps

Pakistan's air quality management system faces multiple structural deficiencies. The National Environmental Quality Standards (NEQS) currently define permissible limits for PM_{10} and $PM_{2.5}$ only, with no reference to ultrafine fractions [17]. This omission stems from limited scientific awareness and technical capacity. Moreover, real-time monitoring infrastructure remains inadequate few continuous air monitoring stations existing, and most lack the sensitivity to detect $PM_{0.1}$.

At the institutional level, coordination among federal, provincial and municipal agencies is weak. While the Clean Air Program and National Climate Change Policy acknowledge particulate pollution, they do not specifically address $PM_{0.1}$. Academic research on ultrafine particles is restricted to small-scale university projects, often unsupported by long-term funding or national datasets. International collaborations remain sporadic, and there is no centralised database for air quality data integration.

2.4. Toward an Integrated Approach

To address $PM_{0.1}$ pollution effectively, Pakistan must adopt a multi-pronged approach integrating technology, policy, and public engagement. Establishing a national ultrafine particle monitoring network using advanced instruments such as CPCs, SMPS, and aerosol mass spectrometers is essential. Data from such networks should inform the revision of NEQS to include $PM_{0.1}$ concentration limits and exposure guidelines.

On the mitigation front, strict vehicular emission standards, phasing out of old vehicles, and promotion of electric mobility could substantially reduce UFP emissions. Industrial modernisation, particularly the replacement of traditional brick kilns with zigzag technology and the adoption of cleaner fuels, should be prioritised. Rural energy transitions such as biogas and solar-based cooking systems would alleviate domestic ultrafine particle exposure.

Furthermore, public awareness campaigns and health advisories should communicate the risks associated with ultrafine particles. Academic institutions should be encouraged to develop interdisciplinary research programs that link environmental science, toxicology, and epidemiology. International partnerships with institutions experienced in UFP assessment (e.g., in Japan, Europe, and the U.S.) could accelerate knowledge transfer and technical training.

3. Conclusion

$PM_{0.1}$ pollution in Pakistan represents a critical yet largely overlooked environmental and public health issue. The limited evidence available indicates that ultrafine particle concentrations in major urban centres such as Lahore, Karachi, and Islamabad far exceed global safety limits, driven by vehicular emissions, industrial combustion, and biomass burning. Owing to their nanoscale size and high surface reactivity, these particles can penetrate deep into the lungs and enter systemic circulation, causing oxidative stress, inflammation, and multi-organ toxicity. However, the absence of standardised monitoring systems, comprehensive chemical characterisation, and health-focused studies severely restricts understanding of their true impact. To mitigate this emerging threat, Pakistan must urgently integrate $PM_{0.1}$ into its national air quality standards, establish advanced monitoring infrastructure, and promote cleaner technologies across transportation, industry, and domestic sectors. A coordinated approach combining scientific research, policy reform, and public awareness is essential to protect population health and ensure sustainable environmental management.

Declaration

Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Issues: There are no ethical issues. All data in this paper is publicly available.

Author Contribution Statement: F.I., R.A., S.A. and R.Z conceived the idea and designed the research; Analysed and interpreted the data and wrote the paper.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Consent to Publish: All authors agree to publish the version of the manuscript in this journal.

References

1. Farahani, Vahid Jalali, et al. "The oxidative potential of particulate matter (PM) in different regions around the world and its relation to air pollution sources." Environmental Science: Atmospheres 2.5 (2022): 1076-1086.
2. Daellenbach, Kaspar R., et al. "Sources of particulate-matter air pollution and its oxidative potential in Europe." Nature 587.7834 (2020): 414-419.
3. Li, Chi, et al. "Reversal of trends in global fine particulate matter air pollution." Nature communications 14.1 (2023): 5349.
4. Brook, Robert D., and Sanjay Rajagopalan. "Particulate matter, air pollution, and blood pressure." Journal of the American Society of Hypertension 3.5 (2009): 332-350.
5. Arif, Muhammad Maaz, Muhammad Abu Bakar, and Sajid Rashid Ahmad. "Impacts of air pollution on human health in Lahore city, Pakistan: A review." Global Research Journal of Natural Science and Technology 1.2 (2023).
6. Nasar-u-Minallah, Muhammad, et al. "Assessing and anticipating environmental challenges in Lahore, Pakistan: future implications of air pollution on sustainable development and environmental governance." Environmental Monitoring and Assessment 196.9 (2024): 865.
7. Shi, Yuan, et al. "Urbanization and regional air pollution across South Asian developing countries–A nationwide land use regression for ambient PM2. 5 assessment in Pakistan." Environmental Pollution 266 (2020): 115145.
8. Ahmad, Shafiq, et al. "Morphological, mineralogical, and biochemical characteristics of particulate matter in three size fractions (PM10, PM2. 5, and PM1) in the urban environment." ACS omega 8.35 (2023): 31661-31674.
9. Benabed, Ahmed, and Amir Boulbair. "PM10, PM2. 5, PM1, and PM0. 1 resuspension due to human walking." Air Quality, Atmosphere & Health 15.9 (2022): 1547-1556.
10. Schraufnagel, Dean E. "The health effects of ultrafine particles." Experimental & molecular medicine 52.3 (2020): 311-317.
11. Phairuang, Worradorn, et al. "Characterizing chemical, environmental, and stimulated subcellular physical characteristics of Size-Fractionated PMs down to PM0. 1." Environmental Science & Technology 58.28 (2024): 12368-12378.
12. Traboulsi, Hussein, et al. "Inhaled pollutants: the molecular scene behind respiratory and systemic diseases associated with ultrafine particulate matter." International journal of molecular sciences 18.2 (2017): 243.
13. Phairuang, Worradorn, et al. "Airborne nanoparticles (PM0. 1) in southeast Asian cities: a review." Sustainability 14.16 (2022): 10074.
14. Sharma, R., et al. "Health implications due to exposure to fine and ultrafine particulate matters: A short review." International Journal of Ambient Energy 45.1 (2024): 2314256.
15. Shabbir, Muneeba, et al. "A paradigm shift: Low-cost sensors for effective air quality monitoring and management in developing countries." Environment International (2025): 109521.
16. Wangchuk, Tenzin, et al. "Mobile assessment of on-road air pollution and its sources along the East–West Highway in Bhutan." Atmospheric Environment 118 (2015): 98-106.
17. Bachmann, John. "Will the circle be unbroken: a history of the US National Ambient Air Quality Standards." Journal of the Air & Waste Management Association 57.6 (2007): 652-697.

Journal of Engineering, Science and Technological Trends (JESTT)

Vol. 2, Issue 3, November 2025
DOI: https://doi.org/10.64060/jestt.v2i3.5

Perspective

AI/ML-Driven Design and Optimisation of Quantum Dots: A Perspective Toward Intelligent Materials Discovery

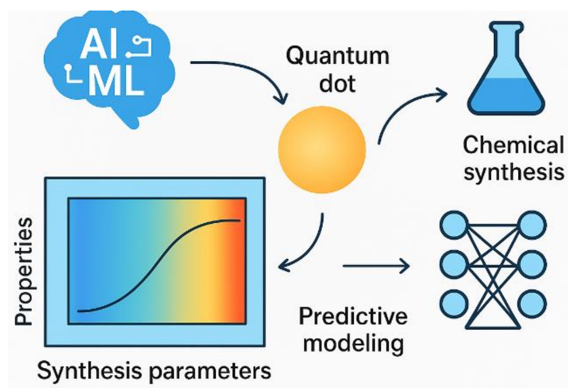
Asma Batool[1]* and Tanzeela Urooj[2]

[1]Faculty of Chemistry, Islamic Azad University Tehran, Iran
[2]Deakin University - Australia
* Corresponding Email: asmahfd1994@gmail.com (A. Batool)

Received: 01 July 2025 / Revised: 29 September 2025 / Accepted: 10 October 2025 / Published online: 15 November 2025

ABSTRACT

Quantum dots (QDs), nanoscale semiconductors with size-dependent and tunable optoelectronic properties, are central to next-generation technologies spanning displays, photovoltaics, bioimaging, and quantum information systems. However, their synthesis and optimisation remain challenging due to the intricate interplay of reaction parameters and nonlinear physicochemical interactions. The integration of artificial intelligence (AI) and machine learning (ML) is redefining this landscape, enabling predictive design, autonomous synthesis control, and accelerated discovery across the QD domain. This Perspective highlights the conceptual advances and methodological innovations driving AI/ML-assisted QD research, emphasising achievements in data-driven modelling, synthesis optimisation, and materials informatics. Persistent challenges, including data scarcity, model transparency, and limited generalizability, are critically examined, alongside emerging strategies toward physics-informed and autonomous discovery frameworks. We propose that the convergence of intelligent algorithms and human expertise will catalyse a paradigm shift from empirical experimentation toward rational, self-evolving materials design in quantum dot science.



Keywords*:* Artificial intelligence (AI); Autonomous synthesis; Data-driven modelling; Intelligent materials discovery; Machine learning (ML); Materials informatics; Optoelectronic properties; Physics-informed learning; Predictive design; Quantum dots (QDs); Synthesis optimization.

1. Introduction

Quantum dots (QDs) represent one of the most profound advancements in nanomaterials science since the discovery of fullerenes and carbon nanotubes [1],[2],[3],[4],[5]. Their exceptional optical tunability, stemming from quantum confinement effects, enables precise control over absorption and emission spectra through manipulation of particle size, shape, and composition. This unique property has revolutionised technologies such as quantum dot light-emitting diodes (QLEDs), high-efficiency solar cells, and nanoscale biosensing systems [6],[7],[8]. However, despite their immense potential, achieving consistent, high-quality QDs remains a scientific and engineering challenge. Their optical and electronic behaviour is governed by a complex interplay of synthesis variables, including precursor ratios, ligand chemistry, solvent polarity, temperature, and nucleation kinetics. Small perturbations in any of these parameters can result in significant variations in size distribution, defect density, and surface passivation, ultimately influencing emission efficiency, stability, and colour purity.

Traditional synthesis strategies have largely relied on chemical intuition, empirical optimisation, and iterative experimentation. While these methods have yielded critical insights, they are inherently limited in scope and efficiency [9],[10]. The multidimensional parameter space governing QD synthesis is vast and nonlinear, making exhaustive exploration through manual experimentation infeasible. Moreover, the lack of standardised protocols and reproducibility between laboratories hinders the translation of empirical results into predictive design rules. This bottleneck has motivated the scientific community to explore artificial intelligence (AI) and machine learning (ML) as transformative approaches for accelerating QD discovery. By analysing large experimental and computational datasets, AI/ML algorithms can identify hidden correlations, construct predictive models for synthesis–structure–property relationships, and autonomously guide optimisation toward desired optoelectronic outcomes.

The integration of AI/ML into QD research marks a paradigm shift from descriptive to predictive materials science [11],[12]. Supervised learning has been employed to forecast band gaps, emission wavelengths, and quantum yields, while unsupervised models have been used to classify synthesis regimes and identify latent material families. Reinforcement learning and Bayesian optimisation techniques enable adaptive experimentation, allowing algorithms to iteratively refine synthesis parameters in real time. These advances have laid the foundation for autonomous laboratories and self-driving experimental platforms that combine robotic synthesis with AI-driven decision-making. Such systems can perform hundreds of synthesis and characterisation cycles with minimal human input, optimising nanomaterial properties far more efficiently than traditional methods. This convergence of automation, computation, and data science is redefining the pace and precision of nanomaterials research.

Nonetheless, several critical challenges persist in realising the full promise of AI/ML-driven QD design. Data scarcity and inconsistency remain major obstacles, as most published datasets are fragmented, unstandardized, and lack essential metadata describing reaction dynamics or surface chemistry. Additionally, the "black box" nature of deep learning models limits interpretability, hindering the extraction of mechanistic insights crucial for scientific understanding. To address these issues, emerging frameworks such as physics-informed machine learning and hybrid AI–quantum mechanical modelling are being developed to merge predictive accuracy with physical transparency. As the field progresses, a synergistic integration of computational intelligence, experimental automation, and human expertise is expected to define the next frontier of QD science, transitioning from empirical discovery toward intelligent, self-evolving materials engineering.

Herein, the study is designed to explore and critically analyse the transformative integration of AI and ML in the rational design, synthesis, and optimisation of QDs. The primary objective is to elucidate how data-driven methodologies can unravel the complex, nonlinear relationships linking synthesis parameters, structural attributes, and optoelectronic performance. This work aims to assess recent advances in predictive modelling, autonomous experimentation, and materials informatics that collectively redefine the landscape of QD research. Furthermore, the study seeks to identify existing challenges, such as data scarcity, model interpretability, and limited cross-domain generalisation that impede widespread adoption of AI/ML frameworks. Ultimately, the study is structured to propose a forward-looking vision wherein physics-informed, human-guided machine learning enables intelligent, autonomous, and sustainable discovery pathways for next-generation quantum dot materials.

2. AI/ML Integration in Quantum Dot Science

The integration of AI/ML into QD research marks a transformative step toward intelligent materials design. By leveraging data-driven algorithms capable of capturing complex, nonlinear relationships between synthesis parameters and material properties, AI/ML approaches overcome the inefficiencies of traditional trial-and-error experimentation. These techniques enable predictive modelling, rapid optimisation, and even autonomous discovery of QDs with tailored optoelectronic characteristics. Through supervised learning, reinforcement learning, and Bayesian optimisation, AI can not only forecast outcomes such as emission wavelength or quantum yield but also guide experimental systems toward optimal synthesis conditions in real time. As a result, the convergence of AI/ML with nanochemistry is accelerating the transition from empirical exploration to intelligent, closed-loop materials engineering, heralding a new era of automated, data-informed innovation in quantum dot science.

The Rationale for AI/ML in Materials Design: The discovery and optimisation of advanced materials are increasingly driven by data-intensive methodologies that transcend conventional experimental paradigms [13]. In traditional synthesis workflows, even modest variations in precursor ratios, ligand chemistry, or temperature profiles can necessitate hundreds of iterative experiments. For QDs, whose properties depend on a highly multidimensional and nonlinear parameter space, such trial-and-error approaches are inherently inefficient and often fail to capture the intricate correlations governing material behaviour. AI/ML introduce a fundamentally different design philosophy, one that leverages computational intelligence to uncover hidden patterns within complex datasets. Algorithms such as random forests, gradient-boosted decision trees, support vector machines (SVMs), and deep neural networks have demonstrated the capacity to learn nonlinear relationships between synthesis parameters (e.g., precursor identity, solvent polarity, ligand functionality) and resultant material properties (e.g., bandgap, emission wavelength, quantum yield, or carrier lifetime). Once trained, these models can rapidly predict QD properties for untested combinations of variables, dramatically reducing the experimental burden. Moreover, when integrated with Bayesian optimisation or reinforcement learning frameworks, AI systems can autonomously explore and refine synthesis pathways through iterative, closed-loop feedback cycles, thereby transforming materials discovery into a self-optimising process.

Data Generation and Representation: Data serves as the cornerstone of all AI/ML-driven materials research [14]. In the context of quantum dot science, data sources are diverse, encompassing experimental measurements, computational simulations, and process-level metadata. Experimental datasets typically include optical absorption and emission spectra, X-ray diffraction (XRD) patterns, transmission electron microscopy (TEM) images, and photoluminescence lifetimes, while computational data may derive from density functional theory (DFT) calculations or molecular dynamics simulations [15]. Process data such as temperature profiles, solvent polarity, and precursor concentrations to further enrich the dataset, providing contextual insight into reaction dynamics. Transforming these heterogeneous data streams into machine-readable formats requires robust feature engineering. Ligand-related descriptors, for instance, can encode steric and electronic effects, while reaction conditions may be normalised or represented through thermodynamic parameters. More recently, representation learning approaches, such as graph neural networks (GNNs), have gained traction for their ability to capture atomic connectivity and electronic structure directly from raw data, reducing reliance on manual feature construction. Despite these advances, data scarcity and inconsistency remain pervasive challenges. The majority of QD-related datasets are small, fragmented, and non-standardised, which undermines the generalizability and transferability of trained models. Establishing community-wide data standards and open repositories is, therefore, a crucial prerequisite for the maturation of AI/ML-enabled QD research.

Predictive Modelling of Quantum Dot Properties: Machine learning models have achieved considerable success in predicting and optimising quantum dot properties by correlating synthesis parameters with optoelectronic performance [16]. For example, random forest models have been employed to predict the emission wavelength and quantum yield of CdSe and InP QDs, yielding mean absolute errors below 10 nm. Bayesian optimisation frameworks have efficiently identified optimal precursor ratios and reaction temperatures for PbS and $CsPbBr_3$ QDs, enhancing particle monodispersity and photoluminescence quantum yield (PLQY). Deep neural networks trained on spectroscopic data have accurately classified QD compositions, defect states, and emission behaviours, providing rapid, data-driven insight into synthesis outcomes. Such predictive frameworks have demonstrated an order-of-magnitude reduction in experimental iterations compared with conventional manual optimisation. Importantly, these approaches shift the paradigm from empirical trial-and-error to hypothesis-free exploration, where relationships between structure and function emerge through statistical inference rather than deterministic modelling. This transition signifies the growing maturity of data-centric materials science, one where discovery is increasingly guided by algorithmic reasoning rather than intuition alone.

Autonomous and Closed-Loop Systems: The integration of AI/ML with robotic and microfluidic experimental platforms represents a defining frontier in quantum dot research. Automated synthesis systems equipped with precise control over flow rates, injection timing, and temperature gradients can execute hundreds of reactions per day, generating consistent and high-resolution datasets [11]. When combined with real-time spectroscopic monitoring, ML algorithms can dynamically adjust reaction parameters to steer synthesis toward pre-defined performance objectives such as emission wavelength, colour purity, or PLQY. Pioneering demonstrations of such self-driving laboratories have successfully achieved autonomous optimisation of perovskite and CdSe QD syntheses. In these systems, reinforcement learning algorithms continuously refine synthesis conditions based on reward functions tied to optical performance metrics, enabling adaptive improvement across successive experiments. This closed-loop paradigm transcends traditional human-in-the-loop workflows. AI not only interprets experimental data but also generates hypotheses, tests them autonomously, and refines its predictive understanding through iterative learning. As this integration matures, autonomous materials discovery is poised to become a cornerstone of next-generation nanoscience, where intelligent systems accelerate innovation through continuous, data-driven experimentation.

3. Critical Analysis of Current Limitations

Despite significant progress, the integration of AI/ML in QD research continues to face systemic barriers that constrain predictive reliability, interpretability, and scalability. While data-centric approaches have accelerated the pace of discovery, fundamental challenges persist in data quality, model transparency, transferability, and reproducibility. A critical evaluation of these issues is essential to guide the next phase of AI/ML-driven QD innovation, as shown in Figure 1.

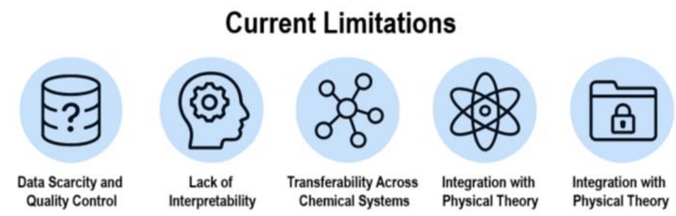


Figure 1: *Limitations in AI/ML-Driven Design and Optimisation of Quantum Dots.*

3.1. Data Scarcity and Quality Control

Reliable and generalizable ML models require large, diverse, and high-fidelity datasets, yet the datasets currently available in QD research are often limited, fragmented, and poorly standardised. Most published datasets contain fewer than a few hundred samples, insufficient for robust statistical learning. Compounding this issue, experimental protocols vary considerably across laboratories, with inconsistencies in reporting key variables such as ligand purity, precursor concentration, ramp rates, and solvent composition. These variations hinder model transferability and cross-study validation.

An additional challenge arises from publication bias: the literature overwhelmingly emphasises successful syntheses while omitting failed or inconclusive experiments. Such selective reporting produces data imbalance, leading ML models to overfit and capture coincidental correlations rather than causal mechanisms. The absence of "negative" examples ultimately compromises predictive robustness. To address these deficiencies, the research community must prioritise data standardisation and transparency, implementing machine-readable metadata formats, uniform experimental documentation, and open-access repositories. Establishing such infrastructure will enable reproducible, transferable, and statistically meaningful AI/ML applications in QD science.

3.2. Lack of Interpretability

Although deep learning architectures and ensemble models have achieved high predictive accuracy in QD research, they often function as **black boxes**, offering limited mechanistic insight into the underlying physicochemical processes. This opacity conflicts with the foundational scientific goal of understanding *why* specific synthesis parameters yield certain outcomes. The inability to trace predictions back to chemical intuition restricts trust and hinders the broader adoption of AI-driven insights in experimental practice.

To mitigate this challenge, interpretable AI frameworks are increasingly being introduced. Techniques such as SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-Agnostic Explanations) and attention-based neural networks can quantify feature importance and provide local explanations for model outputs. These methods facilitate partial mechanistic understanding by revealing which synthesis variables, such as ligand type or temperature, most strongly influence emission wavelength or quantum yield. Nonetheless, achieving true mechanistic interpretability, where predictive models yield chemically meaningful explanations aligned with reaction kinetics or thermodynamics, remains a critical and unresolved frontier.

3.3. Transferability Across Chemical Systems

Another fundamental limitation lies in the restricted transferability of ML models across different QD compositions and chemical systems. Models trained on one material family (e.g., CdSe QDs) often fail to generalise to others (e.g., perovskite, carbon, or III–V semiconductor QDs) due to intrinsic differences in bonding, crystal structure, and defect chemistry. Consequently, feature–property relationships are frequently system-specific, and retraining is required for each new QD class. This lack of cross-domain adaptability limits AI's ability to accelerate the discovery of entirely new materials.

Emerging solutions include transfer learning, wherein knowledge gained from one system is adapted to another through re-training on smaller datasets, and multi-fidelity modelling, which integrates experimental and simulated data at varying levels of accuracy. Furthermore, hybrid AI–physics approaches that couple ML models with quantum mechanical or thermodynamic constraints show great promise. By grounding predictions in physical theory, such frameworks can enhance both generalisation and scientific interpretability, moving the field closer to universal predictive models for QD behaviour.

3.4. Integration with Physical Theory

Purely data-driven methodologies, while powerful, can occasionally produce predictions that defy fundamental chemical or physical principles. To ensure physical plausibility, researchers are increasingly incorporating domain knowledge directly into ML architectures through physics-informed machine learning (PIML). These models embed constraints such as conservation laws, thermodynamic boundaries, and quantum-mechanical relationships into the training process. The result is a synthesis of empirical learning and theoretical understanding that enhances both predictive accuracy and scientific validity. By uniting data-centric and physics-based reasoning, PIML represents a critical step toward bridging the long-standing divide between empirical modelling and first-principles theory in QD science.

3.5. Reproducibility and Transparency

Reproducibility remains a cornerstone of scientific credibility, yet it continues to be a major concern in AI/ML-based QD research. Many studies fail to provide comprehensive documentation of model architectures, hyperparameter configurations, training datasets, or pre-processing protocols, making independent replication difficult. Moreover, the lack of open-source access to data and code limits transparency and hinders collective progress.

To foster a culture of reproducibility and openness, the QD community must align with the FAIR data principles, ensuring that datasets and models are *Findable, Accessible, Interoperable,* and *Reusable*. Publicly available repositories, standardised benchmarking datasets, and shared codebases will be essential to validate AI/ML findings and build cumulative knowledge. Only through transparent, reproducible research practices can AI truly mature from a computational tool into a reliable scientific partner in the discovery and design of next-generation quantum dot materials.

4. The Emerging Landscape and Future Outlook

The future of AI/ML-driven quantum dots research lies in the establishment of standardised, open, and interoperable data ecosystems that promote reproducibility and cross-disciplinary collaboration. Building comprehensive databases that include both successful and failed syntheses, complete with detailed metadata and standardised descriptors, will enable statistically robust model training and benchmarking. Leveraging natural language processing (NLP) for automated data extraction from the scientific literature can further expand datasets and minimise manual curation. Concurrently, the integration of physics-informed and explainable AI architectures represents a transformative step toward interpretable and physically grounded predictions. Physics-informed neural networks (PINNs), graph neural networks (GNNs), and hybrid ML–DFT models can incorporate fundamental laws of chemistry and thermodynamics, ensuring both accuracy and chemical plausibility. Explainable AI frameworks will further elucidate the governing parameters that influence QD synthesis and performance, allowing researchers to move from empirical prediction toward mechanistic understanding.

In parallel, the rise of autonomous and self-learning laboratories promises to redefine materials discovery by coupling robotic synthesis, in situ characterisation, and reinforcement learning into closed-loop experimental systems capable of continuous self-improvement. These platforms can explore vast chemical spaces, iteratively refine synthesis conditions, and dramatically accelerate optimisation, transforming months of manual experimentation into days. Furthermore, sustainability will emerge as a guiding principle, as AI enables multi-objective optimisation to balance optoelectronic performance with environmental responsibility, identifying lead-free or cadmium-free alternatives. The integration of QD-level predictive models with device-level simulations spanning QLEDs, photovoltaics, and sensors will bridge the gap between nanoscale design and macroscopic performance, ushering in an era of intelligent, sustainable, and fully automated quantum dot engineering.

4. Conclusion

The integration of artificial intelligence and machine learning with quantum-dot science has progressed from proof-of-concept regressions to autonomous, closed-loop platforms that can optimize emission colour within a single daylight cycle. Nonetheless, the field remains in an adolescent state: models are accurate yet fragile, datasets are growing yet biased, and synthesis successes are reported while thousands of unwritten failures accumulate in laboratory notebooks. Realizing the full potential of AI-driven quantum-dot discovery will require community adoption of standardized, FAIR experimental logs that record not only the triumphant synthesis but also the silent, informative failures; interpretable algorithms that map latent variables to chemically meaningful descriptors such as ligand steric constants or surface reaction enthalpies; and physics-informed architectures that embed thermodynamic conservation laws and quantum-mechanical boundary conditions directly into the loss function. When these elements converge, the next decade could witness self-driving laboratories that explore lead-free, earth-abundant compositions while simultaneously optimizing device-level efficiency, colour purity and environmental sustainability. Such autonomous systems will not replace human creativity they will amplify it, freeing researchers to pose bolder questions about quantum confinement, energy transfer and emergent nanoscale phenomena. The future of quantum dots, therefore, lies at the nexus of human curiosity and machine persistence, where intelligent algorithms become the invisible yet indispensable reagents of every successful synthesis.

Declaration

Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Issues: There are no ethical issues. All data in this paper is publicly available.

Author Contribution Statement: A.B and T.U conceived the idea and designed the research; Analysed and interpreted the data and wrote the paper.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Consent to Publish: All authors agree to publish the version of the manuscript in this journal.

Data Availability Statement: Data will be made available on request from the corresponding author.

References

1. Georgakilas, Vasilios, et al. "Broad family of carbon nanoallotropes: classification, chemistry, and applications of fullerenes, carbon dots, nanotubes, graphene, nanodiamonds, and combined superstructures." Chemical reviews 115.11 (2015): 4744-4822.
2. García de Arquer, F. Pelayo, et al. "Semiconductor quantum dots: Technological progress and future challenges." Science 373.6555 (2021): eaaz8541.
3. Kovalenko, Maksym V. "Opportunities and challenges for quantum dot photovoltaics." Nature Nanotechnology 10.12 (2015): 994-997.
4. Resch-Genger, Ute, et al. "Quantum dots versus organic dyes as fluorescent labels." Nature methods 5.9 (2008): 763-775.
5. Li, Xu-Bing, Chen-Ho Tung, and Li-Zhu Wu. "Semiconducting quantum dots for artificial photosynthesis." Nature Reviews Chemistry 2.8 (2018): 160-173.
6. Medintz, Igor L., et al. "Quantum dot bioconjugates for imaging, labelling and sensing." Nature materials 4.6 (2005): 435-446.
7. Matea, Cristian T., et al. "Quantum dots in imaging, drug delivery and sensor applications." International journal of nanomedicine (2017): 5421-5431.
8. Kargozar, Saeid, et al. "Quantum dots: a review from concept to clinic." Biotechnology Journal 15.12 (2020): 2000117.
9. Mohamed, Walied AA, et al. "Quantum dots synthetization and future prospect applications." Nanotechnology Reviews 10.1 (2021): 1926-1940.
10. Agarwal, Kushagra, Himanshu Rai, and Sandip Mondal. "Quantum dots: an overview of synthesis, properties, and applications." Materials Research Express 10.6 (2023): 062001.
11. Chugh, Vibhas, et al. "Employing nano-enabled artificial intelligence (AI)-based smart technologies for prediction, screening, and detection of cancer." Nanoscale 16.11 (2024): 5458-5486.
12. Diao, Shanhui, et al. "From synthesis to properties: expanding the horizons of machine learning in nanomaterials research." Materials Horizons 12.12 (2025): 4133-4164.
13. Guo, Kai, et al. "Artificial intelligence and machine learning in design of mechanical materials." Materials Horizons 8.4 (2021): 1153-1172.
14. Dunjko, Vedran, and Hans J. Briegel. "Machine learning\& artificial intelligence in the quantum domain." arXiv preprint arXiv:1709.02779 (2017).
15. Haram, Santosh K., et al. "Quantum confinement in CdTe quantum dots: investigation through cyclic voltammetry supported by density functional theory (DFT)." The Journal of Physical Chemistry C 115.14 (2011): 6243-6249.
16. Duman, Ali Nabi, and Almaz S. Jalilov. "Machine learning for carbon dot synthesis and applications." Materials Advances 5.18 (2024): 7097-7112.