ISBN: 978-93-47587-99-3

CONTEMPORARY INNOVATIONS IN SCIENCE, ENGINEERING AND TECHNOLOGY

EDITORS:
DR. SUKANTA DAS
DR. SHWETA TANWAR
MR. HUSSAIN M.
DR. R. KARTHICK MANOJ

Bhumi Publishing, India

First Edition: January 2026

Contemporary Innovations in Science, Engineering and Technology

(ISBN: 978-93-47587-99-3)

DOI: https://doi.org/10.5281/zenodo.18449095

Editors

Dr. Sukanta Das	Dr. Shweta Tanwar
Department of Veterinary Anatomy, College of Veterinary Sciences and Animal Husbandry, R.K. Nagar, Tripura West	Medical Scientist and Technical Editor, The Indian Journal of Medical Research, ICMR Headquarters, New Delhi
Mr. Hussain M.	**Dr. R. Karthick Manoj**
Department of Physics, Maharashtra College of Arts, Science and Commerce, Nagpada, Mumbai, M.S.	Department of Electrical and Electronics Engineering, AMET Deemed to be University, Chennai, Tamil Nadu

Bhumi Publishing

January 2026

Title: Contemporary Innovations in Science, Engineering and Technology

Editors: Dr. Sukanta Das, Dr. Shweta Tanwar, Mr. Hussain M., Dr. R. Karthick Manoj

First Edition: January 2026

ISBN: 978-93-47587-99-3

ISBN 978-93-475-8799-3

9 789347 587993

DOI: https://doi.org/10.5281/zenodo.18449095

Published by Bhumi Publishing,

a publishing unit of Bhumi Gramin Vikas Sanstha

Bhumi Publishing

Nigave Khalasa, Tal – Karveer, Dist – Kolhapur, Maharashtra, INDIA 416 207

E-mail: bhumipublishing@gmail.com

PREFACE

The rapid pace of scientific discovery and technological advancement has fundamentally reshaped the way society's function, industries evolve, and knowledge is generated. In this context, *Contemporary Innovations in Science, Engineering and Technology* has been conceived as a comprehensive volume that captures the dynamic and interdisciplinary nature of modern research and development. The book brings together diverse perspectives that reflect how innovation today emerges at the intersection of multiple disciplines rather than within isolated domains.

This volume aims to provide readers with a balanced blend of theoretical foundations, experimental investigations, and applied technologies that address real-world challenges. The chapters included in this book explore recent advancements across core areas of science, cutting-edge engineering solutions, and transformative technologies that are redefining productivity, sustainability, and quality of life. Emphasis has been placed on contemporary themes such as smart systems, advanced materials, computational approaches, automation, artificial intelligence, energy-efficient technologies, and sustainable engineering practices.

The contributors to this book are academicians, researchers, and industry professionals who have shared their expertise and insights based on rigorous research and practical experience. Each chapter has been carefully reviewed to ensure clarity, originality, and relevance, making the content accessible to postgraduate students, research scholars, educators, and practicing engineers. The interdisciplinary approach adopted in this volume encourages readers to think beyond conventional boundaries and to appreciate the collaborative nature of modern innovation.

Contemporary Innovations in Science, Engineering and Technology is not only intended to serve as a reference source but also as a platform to inspire future research and innovation. By highlighting emerging trends, novel methodologies, and practical applications, the book seeks to stimulate critical thinking and foster a deeper understanding of how scientific and technological advancements can contribute to societal and industrial progress.

We sincerely hope that this volume will prove valuable to the academic community and industry alike, and that it will encourage readers to engage actively with innovation-driven research, paving the way for sustainable and technologically advanced solutions for the future.

- Editors

TABLE OF CONTENT

ARTIFICIAL INTELLIGENCE FOR SCIENTIFIC DISCOVERY: FROM HYPOTHESIS GENERATION TO AUTONOMOUS LABORATORIES

Rajesh Kumar Mishra[1], Divyansh Mishra[2] and Rekha Agarwal[3]

[1]ICFRE-Tropical Forest Research Institute

(Ministry of Environment, Forests & Climate Change, Govt. of India),

P.O. RFRC, Mandla Road, Jabalpur, MP-482021, India

[2]Department of Artificial Intelligence and Data Science,

Jabalpur Engineering College, Jabalpur (MP)

[3]Government Science College, Jabalpur, MP, India- 482 001

Corresponding author E-mail: rajeshkmishra20@gmail.com, rajesh.mishra0701@gov.in, divyanshspps@gmail.com, rekhasciencecollege@gmail.com

Abstract:

Artificial Intelligence (AI) has transitioned from a peripheral analytical tool to a central methodological engine in scientific discovery and engineering innovation. Advances in machine learning, deep learning, and generative modeling now enable AI systems to formulate hypotheses, design and execute experiments autonomously, construct high-fidelity digital twins, and support real-time decision-making in complex cyber-physical systems. This chapter presents a comprehensive and integrated treatment of AI-driven scientific discovery and intelligent engineering systems. It examines the full lifecycle—from hypothesis generation and self-driving laboratories to AI-enabled digital twins, generative design, edge intelligence, and federated learning—while addressing explainability, trustworthiness, governance, and sustainability. Emphasis is placed on hybrid physics–AI models, uncertainty-aware learning, and assurance-driven deployment in safety-critical domains. The chapter concludes with reference architecture, research challenges, and future directions that position AI as a foundational infrastructure for next-generation science and engineering.

Keywords: Artificial Intelligence for Science, Autonomous Laboratories, Digital Twins, Generative AI, Explainable AI, Edge AI, Federated Learning, Sustainable AI

Introduction:

Scientific and engineering discovery has traditionally relied on human intuition, domain expertise, and iterative experimentation. While computational tools have long assisted analysis and simulation, the emergence of modern AI marks a structural shift: intelligent systems can now actively participate in hypothesis formulation, experimental design, execution, and interpretation. This transformation is particularly evident in data-intensive domains such as materials science, manufacturing, energy systems, and infrastructure engineering. Artificial Intelligence (AI) has transcended from being a theoretical concept to a cornerstone of technological advancement. The integration of AI across industries demonstrates its potential to revolutionize processes, systems, and services (Mishra *et al.,* 2024). Artificial Intelligence (AI) has witnessed rapid advancements in recent years, transforming various sectors by enhancing efficiency, automating tasks, and enabling more intelligent decision-making processes (Mishra *et al.,* 2025a; Mishra *et al.,* 2025b; Mishra *et al.,* 2025c; Mishra *et al.,* 2025d; Mishra *et al.,* 2025e; Mishra *et al.,* 2025f; Mishra *et al.,* 2025g; Mishra *et al.,* 2025h; Mishra *et al.,* 2025i). In sum, AI represents the intelligent layer of the materials science trinity. It transforms how knowledge is generated, validated, and applied — bridging microscopic theory and macroscopic application through automation, reasoning, and prediction. As the third pillar, AI not only accelerates discovery but redefines the scientific method itself, heralding a new era of autonomous, explainable, and intelligent materials innovation. AI-for-Science represents a convergence of machine learning, automation, robotics, and domain theory, enabling closed-loop discovery pipelines that dramatically compress innovation timelines. Rather than replacing scientists and engineers, AI systems act as amplifiers of human capability—exploring vast design spaces, uncovering non-obvious relationships, and supporting decisions under uncertainty.

Artificial Intelligence for Scientific Discovery

Artificial Intelligence (AI) for scientific discovery represents a fundamental paradigm shift in the way knowledge is generated, validated, and refined. Traditionally, scientific discovery has relied heavily on human intuition, domain expertise, and iterative experimentation, all of which are constrained by time, cost, and cognitive limitations. AI augments—and in some cases partially automates—this process by learning from large, heterogeneous datasets, uncovering latent patterns, generating testable hypotheses, and optimizing experimental and analytical workflows. A defining feature of AI-driven discovery is closed-loop intelligence, where models not only analyze data but also propose, evaluate, and iteratively refine scientific actions. Artificial Intelligence (AI) is rapidly transforming global production systems, learning environments, and scientific research practices. Although individual AI applications in industry, education, and research have been widely explored, an integrated perspective that unifies these three pillars of societal and knowledge development remains largely absent from the literature (Mishra *et al.,* 2025j).

At its core, AI-for-Science combines data-driven inference with theory-driven constraints. Unlike purely statistical approaches, scientifically meaningful AI systems must adhere to physical laws, conservation principles, causality, and experimental feasibility. This has encouraged the emergence of hybrid paradigms that integrate machine learning with symbolic reasoning, physics-based simulations, and probabilistic inference. Artificial Intelligence (AI), rooted in computer science, mathematics, cognitive science, and engineering, has evolved

(ISBN: 978-93-47587-99-3)

through foundational paradigms such as symbolic AI, connectionism, and statistical learning, shaped by Turing's computational theory and the rise of machine learning (Mishra *et al.,* 2025k). Today, AI and intelligent systems constitute one of the most transformative technological forces of the twenty-first century, reshaping scientific discovery, industrial processes, governance, and human–machine interaction (Mishra *et al.,* 2025l).

Hypothesis Generation and Knowledge Synthesis

One of the most transformative roles of AI is in hypothesis generation. Large-scale text mining, natural language processing, and knowledge graph construction allow AI systems to synthesize decades of scientific literature, experimental databases, and simulation outputs. These systems can identify non-obvious correlations, propose mechanistic explanations, and highlight unexplored regions of parameter space. Modern approaches treat hypotheses as probabilistic models rather than deterministic statements. Bayesian frameworks allow AI to assign confidence levels to competing hypotheses, update beliefs as new evidence arrives, and explicitly manage uncertainty. Causal inference methods further enhance this capability by distinguishing correlation from causation—an essential requirement for scientific validity.

Importantly, AI-generated hypotheses are increasingly actionable: they are formulated in a way that maps directly to experimental variables, measurement strategies, and feasibility constraints, enabling seamless integration into experimental workflows.

Active Learning and Experimental Design

Scientific discovery is fundamentally constrained by experimental resources. AI addresses this bottleneck through active learning and Bayesian optimization, which prioritize experiments that are expected to yield the highest information gain. Rather than exhaustively exploring all possibilities, AI strategically selects experiments that maximize learning efficiency.

In practice, this involves:

- Building surrogate models that approximate complex physical or chemical systems,
- Quantifying predictive uncertainty,
- Selecting experiments that reduce uncertainty or optimize target objectives under constraints.

This approach has demonstrated the ability to reduce experimental effort by orders of magnitude in domains such as materials discovery, chemical synthesis, and biological screening. The scientific implication is profound: discovery becomes adaptive and data-efficient, rather than brute-force.

Autonomous and Self-Driving Laboratories

The convergence of AI, robotics, and automation has given rise to self-driving laboratories (SDLs)—systems capable of autonomously executing the scientific method. In an SDL, AI algorithms design experiments, robotic platforms execute them, sensors collect data, and models update themselves in real time, closing the loop without continuous human intervention.

A typical SDL architecture includes:

- Robotic experiment execution and sample handling,
- Automated characterization and data acquisition,
- AI-driven planning, optimization, and interpretation,
- Human-in-the-loop oversight for validation and safety.

These systems are particularly impactful in fields where experiments are slow, expensive, or hazardous. From a scientific perspective, SDLs transform discovery from a linear process into a continuous, adaptive feedback system, dramatically accelerating innovation cycles.

Integration with Simulation and Theory

AI does not replace theory or simulation; rather, it complements them. Physics-based models provide interpretability and extrapolation capability, while AI provides scalability and adaptability. Physics-informed neural networks (PINNs), neural operators, and hybrid surrogate models embed governing equations directly into learning architectures, ensuring consistency with known laws.

This integration enables:

- Rapid approximation of high-fidelity simulations,
- Real-time inference for complex systems,
- Discovery of effective reduced-order models.

Such hybrid approaches are essential for scientific credibility, particularly when AI models are deployed beyond the range of available data.

Reliability, Reproducibility, and Scientific Trust

Despite its promise, AI-driven discovery introduces new risks. Model bias, data leakage, distribution shift, and automation-induced error propagation can undermine scientific reliability. Consequently, AI-for-Science must prioritize:

- Uncertainty quantification and confidence calibration,
- Explainability and interpretability of model outputs,
- Rigorous data provenance and reproducibility standards,
- Human oversight and validation checkpoints.

Scientific trust in AI systems is not achieved through performance alone but through transparent, auditable, and reproducible workflows.

Artificial Intelligence is evolving into a foundational infrastructure for science, analogous to the role of computing in the late 20th century. Its most enduring impact will not be isolated breakthroughs, but the transformation of discovery itself—from a slow, intuition-limited process to a scalable, adaptive, and uncertainty-aware enterprise.

In my view, the future of scientific discovery lies in human–AI symbiosis, where AI handles exploration, optimization, and pattern discovery, while humans provide conceptual framing, ethical judgment, and theoretical insight. Success will depend on how effectively AI systems are grounded in scientific principles, governed responsibly, and integrated into the culture of research.

Machine Learning and Deep Learning for Complex Engineering Systems

Machine Learning (ML) and Deep Learning (DL) are now core enablers for modeling, analysis, control, and optimization of complex engineering systems marked by nonlinearity, high dimensionality, uncertainty, and strong subsystem coupling. Unlike traditional physics-based or rule-driven models, ML and DL learn system behavior directly from data, supporting prediction, adaptation, and self-improvement in manufacturing, power grids, transportation, aerospace, and smart infrastructure. Deep architectures, including convolutional, sequence, and graph-based models, effectively capture spatial, temporal, and networked interactions. Hybrid approaches integrating physics-based constraints improve data efficiency, physical consistency, reliability, and suitability for safety-critical, real-world engineering applications and long-term sustainable system-level performance globally applicable.

Architectural Paradigms

The effectiveness of machine learning (ML) and deep learning (DL) in complex engineering systems is strongly influenced by architectural choice, because engineering problems involve structured interactions, temporal evolution, physical constraints, and uncertainty. Modern paradigms therefore introduce inductive biases aligned with system structure. Convolutional Neural Networks (CNNs) address spatial locality and translational invariance in applications such as visual inspection, non-destructive testing, thermal imaging, and infrastructure monitoring, enabling scalable defect detection and pattern recognition. Recurrent Neural Networks, Long Short-Term Memory models, and Transformers capture temporal dynamics in multivariate sensor data and are widely used for predictive maintenance, load forecasting, traffic prediction, and process monitoring, with Transformers excelling at long-horizon dependencies. Graph Neural Networks (GNNs) model inherently networked systems, including power grids, transportation networks, and structural frameworks, by encoding relational and topological information for fault propagation, resilience analysis, and optimization. Physics-Informed Neural Networks and hybrid physics–ML models embed governing equations and constraints, improving data efficiency, generalization, and physical consistency when data are limited. Neural operators and surrogate models learn mappings between function spaces, enabling fast approximation of numerical solvers and real-time digital twins. Finally, probabilistic and uncertainty-aware architectures, such as Bayesian networks and ensembles, quantify predictive confidence, supporting risk-aware decisions, anomaly detection, applications.

Learning for Control and Optimization

Learning for control and optimization is among the most impactful applications of machine learning (ML) and deep learning (DL) in complex engineering systems, where the goal extends beyond prediction to active, constrained decision-making. Engineering control problems involve nonlinear dynamics, partial observability, delays, safety limits, and competing objectives related to performance, cost, energy efficiency, reliability, and sustainability. Classical approaches such as PID control, optimal control, and model predictive control (MPC) are well established and interpretable but rely heavily on accurate system models, which are difficult to maintain in large-scale, time-varying environments. ML-based control mitigates these limitations by learning dynamics, policies, or optimization surrogates directly from data and adapting to evolving conditions.

A key contribution of ML is data-driven system identification, where neural networks, Gaussian processes, and learned state-space models approximate unknown dynamics. These models are often embedded within MPC frameworks, combining adaptability with explicit constraint handling and stability guarantees. Deep learning further enables reinforcement learning (RL), allowing control policies to be learned through interaction in high-dimensional systems such as robotics, autonomous vehicles, and energy systems. To address safety concerns, current research emphasizes constrained and risk-aware RL. Learning-based optimization also benefits engineering practice through surrogate modeling and Bayesian optimization, enabling efficient design-space exploration and uncertainty-aware decision-making. Successful deployment ultimately depends on rigorous validation, integration with control theory, and system-level oversight.

Explainable and Trustworthy AI in Safety-Critical Applications

The deployment of Artificial Intelligence (AI) in safety-critical applications—including aerospace, autonomous and assisted transportation, industrial control, power grids, medical devices, and critical infrastructure—requires reliability, transparency, and accountability far beyond conventional analytics. Failures in these domains can result in loss of life, environmental damage, or economic disruption. Explainability and trustworthiness are therefore core engineering requirements, on par with robustness, fault tolerance, and verification. Explainable AI (XAI) addresses the opacity of high-performing ML and DL models, providing insights into predictions and control actions. Explanation methods include post-hoc local techniques (feature attribution, sensitivity analysis), global interpretability approaches (simplified surrogate models, rule extraction), and intrinsically interpretable architectures aligned with domain concepts.

Safety-critical systems increasingly favor intrinsic or hybrid models for stable, auditable explanations.

Trustworthy AI extends beyond interpretability to robustness, reliability, uncertainty quantification, and governance. Real-world systems face sensor noise, environmental variability, component degradation, and distributional shifts. AI models must express confidence, trigger human-in-the-loop interventions under uncertainty, and support runtime monitoring. Certification requires traceability, validation, version control, and post-deployment oversight. Successful adoption depends on designing AI that is interpretable by design, uncertainty-aware by default, and integrated into safety-engineered processes, ensuring resilient, accountable, and auditable performance in high-stakes environments.

Explainability as an Engineering Requirement

In safety-critical and high-assurance engineering systems, explainability must be regarded as a formal engineering requirement, on par with reliability, robustness, maintainability, and safety certification. In domains such as aerospace control, nuclear power, autonomous transportation, medical diagnostics, and industrial automation, AI outputs directly influence decisions with potentially irreversible consequences. A model lacking explainability cannot be fully validated, certified, or trusted, regardless of empirical performance. Explainability ensures traceability, enabling engineers to map system-level requirements to model behavior and decision logic, identify root causes, and perform systematic debugging and lifecycle maintenance.

Explainable AI is essential for safety assurance and regulatory compliance, as auditors must verify that automated systems operate predictably within defined safety envelopes. Interpretable models facilitate risk assessment and verification under rare or extreme conditions. Additionally, explainability strengthens human–AI collaboration, allowing operators to understand and interrogate AI decisions during abnormal or emergency events, thereby improving situational awareness and preventing misuse.

Designing AI for explainability at development time—rather than relying on post-hoc explanations—encourages architectures that are interpretable by design, incorporate physical or logical constraints, and expose meaningful internal representations. The maturation of AI in safety-critical engineering depends on this shift, ensuring AI systems are accountable, auditable, and fit for purpose throughout their operational lifecycle.

Explainability Techniques

Explainability techniques form the methodological foundation for rendering Artificial Intelligence (AI) systems transparent, interpretable, and auditable, especially in safety-critical and high-assurance engineering applications. These techniques aim to clarify how and why a model produces a given decision, enabling engineers, operators, and regulators to assess correctness, reliability, and compliance. Explainability is best treated as a toolbox rather than a single method, with approaches chosen according to system requirements, risk levels, and lifecycle stage.

Post-hoc local explanations, such as gradient-based saliency, SHAP, and LIME, provide insight into individual predictions by highlighting the influence of input variables, supporting debugging, fault investigation, and operator decision-making. Global explanation methods, including rule extraction, surrogate models, monotonic constraints, and partial dependence analysis, reveal systematic patterns and dominant decision rules, aiding validation, safety assessment, and regulatory audits. Intrinsic interpretability embeds explainability directly in model architectures, as seen in linear and generalized additive models, attention-based networks, concept bottleneck models, and physics-informed neural networks, aligning internal representations with human-understandable operational or physical concepts. Counterfactual explanations provide actionable "what-if" insights, while uncertainty-aware and causal methods communicate reliability and causal mechanisms.

For complex, safety-critical systems, layered strategies combining local and global explanations, intrinsic interpretability, uncertainty quantification, and human oversight are essential. Explainability must be rigorously selected, tested, and documented, ensuring AI systems remain transparent, auditable, and trustworthy throughout their operational lifecycle.

AI-Driven Digital Twins for Smart Manufacturing and Infrastructure

AI-driven digital twins have become a cornerstone of smart manufacturing and modern infrastructure, enabling real-time synchronization between physical assets and their virtual counterparts. Unlike static simulations, a digital twin is a dynamic, data-driven model that evolves continuously via sensor inputs, operational logs, and historical knowledge. When integrated with Artificial Intelligence (AI), digital twins move beyond monitoring to decision support and autonomous optimization, predicting system behavior, detecting faults, and recommending or executing control actions.

In manufacturing, AI-enhanced twins aggregate data from machines, production lines, quality systems, and enterprise platforms to build multi-scale operational models. Machine learning and deep learning support predictive maintenance, remaining useful life estimation, and early fault detection by capturing complex degradation patterns. AI also enables real-time optimization of throughput, energy, and quality, particularly in flexible, high-mix environments where static models fail.

For infrastructure systems—power grids, transport networks, water distribution, and built environments—digital twins integrate sensors, geospatial data, and operational logs with AI-based forecasting, anomaly detection, and "what-if" simulations. Neural operators and surrogate models allow near-real-time simulation of complex processes. Adaptive learning supports lifecycle management, extending asset longevity, reducing downtime, and enabling semi-autonomous operations through reinforcement learning or optimization.

Challenges include data quality, uncertainty management, IT–OT integration, and model interpretability, alongside governance and accountability. Responsible, engineered deployment with validation, explainability, cybersecurity, and standards-based oversight is essential. Properly implemented, AI-driven digital twins can transform manufacturing and infrastructure into intelligent, resilient, and sustainable systems aligned with Industry 4.0.

AI-Enhanced Twin Capabilities

AI integration significantly elevates the functionality of digital twins, transforming them from descriptive or diagnostic models into predictive, adaptive, and decision-capable systems. By embedding machine learning, deep learning, and optimization algorithms, AI-enhanced digital twins extract actionable intelligence from continuous data streams, providing advanced capabilities crucial for smart manufacturing and resilient infrastructure.

Predictive analytics is a core strength, enabling predictive maintenance and remaining useful life estimation. AI models learn complex degradation patterns from historical and real-time sensor data, anticipating failures before they occur, reducing unplanned downtime, extending asset life, lowering operational costs, and improving safety.

Adaptive system modeling allows digital twins to continuously recalibrate as systems age or operating conditions change. Unlike static models, AI-driven twins evolve with new data, accounting for wear, environmental variability, and process drift, ensuring accuracy over long lifecycles.

Real-time optimization and decision support are further enhanced. AI evaluates multiple scenarios, balancing objectives like efficiency, quality, energy consumption, and emissions. Reinforcement learning and surrogate-assisted optimization allow recommendations for control actions, production schedules, or resource allocation, which can be executed autonomously within safe operational limits.

AI-driven twins also detect anomalies, support root-cause analysis, and improve operator trust through explainable AI. Scenario analysis and resilience planning further enable risk assessment, strategic mitigation, and robust system design.

Together, these capabilities make AI-driven digital twins central intelligence hubs, capable of learning, predicting, and guiding decisions throughout the asset lifecycle, making them indispensable in next-generation engineering systems.

Challenges

AI-driven digital twins offer transformative potential in smart manufacturing and infrastructure systems, yet their deployment faces substantial technical, organizational, and governance challenges. A primary technical challenge is ensuring data quality, availability, and integration.

Digital twins require continuous, high-fidelity streams from heterogeneous sensors, legacy machinery, and enterprise systems, but real-world data are often noisy, incomplete, or affected by sensor drift. Integrating operational technology (OT) with information technology (IT) systems further complicates pipelines, creating interoperability issues and high costs. Poor data quality can propagate through AI models, undermining predictions and twin reliability.

Model fidelity, generalization, and uncertainty management present additional challenges. AI models trained under specific operating conditions may fail under distribution shifts caused by equipment aging, process changes, or extreme events. Overfitting, insufficient uncertainty quantification, and lack of physical grounding can yield unsafe recommendations. Real-time requirements compound the difficulty of balancing accuracy with computational efficiency.

Scalability also poses challenges: large-scale plants and infrastructure networks generate massive, high-frequency data, demanding efficient architectures, edge–cloud coordination, and substantial computational resources. Sustainability concerns arise from energy-intensive AI models.

Explainability, trust, and human acceptance are critical. Opaque AI outputs may be over- or under-trusted by operators, requiring interpretable insights, clear visualization, and effective human–machine interfaces. Governance challenges include accountability, cybersecurity, regulatory compliance, and lifecycle management.

Addressing these issues requires rigorous systems engineering, robust data and model governance, explainable AI, cybersecurity safeguards, and standards-aligned institutional frameworks. Collaboration among engineers, data scientists, domain experts, and policymakers is essential for safe, reliable, and scalable deployment of AI-driven digital twins.

Generative AI in Engineering Design, Simulation, and Optimization

Generative Artificial Intelligence (AI) is transforming engineering by shifting focus from incremental design improvement to systematic exploration of entire design spaces. Unlike traditional workflows, where engineers manually propose and refine designs, generative AI learns underlying design distributions and constraints from data, simulations, and domain knowledge. This enables automated creation of novel, high-performing alternatives, particularly in complex, multi-objective problems. Variational autoencoders, generative adversarial networks, and diffusion-based models produce geometries, layouts, microstructures, and system configurations satisfying functional and manufacturability constraints. Coupled with physics-based solvers or surrogate models, these approaches allow rapid evaluation of performance metrics such as strength, thermal efficiency, and fluid flow. Generative AI also accelerates simulations by approximating expensive numerical solvers, enabling real-time digital twins, sensitivity analysis, and scenario evaluation. In optimization, generative models learn distributions over high-quality solutions, facilitating efficient sampling and exploration, especially in constrained, multi-objective problems. Challenges include ensuring compliance with safety standards, physical laws, and manufacturability, as well as increased verification needs. Hybrid workflows integrating generative creativity with physics-based validation, explainable AI, and human expertise offer the most promising pathway for responsible, high-impact engineering innovation.

Generative Design

Generative design is an advanced computational paradigm in which Artificial Intelligence (AI) systems automatically generate, evaluate, and refine design alternatives to satisfy functional requirements, constraints, and performance objectives. Unlike traditional workflows, where engineers manually create and iteratively adjust designs, generative design explores vast, high-dimensional design spaces, producing thousands of feasible solutions. This is particularly valuable in contexts with complex trade-offs among performance, cost, weight, energy efficiency, manufacturability, safety, and sustainability.

Technically, generative design integrates AI-driven models—such as variational autoencoders, generative adversarial networks, diffusion models, and evolutionary algorithms—with physics-based analysis and optimization. Candidate designs are evaluated via simulations or surrogate models for structural integrity, thermal behavior, fluid flow, and other domain-specific metrics. Feedback refines the generative process, forming a closed-loop design–evaluate–learn cycle.

A key strength is discovering non-intuitive, efficient structures, often with organic or lattice-like geometries, impactful in mechanical, aerospace, architecture, civil, and materials engineering. Combined with additive manufacturing, complex designs become fabricable. Engineers focus on defining objectives and constraints, while AI explores solutions. Challenges remain in validation, explainability, and certification. Effective frameworks integrate AI creativity with physics-informed constraints, human oversight, and explainable decision support, enabling innovation while ensuring rigor, safety, and accountability.

Generative AI for Simulation and Decision Support

Generative Artificial Intelligence (AI) is transforming engineering simulation and decision support by addressing the computational bottlenecks of physics-based modeling. Traditional simulations—such as finite element analysis, computational fluid dynamics, and multiphysics solvers—offer high-fidelity insights but are often too slow for real-time decision-making, large-scale design exploration, or operational optimization. Generative AI complements these approaches by learning surrogate models that approximate simulator behavior, enabling fast, data-driven inference while preserving essential physical relationships.

In simulation, generative models—including variational autoencoders, generative adversarial networks, diffusion models, and neural operators—learn mappings from input conditions to system responses using data from high-fidelity simulations or experiments. Once trained, they provide near-instantaneous predictions, supporting sensitivity analysis, uncertainty propagation, and scenario evaluation. In digital twins, generative surrogates enable real-time updating and interactive "what-if" exploration, which would be computationally infeasible with traditional solvers.

For decision support, generative AI facilitates multi-objective trade-off analysis, evaluating alternatives in terms of performance, cost, safety, energy, and environmental impact. Coupled with optimization, reinforcement learning, or Bayesian methods, it proposes actionable strategies while augmenting human judgment. Challenges include extrapolation errors and probabilistic interpretation, requiring integration with uncertainty quantification, explainable AI, and human-in-the-loop validation.

Hybrid architectures combining generative surrogates with physics-based models and domain constraints offer the most effective path, delivering both rigor and scalability for intelligent engineering systems.

Edge AI and Federated Learning for Cyber-Physical Systems

Cyber-Physical Systems (CPS) integrate computational intelligence with physical processes through networks of sensors, actuators, and control logic, forming the foundation of smart manufacturing, autonomous transportation, energy systems, and critical infrastructure. These systems require real-time, reliable, and safe operation, challenging traditional cloud-centric AI approaches. Edge AI and Federated Learning (FL) have emerged as key enablers, bringing intelligence closer to physical assets while preserving privacy, reducing latency, and enhancing resilience.

Edge AI deploys machine learning models directly on edge devices—such as embedded controllers, industrial gateways, and smart sensors—allowing low-latency inference for closed-loop control, anomaly detection, and safety monitoring. Local computation reduces reliance on continuous network connectivity and minimizes bandwidth use. Federated Learning complements this by enabling distributed model training across CPS nodes without centralizing raw data. Nodes share only model updates with an aggregator, maintaining data sovereignty and supporting collaboration across sites while improving shared models.

Together, edge AI and FL allow CPS networks to adapt and scale, balancing real-time responsiveness with global coordination. Hierarchical architectures spanning device-level, edge aggregation, and cloud coordination enable predictive maintenance, distributed fault detection, energy management, and cooperative robotics. Challenges include heterogeneous data, hardware constraints, communication limits, and adversarial risks, necessitating robust aggregation, uncertainty-aware learning, secure protocols, and integration with safety mechanisms.

When designed systematically, edge AI and FL enable CPS to operate as intelligent, privacy-preserving, and resilient systems in dynamic, real-world environments.

Edge Intelligence

Edge intelligence refers to deploying artificial intelligence—particularly machine learning and deep learning—directly at or near data sources within cyber-physical systems, rather than relying solely on centralized cloud infrastructure. In engineering systems such as smart manufacturing, autonomous vehicles, power grids, and critical infrastructure, this paradigm addresses the need for low-latency decisions, high reliability, data privacy, and operational resilience. By embedding intelligence into edge devices like sensors, controllers, and gateways, edge intelligence enables real-time perception, analysis, and control tightly integrated with physical processes.

A key advantage is support for closed-loop control and safety-critical operations. Systems operating on millisecond time scales cannot tolerate delays from remote cloud servers. Edge-deployed models perform anomaly detection, state estimation, and control locally, ensuring timely responses under network disruptions, maintaining stability in robotics, industrial automation, and autonomous mobility.

Edge intelligence also reduces bandwidth usage and preserves data privacy by transmitting only relevant insights rather than raw sensor data. Architectural design emphasizes lightweight, hardware-aware models, compression techniques, and efficient inference, enabled by TinyML and specialized accelerators. Combined with federated learning and digital twins, edge intelligence allows cyber-physical systems to function as autonomous, adaptive, and resilient entities.

Federated Learning

Federated Learning (FL) is a distributed machine learning paradigm that enables collaborative model training across multiple decentralized data sources while keeping raw data local. This approach is particularly valuable in cyber-physical systems, smart manufacturing, healthcare infrastructure, energy systems, and other engineering domains where data are sensitive, proprietary, or regulated. Unlike centralized learning, FL trains a shared global model by aggregating locally computed updates from edge or site-level nodes, preserving data privacy and ownership.

From an engineering perspective, FL addresses key constraints of data sovereignty, scalability, and heterogeneity. Distributed industrial systems often have diverse operating conditions, equipment, and usage patterns. FL allows models to leverage collective learning across these environments without exposing confidential operational data, enhancing applications such as predictive maintenance, fault diagnosis, and quality monitoring.

Technically, FL introduces challenges: data across nodes are often non-independent and non-identically distributed, slowing convergence and affecting performance. Devices vary in computation, bandwidth, and availability. Adaptive aggregation, partial participation, communication-efficient updates, and privacy-preserving mechanisms like secure aggregation and differential privacy are employed to mitigate these issues.

In cyber-physical systems, FL must align with real-time and safety constraints. Hierarchical architectures integrating on-device, edge, and cloud learning balance responsiveness, optimization, and governance, ensuring reliable, secure, and scalable distributed intelligence.

Ethical, Responsible, and Sustainable AI

Ethical, Responsible, and Sustainable Artificial Intelligence (AI) is essential for the long-term deployment of AI in science, engineering, and society, particularly as AI increasingly influences safety-critical decisions, resource allocation, and environmental outcomes. Ethical AI ensures alignment with human values—fairness, transparency, accountability, and autonomy—while responsible AI operationalizes these values through governance, technical controls, and lifecycle practices. Sustainability extends the scope to environmental and societal impacts, addressing energy use, carbon footprint, and resource efficiency.

Responsible AI requires robust data governance, representative and bias-free datasets, traceable model development, and auditable decision logic, ensuring accountability in regulated and high-stakes domains. Human-in-the-loop supervision, explainability, and uncertainty awareness enable informed reliance and operational oversight. Sustainable AI emphasizes energy-efficient architectures, model compression, edge deployment, and lifecycle-aware optimization, while also leveraging AI to reduce emissions, optimize energy, and improve infrastructure resilience. Adaptive governance frameworks—covering monitoring, auditing, and compliance—are critical to maintaining trust, legitimacy, and the ethical, responsible, and sustainable operation of AI systems over time.

Governance and Regulation

Governance and regulation are critical for ensuring that Artificial Intelligence (AI) systems—especially in scientific, engineering, and safety-critical domains—operate lawfully, accountably, and in alignment with societal values. Informal ethical principles are insufficient for AI that directly affects infrastructure, industrial operations, healthcare, and environmental management. Formal governance frameworks and regulatory regimes translate high-level values such as safety, transparency, fairness, and accountability into enforceable operational controls.

From an engineering perspective, governance covers the entire AI lifecycle, including data acquisition, model development, validation, deployment, monitoring, and retirement. Clear role assignments ensure human institutions, not opaque algorithms, remain accountable. Governance mechanisms encompass data stewardship, risk assessment, model versioning, documentation, and incident response, integrating with safety management and quality assurance systems.

Regulatory approaches are increasingly risk-based, with high-risk applications—such as critical infrastructure or healthcare—requiring rigorous validation, human oversight, traceability, and post-deployment monitoring. Adaptive AI necessitates continuous compliance through monitoring, logging, and periodic re-certification. Governance and regulation are enablers, supporting trustworthy, scalable, and societally responsible AI adoption across engineering domains.

Sustainability Considerations

Sustainability has become a critical consideration in the design, deployment, and governance of Artificial Intelligence (AI) systems, particularly as AI is increasingly embedded in large-scale engineering infrastructures and scientific workflows. While AI provides powerful tools for optimization, prediction, and decision-making, it also imposes environmental, economic, and social costs that must be managed to ensure long-term viability and societal acceptance. Sustainable AI requires a lifecycle-oriented approach, balancing performance with resource efficiency and broader environmental impact.

Key sustainability challenges include the energy and carbon footprint of large-scale AI models. Training and deployment consume substantial computational resources, driving energy demand and emissions. Techniques such as model compression, pruning, quantization, transfer learning, and edge deployment reduce computational costs while maintaining performance. Lifecycle-aware optimization—covering data collection, model development, maintenance, and decommissioning—further enhances efficiency by encouraging incremental learning, selective retraining, and model reuse.

Importantly, AI can also enable sustainability outcomes through predictive maintenance, energy optimization, and waste reduction. Integrating human-centered design, workforce planning, and governance ensures that sustainable AI is technically effective, socially responsible, and environmentally beneficial.

Integrated Reference Architecture

An integrated reference architecture is critical for deploying Artificial Intelligence (AI) systematically across the lifecycle of scientific discovery, engineering design, and operational decision-making. Modern engineering systems are complex, heterogeneous, and often safety-critical, requiring AI solutions embedded in a coherent, layered framework that aligns data, models, physical systems, and governance mechanisms. Such an architecture ensures technical integration while providing reliability, explainability, scalability, and regulatory compliance.

At its foundation is the physical and sensing layer, comprising machines, sensors, actuators, and controllers responsible for high-fidelity data acquisition and real-time feedback. Above this is the data and integration layer, managing ingestion, preprocessing, storage, and provenance, integrating heterogeneous operational and IT data while enforcing FAIR principles.

The modeling and intelligence layer hosts machine learning, hybrid, generative, and uncertainty-aware models for discovery, simulation, and optimization. The simulation and digital twin layer enables scenario analysis and real-time decision support, while the decision-making and control layer translates insights into safe, actionable operations.

The edge–cloud continuum supports scalable deployment, federated learning, and low-latency inference. Governance, assurance, and sustainability form a transversal layer enforcing ethics, regulatory compliance, cybersecurity, lifecycle management, and environmental responsibility.

This architecture ensures AI solutions are trustworthy, scalable, and sustainable, enabling integrated, end-to-end intelligent engineering systems.

Future Research Directions

Future research in Artificial Intelligence (AI) for science, engineering, and cyber-physical systems must address foundational limitations while extending capabilities toward trustworthy, sustainable, and autonomous ecosystems. A critical direction is the development of physics-aware and scientifically grounded AI models. Data-driven approaches excel within training regimes but struggle under distribution shifts; integrating physical laws, causal structures, and domain constraints can improve extrapolation, scientific validity, and reliability. Advances in physics-informed learning, neural operators, and hybrid symbolic–numeric reasoning are essential.

Robust autonomy in digital twins and self-driving laboratories is another priority. AI systems must manage uncertainty, sensor drift, unmodeled dynamics, and rare events without compromising safety or reproducibility. Research should focus on adaptive experiment planning, fault-tolerant autonomy, and lifelong learning, while ensuring auditability of autonomous pipelines.

Certifiable and explainable AI is critical for safety-critical applications, requiring verifiable guarantees, runtime monitoring, and formal verification. Edge intelligence and federated learning pose additional challenges in scalability, security, non-IID data, and distributed stability, calling for hierarchical learning, secure aggregation, and co-design with control systems.

Sustainability and governance require lifecycle-aware AI design, energy-efficient models, and adaptive regulatory frameworks. Human–AI collaboration, cognitive ergonomics, and organizational integration remain central for responsible deployment, ensuring AI augments human expertise within complex socio-technical systems.

Conclusion:

This chapter has shown that Artificial Intelligence (AI) is not just a set of algorithms but a transformative, system-level capability reshaping science, engineering, and complex infrastructure management. By spanning autonomous experimentation, advanced learning architectures, AI-driven digital twins, generative design, edge intelligence, federated learning, and responsible governance, AI enables cohesive integration of data, physical principles, and human expertise. Its full potential emerges when hybrid approaches combine data-driven learning with physics-informed reasoning, uncertainty quantification, and human oversight, ensuring safety, robustness, and accountability in long-lived or safety-critical systems.

Adaptive intelligence through digital twins, generative AI, and edge-enabled architectures supports continuous learning, real-time optimization, and resilience under uncertainty, which are vital for smart manufacturing, infrastructure, and sustainable development. However, these capabilities bring challenges in validation, lifecycle management, cybersecurity, and environmental impact, highlighting the need for integrated reference architectures and standards-based governance.

References:

1. Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Review Press.
2. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
3. Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
4. Bommasani, R. *et al.* (2022). On the opportunities and risks of foundation models. arXiv:2108.07258.
5. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., … Liang, P. (2022). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.
6. Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In K. Frankish & W. Ramsey (Eds.), *The Cambridge handbook of artificial intelligence* (pp. 316–334). Cambridge University Press.
7. Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science, 16*(3), 199–231.
8. Cai, C. J., Winter, S., Steiner, D., Wilcox, L., & Terry, M. (2019). "Hello AI": Uncovering the onboarding needs of medical practitioners for human–AI collaborative decision-making. *Proceedings of the ACM on Human-Computer Interaction, 3* (CSCW), 1–24.
9. Chen, T., Guestrin, C., & Assoc. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
10. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.
11. European Commission. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union.
12. European Parliament. (2025). *EU Artificial Intelligence Act: Key provisions and implementation timeline*. Publications Office of the European Union.

13. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
14. Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary perspectives on complex systems* (pp. 85–113). Springer.
15. Ha, D., & Schmidhuber, J. (2018). Recurrent world models facilitate policy evolution. *Advances in Neural Information Processing Systems*, 31, 2450–2462.
16. International Organization for Standardization. (2021). *ISO 23247-1: Automation systems and integration — Digital twin framework for manufacturing — Part 1: Overview and general principles*. ISO.
17. Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. SAGE Publications.
18. Lecun, Y. (2022). A path towards autonomous machine intelligence. *Meta AI Research White Paper*.
19. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.
20. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2017). Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research, 18*(185), 1–52.
21. Mersha, M., Singh, P., & Verma, A. (2024). Explainable artificial intelligence: A comprehensive survey of methods, applications, and challenges. *Artificial Intelligence Review, 57*(2), 1–45.
22. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2024), Recent trends in artificial intelligence and its applications, In: Artificial Intelligence- Trends and Applications, First Edition: DEcember 2024, ISBN: 978-93-95847-63-6, 73-106.
23. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025a), Environmental sustainability and ecological balance, In: Implementation of Innovative Strategies in Integral Plant Protection, First Edition: January 2025, ISBN: 978-93-48620-22-4, 81-96.
24. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025b), Advanced simulation techniques for forest fire and natural hazard prediction: A computational science perspective, Journal of Science Research International (JSRI), Vol. 11 (4) June 2025, 20-34.
25. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025c), Digital Guardians of Nature: Emerging AI Technologies in Plant and Animal Surveillance, In: Advances in Plant and Animal Sciences, First Edition: May 2025, ISBN: 978-93-49938-62-5, 12-35.
26. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025d), Artificial Intelligence and Machine Learning in Plant Identification and Biodiversity Conservation: Innovations,

Challenges, and Future Directions, In: Botanical Insights: From Traditional Knowledge to Modern Science, Volume I: May 2025, ISBN: 978-81-981142-3-5, 7-31.

27. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025e), Digital Guardians of Nature: Emerging AI Technologies in Plant and Animal Surveillance, In: Advances in Plant and Animal Sciences, Volume I: May 2025, ISBN: 978-93-49938-62-5, 12-35.
28. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025f), Advanced simulation techniques for forest fire and natural hazard prediction: A computational science perspective, Journal of Science Research International (JSRI), 11 (4): June 2025, 20-34.
29. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025g), Forest Health Monitoring Using AI and Remote Sensing, ISBN (PDF) 9783389142202 ISBN (Book) 9783389142219, July 2025.
30. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025h), Artificial intelligence and big data in environmental monitoring and decision support: revolutionizing ecosystem management, Journal of Science Research International (JSRI), 11 (5): July 2025, 28-39.
31. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025i), Climate change, biodiversity and ecological resilience, In: Green Footprints: Bridging Environment and Sustainability, First Edition: July 2025, ISBN: 978-81-989981-8-7, 25-47.
32. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025j), AI for industry, education, and research: Transforming knowledge, innovation and society, In: AI for Industry, Education and Research, First Edition: September 2025, ISBN: 978-81-993182-6-7, 41-64.
33. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025k), Foundations and Core Concepts of Artificial Intelligence, In: Concept, Techniques and Applications, An International Edition: October 2025, ISBN: 978-93-49938-72-4, 31-49.
34. Mishra, R.K., Mishra, Divyansh and Agarwal, R. (2025l), Artificial intelligence and intelligent systems, In: Innovations and Research in Science and Technology Volume III: December 2025, ISBN: 978-93-47587-40-5, 26-47.
35. Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Farrar, Straus and Giroux.
36. Nguyen, P., Zhang, Y., Li, X., & Kumar, R. (2025). AI-driven digital twins for smart manufacturing: A comprehensive review. *Sensors, 25*(4), 1123.
37. Rai, A., & Sinha, R. (2024). Artificial intelligence for safety-critical systems in industrial and transportation domains. *ACM Computing Surveys, 56*(3), 1–38.
38. Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press.
39. Reichstein, M. *et al.* (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566, 195–204.

40. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature, 566*(7743), 195–204.
41. Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
42. Schmidt, J. *et al.* (2019). Machine learning in materials science. npj Computational Materials, 5, 83.
43. Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials, 5*(83).
44. Silver, D., Sutton, R. S., Müller, M., Schaul, T., Hassabis, D., & others. (2018). The reinforcement learning revolution. *Nature, 558*(7709), 232–234.
45. Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science, 361*(6404), 751–752.
46. Tom, G. *et al.* (2024). Self-driving laboratories for chemistry and materials science. Chemical Reviews, 124, 1–58.
47. Tom, G., Aspuru-Guzik, A., & coworkers. (2024). Self-driving laboratories for chemistry and materials science. *Chemical Reviews, 124*(1), 1–58.
48. Vapnik, V. (1998). *Statistical learning theory*. Wiley.
49. Vickers, J. (2024). Digital twins in a nutshell. *DOE–NSF Workshop on Digital Twins*, NASA.
50. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology, 10*(2), 1–19.
51. Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2021). Understanding deep learning requires rethinking generalization. *Communications of the ACM, 64*(3), 107–115.
52. Zhou, Q., Wang, J., & Li, Y. (2025). Generative artificial intelligence for engineering design and optimization: A systematic review. *Engineering Applications of Artificial Intelligence, 121*, 106039.

A SHORT NOTES ON FLUID AND ITS APPLICATIONS

Paban Dhar
Department of Mathematics,
Karimganj College, Sribhumi, Assam-788710
Corresponding author E-mail: paban85dhar@gmail.com

Continuum Hypothesis:

At the macroscopic level, the description of the motion of a fluid involves a study of the behavior of the entire discrete molecule which makes up the fluid. However, when one is dealing with problems in which some characteristic length in the flow is very large compared with molecular distances, it is convenient to think of a lump of fluid sufficiently small from macroscopic point of view but large enough at the microscopic level so as to contain a large number of molecules and to work with the average statistical properties of such large number of molecules. In such a case the detailed molecular structure is washed out completely and is replaced by a continuous model of matter having appropriate continuum properties so defined as to ensure that on the macroscopic scale the behavior of the model resembles with the behavior of the real fluid. When the characteristic length in the flow is not large compared with molecular distances, the continuum model is invalid and the flow must be analyzed on the molecular scale.

The smallest lump of fluid material having sufficiently large number of molecules to allow statistically of a continuum interpretation is called a fluid particle. This study deals with fluid obeying continuum hypothesis. Further, it is assumed that the fluid properties are the same at all points in the fluid and are identical in all directions from any specified point. These stipulations make the fluid both homogeneous and isotropic.

Classical Theory:

Newton (1687) performed an experiment by considering a flow of water between two parallel plates where the upper plate moves in water with velocity V and d is the distance between the parallel plates. He observed that the force experienced by the upper plate is in the direction of motion and the force is directly proportional to the velocity but it is inversely proportional to the distance between the plates. Mathematically, Newton established the relation between the stress tensor and strain rate tensor as

$$\tau_{ij} = 2\mu e_{ij}$$

where

$$e_{ij} = \frac{1}{2}\left(v_{i,j} + v_{j,i}\right)$$

v_i is the velocity component in the direction of i. τ_{ij} is the extra stress.

The constitutive equation for isotropic, viscous, incompressible fluid is defined as

$$\sigma_{ij} = -p\delta_{ij} + 2\mu e_{ij}$$

where σ_{ij} is the stress tensor, δ_{ij} is the Kronecker's delta, p is hydrostatic pressure and μ is the coefficient of viscosity. This linear relation between the stress and strain is based on the Newton's law of viscosity. The fluid model based on the linear constitutive equation between stress and strain rate is called Newtonian fluid.

The classical theory of fluid mechanics is based on the linear constitutive equation given by Newton. This constitutive equation discussed perfectly the motion of water, glycerin and many thin oils. Also, on the basis of the linear relation it was possible to explain the theory of formation of drag, lift, skin friction, separation of the fluid flow. Many theoretical works on the fluid flows have been done by applying the linear constitutive equation. But this classical theory is incapable to explain a number of phenomena observed in a large number of liquids.

Generalization of Classical Linear Theory:

Merrington (1943) noticed that when a solution of rubber in mineral oil is forced through a straight pipe, the fluid swells after emerging out of the tube. In 1946, Garner and Nissan observed that when a cylindrical rod is rotated in a container containing metallic soap, hydro-carbon gels, high polymer solutions and a variety of other materials, the fluid rises up to the rod to a considerable extent. This phenomenon was demonstrated by Weissenberg (1947, 1949) and is known as Weissenberg effect. Weissenberg (1947, 1949) performed experiments with various types of fluids. In his experiments, the liquids were sheared in a gap between an outer vessel rotated at various constant angular velocities and inner cylinder that could be either fixed rigidly or allowed to move up and down. The phenomenon of rising the liquid in a direction perpendicular to the plane of shearing is called normal stress effect (Poynting effect) and that of drawing of the liquid towards the axis of rotation against the action of centrifugal force when the liquid is sheared between two rotating plates is called centripetal pump effect. Reiner (1957) found that this effect is present even in air if the distance between the bases of the cylinders is less than a certain critical distance. The physics behind these experimental results could not be described with the help of the linear constitutive equation given by Newton.

This inadequacy of the classical linear theory of Newtonian fluids generalized the linear relationship between stress and strain rate tensor. Oldroyd, Walters and their co-workers generalized the linear relation between stress and strain by experimental results. They introduced Oldroyds liquids and Walters liquids etc. to account for the non-Newtonian phenomena. Another group of scientists viz. Truesdell, Noll, Coleman, Ericksen and their co-workers obtained the constitutive equations to account for the non-Newtonian phenomena by using the general idea of fluidity and then approximated these constitutive relations to give the behavior of certain fluids. Below, we present some constitutive models which are frequently used in the literature.

Applications:

Fluid dynamics is one of the most important part of the recent interdisciplinary activities concerning engineering and technological activities. It is one of the oldest branches of applied mathematics in which some of the most significant advances have been made during the last five decades. Nearly two hundred years ago, man thought of laying down scientific rules governing the motion of fluids, water and air, mainly to use the rules to understand these elements so that he could not only protect himself from their fury during natural calamities but also utilize their powers to develop fields like civil engineering and naval architecture. In spite of this common origin, two distinct thoughts gradually developed. On the one hand, through the concept of an ideal fluid, mathematical physicists developed the theoretical science known as classical hydrodynamics. On the other hand, realizing that idealized theories were of no practical application without empirical correction factors, engineers developed from experimental studies the applied science known as hydraulics, for the scientific fields of irrigation, water supply, river flow control, hydraulic power and so on. Further, the development of aeronautical, chemical and mechanical engineering during the last few decades and the exploration of space from 1960s have increased the interest in the study of fluid mechanics.

The science of fluid mechanics has been extended into fields like regimes of hypervelocity flight and flow of electrically conducting fluids. This has introduced new fields of interest such as hypersonic flow and magneto fluid dynamics. In this connection, it has become essential to combine the knowledge of thermodynamics, heat transfer, mass transfer, electromagnetic theory and fluid mechanics for the complete understanding of the physical phenomenon involved in any flow process.

Fluid dynamics is one of the rapidly growing basic sciences whose principles find application even in daily life. For instance, the flight of a bird in air and the motion of fish in water are governed by the fluid dynamic rules. The design of various types of aircraft and ships are based on the principles of fluid dynamics. Even natural phenomena like tornadoes and hurricanes can also be explained by the science of fluid dynamics. In fact, the science of fluid dynamics dealing with such natural phenomena has been developed to such an extent that they can be predicted well in advance. Since the earth is surrounded by an environment of air and water to a very large extent, almost everything that happens on earth and its atmosphere, in some way or the other is associated with the science of fluid dynamics.

References:

1. Merrington, A. C. (1943). *Nature, 152*, 663.
2. Garner, F. H., & Nissan, A. H. (1946). *Nature, 158*, 634.
3. Weissenberg, K. (1947). *Nature, 159*, 310.
4. Weissenberg, K. (1949). *Proceedings of the International Congress of Rheology, 1*, 29.
5. Reiner, M. (1957). *Proceedings of the Royal Society A, 240*, 553.

QUANTUM COMMUNICATION AND QUANTUM - SAFE NETWORKING FOR FUTURE WIRELESS SYSTEMS

Allanki Sanyasi Rao
Christu Jyothi Institute of Technology & Science,
Jangaon-506167, Telangana, India
Corresponding author E-mail: srao_allanki@cjits.org

Abstract:
Quantum communication leverages core principles of quantum mechanics—such as superposition, entanglement, and the no-cloning theorem—to enable highly secure data transmission that surpasses the guarantees of conventional cryptographic methods. In parallel, wireless networks beyond 5G are evolving to provide exceptional security, reliability, and availability for applications including critical infrastructure, healthcare, autonomous systems, and large-scale sensing. These developments are increasingly converging, leading to future wireless architectures that are quantum-aware and resilient to quantum-era threats. This chapter outlines key quantum communication fundamentals, including single- and multi-qubit state preparation, Bell-state measurements, and measurement-induced disturbances. It reviews major protocols such as BB84 and E91 Quantum Key Distribution, along with quantum teleportation and superdense coding, from both security and implementation viewpoints. Post-quantum cryptographic methods are also introduced as practical classical alternatives where quantum infrastructure is lacking. Finally, the chapter discusses integrated architectures that merge quantum links, repeaters, and satellite systems with terrestrial 5G/6G networks, highlighting deployment challenges, standardization efforts, and open research directions.

Keywords: Quantum Communication, Entanglement, Quantum Key Distribution, Quantum Teleportation, Post-Quantum Cryptography, Quantum Repeaters, Quantum-Safe Networking.

1. Introduction:

1.1 Background and Motivation

Classical information theory, introduced by Claude Shannon, establishes the basis for communication over noisy channels by representing information as bits. Conventional cryptographic methods such as RSA and elliptic curve cryptography depend on the computational difficulty of problems like integer factorization and discrete logarithms on classical computers. The development of scalable quantum computers capable of executing Shor's algorithm poses a serious risk to these schemes by making key-breaking computationally feasible. At the same time, modern wireless networks operating in millimeter-wave and sub-terahertz frequencies, supported by massive antenna systems, software-defined radios, and artificial intelligence, process increasing volumes of sensitive data in areas such as healthcare, finance, industry, and autonomous transportation. The move toward edge-centric architectures, network slicing, and distributed computing further raises system complexity and expands potential attack surfaces.

Quantum communication offers a fundamentally new security paradigm by exploiting quantum-mechanical principles that do not rely on computational hardness. By using measurement-induced disturbances and the no-cloning property of quantum states, it enables reliable detection of eavesdropping. Supported by quantum repeaters and satellite links, quantum communication facilitates secure key distribution over large distances, forming the foundation of a future quantum-secured internet.

1.2 Quantum Phenomena and Communication

Quantum communication architectures are founded on five fundamental quantum mechanical effects.

- **Superposition:** A quantum system can exist in multiple basis states simultaneously, allowing a single qubit to represent both logical 0 and 1 at the same time and thereby providing higher information capacity than classical bits.
- **Entanglement:** Two or more qubits can share strong non-classical correlations, where measurements on spatially separated particles produce correlated outcomes that are valuable for secure communication protocols.
- **No-Cloning Theorem:** Quantum mechanics forbids the perfect duplication of an unknown quantum state, meaning that any eavesdropping attempt introduces disturbances that can be detected by legitimate users.
- **Measurement-Induced Disturbance:** Observing a quantum state forces it to collapse, irreversibly destroying information linked to other measurement bases and exposing unauthorized access.
- **Bell Inequality Violation:** Certain entangled states generate correlations beyond classical limits, confirming that these correlations arise from genuine quantum entanglement rather than classical shared information.

2. Fundamentals of Quantum Information for Communication

2.1 Qubits and Quantum State Representation

A quantum bit, known as a qubit, is the fundamental unit of quantum information and is realized using a physical system with two distinct quantum states. In quantum communication, qubits are typically implemented using different photonic encoding techniques, such as:

- **Photon Polarization:** Qubit states are represented using orthogonal polarization modes, such as horizontal and vertical orientations. This approach is widely adopted due to its simplicity and ease of state preparation and measurement.

- **Photon Time-Bin Encoding:** Information is encoded based on a photon's arrival in either an early or late time slot. This technique is particularly resilient to polarization fluctuations in optical fiber links.
- **Photon Path Encoding:** Here, the qubit value is determined by the spatial path followed by a photon, typically implemented using beam splitters or integrated photonic waveguides.
- **Phase Encoding:** Data is carried through the relative phase difference between interferometer paths, providing enhanced stability for long-distance fiber-based quantum communication networks.

A general single-qubit state can be described as a normalized linear superposition of the computational basis states:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

where the complex coefficients α and β satisfy the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. The values $|\alpha|^2$ and $|\beta|^2$ correspond to the probabilities of obtaining the measurement outcomes $|0\rangle$ and $|1\rangle$ respectively when the qubit is measured in the computational basis.

States involving multiple qubits reside in a tensor product Hilbert space. For example, a two-qubit quantum state is represented as:

$$|\psi\rangle_{12} = \sum_{i,j\in\{0,1\}} c_{ij}|i\rangle_1 \otimes |j\rangle_2$$

where $\sum_{i,j} |c_{ij}|^2 = 1$

To describe realistic quantum systems that may undergo noise and decoherence, the density matrix framework is used to generalize the state vector description to mixed states, which are statistical ensembles of pure states:

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$$

where $p_i \geq 0$ are probabilities and $\sum_i p_i = 1$. Pure quantum states satisfy $\mathrm{Tr}(\rho^2) = 1$, while mixed states satisfy $\mathrm{Tr}(\rho^2) < 1$.

2.2 Entanglement and Bell States

Entanglement is the signature resource of quantum information. A two-qubit state is entangled if it cannot be factorized as a product of single-qubit states.

The four maximally entangled Bell states form an orthonormal basis for the two-qubit Hilbert space:

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

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

$$|\Psi^{+}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$

$$|\Psi^{-}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

These states exhibit perfect correlation or anti-correlation when measured in the computational basis. For the state $|\Phi^{+}\rangle$, measurements of qubits 1 and 2 always yield identical outcomes, yet the state is a superposition where the outcomes are undefined until measurement.

Entanglement can be quantified using the concurrence $C(\rho)$, defined for a two-qubit density matrix ρ as:

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}$$

where λ_i denotes the eigenvalues of the matrix $R = \sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$ in descending order, and $\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^{*}(\sigma_y \otimes \sigma_y)$ with σ_y the Pauli-Y matrix. For pure entangled states $C = 1$, whereas for separable states $C = 0$.

2.3 The No-Cloning Theorem and Eavesdropping Detection

The no-cloning principle, first demonstrated by Wootters and Zurek, states that an exact duplication of an unknown quantum state cannot be achieved:

Theorem: There is no unitary operation U that can clone an arbitrary qubit state $|\psi\rangle$ such that $U|\psi\rangle|\phi_0\rangle = |\psi\rangle|\psi\rangle$ for all possible input states $|\psi\rangle$, where $|\phi_0\rangle$ is an initial "blank" state.

This phenomenon has important implications for secure communication. If an eavesdropper, typically referred to as Eve, attempts to intercept and measure quantum signals, she may obtain an outcome but cannot reproduce and resend the original quantum state to the receiver, Bob, without introducing observable disturbances. As a result, the communicating parties can detect such intrusion by observing an unexpected rise in the quantum bit error rate.

2.4 Quantum Measurement and State Collapse

The quantum measurement postulate specifies that if a system prepared in the state $|\psi\rangle$ is measured using an observable whose eigenvalues $\{a_n\}$ are with associated orthonormal eigenbasis $\{|\phi_n\rangle\}$, the result of the measurement a_n is with probability $p_n = |\langle\phi_n|\psi\rangle|^2$, and the system subsequently reduces to the state $|\phi_n\rangle$.

Crucially, when a quantum state $|\psi\rangle$ is observed using a basis different from the one in which $|\psi\rangle$ was originally prepared, the measurement outcome becomes effectively random and the original state is permanently disturbed. This behavior forms the foundation of eavesdropping detection in prepare-and-measure QKD protocols.

2.5 Quantum Channels and Decoherence

Quantum communication channels include:

- **Optical fiber:** Quantum information carried by single-photon states travels through optical fibers and experiences attenuation described by a loss coefficient α (dB/km). As a result, the photon intensity at the receiver decreases according to $I(z) = I_0 \exp(-\alpha z)$. For standard single-mode optical fibers operating at the 1550 nm telecommunications wavelength, typical attenuation values are approximately $\alpha \approx 0.17$ dB/km.

- **Free-space optical (FSO):** Free-space atmospheric links are affected by effects such as turbulence-driven beam fluctuations, scattering, and absorption losses. The overall atmospheric transmittance varies with the operating wavelength, prevailing weather conditions, and the length of the transmission path.
- **Satellite quantum links:** Quantum signals transmitted from satellites in orbit to terrestrial ground stations rely on free-space optical (FSO) propagation while also being influenced by the curvature of the Earth and the continuously changing geometry of the communication link due to satellite motion and orbital dynamics.

The primary quantum channel imperfections are:

- **Attenuation/Loss:** During transmission, there is a probability $L = 1 - \exp(-\alpha L_{\text{channel}})$ that photons are lost over a propagation distance L_{channel} which directly reduces the number of photons successfully reaching the receiver.
- **Decoherence:** Interactions between the quantum system and its surrounding environment lead to a gradual loss of quantum coherence. In the case of a dephasing channel, after a time t, the off-diagonal elements of the density matrix decay as $\exp(-t/T_2)$ where T_2 represents the characteristic dephasing time of the system
- **Noise:** Unwanted photons originating from ambient illumination, background radiation, or thermal sources introduce errors in detection and degrade the reliability of quantum communication links.

3. Quantum Key Distribution Protocols

3.1 The BB84 Protocol: Prepare-and-Measure QKD

The Bennett–Brassard 1984 (BB84) protocol is recognized as the earliest practical and widely implemented scheme for quantum key distribution. It operates by making use of two mutually conjugate measurement bases: the computational or rectilinear basis $\{|0\rangle, |1\rangle\}$) and the diagonal basis $\{|+\rangle, |-\rangle\}$ in which the states are represented as $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$[2]. The BB84 protocol generates a series of qubits with randomized lengths, as dictated by the encryption algorithm.

3.1.1 Protocol Steps

The BB84 protocol proceeds as follows:

- **State preparation (Alice):** Alice creates a random sequence of bits $\{b_1, b_2, \dots, b_n\}$ along with an independent random sequence of basis selections $\{B_1, B_2, \dots, B_n\}$ where each $B_i \in$ {rectilinear,diagonal}. Based on each bit b_i and the corresponding chosen basis B_i, Alice prepares the appropriate quantum state:

If B_i = rectilinear and $b_i = 0$: prepare $|0\rangle$

If B_i = rectilinear and $b_i = 1$: prepare $|1\rangle$

If B_i = diagonal and $b_i = 0$: prepare $|+\rangle$

If B_i = diagonal and $b_i = 1$: prepare $|-\rangle$

- **Transmission:** Alice then transmits the entire sequence of prepared qubits to Bob over the quantum communication channel, allowing the quantum states to propagate from the sender to the intended receiver.
- **Measurement (Bob):** Bob receives the incoming qubits and, for each one, independently chooses a measurement basis $\{B'_1, B'_2, \ldots, B'_n\}$ at random and performs the appropriate measurement. He then records both the selected measurement bases and the resulting outcomes $\{r_1, r_2, \ldots, r_n\}$.
- **Basis reconciliation (public channel):** Alice and Bob disclose their chosen measurement bases $\{B_i\}$ and $\{B'_i\}$, through an authenticated public classical channel, while keeping the actual quantum measurement results secret. They then determine the set of indices $S = \{i: B_i = B'_i\}$ for which their basis selections agree. The bits corresponding to these positions S constitute the sifted key.
- **Error estimation:** Alice and Bob randomly select and discard a portion of the sifted key bits for the purpose of estimating the quantum bit error rate (QBER). For these chosen positions, they openly reveal and compare the corresponding bit values over a public channel, and based on this comparison, they calculate:

$$\text{QBER} = \frac{\text{Number of disagreements}}{\text{Number of sacrificed bits}}$$

If the quantum bit error rate (QBER) surpasses a predefined threshold—commonly set at 11% for the BB84 protocol—the protocol is terminated to prevent the distribution of a key that may have been intercepted or compromised.

- **Error correction and privacy amplification:** The sifted key bits that remain are subjected to classical error-correction techniques, such as Hamming codes, to fix any errors introduced during transmission. Additionally, if required, privacy amplification methods—like Toeplitz-matrix-based compression—are applied to minimize any potential information that could have been gained by an eavesdropper.

3.1.2 Security Analysis

The security of the BB84 protocol is fundamentally based on the principles of the no-cloning theorem and the disturbance caused by quantum measurements. Consider an eavesdropper, Eve, who tries to intercept the qubits transmitted through the quantum channel in an attempt to gain information about the key.

Eve's best approach is to measure each qubit she receives using a randomly selected basis and then forward the resulting state to Bob. Because there is a 50% chance that she chooses the wrong basis, her measurements inevitably introduce errors into the transmission. On average, the interference from Eve's actions results in a quantum bit error rate (QBER) given by:

$$\text{QBER}_{\text{Eve}} = \frac{1}{4}$$

This occurs because Eve selects the incorrect basis for half of the qubits, and in those cases, the outcome of her measurement has a 50% chance of differing from the bit originally prepared by Alice. Thus $\text{QBER} = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

Eve's intrusion is revealed when the measured quantum bit error rate (QBER) exceeds the level that would be expected from normal channel noise. The rate at which a secure key can be generated, denoted as *r* (in bits per qubit), is expressed as:

$$r = 1 - H(QBER) - H(\text{QBER}_{\text{Eve}})$$

where, $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ represents the binary Shannon entropy. In the case of a perfect channel with QBER = 0, $r = 1 - 0 - H(1/4) \approx 0.19$ bits per qubit can be obtained after applying privacy amplification.

3.2 Entanglement-Based QKD: The E91 Protocol

The Ekert 1991 (E91) protocol offers a different method for quantum key distribution by employing entangled photon pairs instead of relying on prepare-and-measure quantum states.

3.2.1 Protocol Steps

Entangled pair generation: A source generates pairs of entangled photons in the quantum state:

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

In this scheme, the qubits are encoded using either the polarization or the time-bin of the photons, with one photon directed to Alice and the other to Bob.

Random measurement basis selection: Alice and Bob each select their measurement bases independently and at random from three possible choices (e.g., 0°, 45°, 90° for photon polarization):

- Basis A_1 or A_2 for Alice
- Basis B_1 or B_2 for Bob

Measurement and outcome recording: Both parties perform a measurement on their respective photons and record the resulting value, either 0 or 1.

Public basis announcement: Alice and Bob share their selected measurement bases with each other over a public channel, making their basis choices known to both parties, while carefully keeping the actual results of their measurements secret and undisclosed.

Bell inequality test (optional): Alice and Bob can use a portion of their shared data to perform tests of Bell inequalities, such as the CHSH inequality, in order to confirm that true quantum entanglement exists between their particles.

$$S = E(A_1B_1) + E(A_1B_2) + E(A_2B_1) - E(A_2B_2)$$

where $E(A_iB_j)$ represents the expected value of the product of measurement results obtained in the chosen bases. For genuinely entangled states, $|S| \leq 2\sqrt{2}$, whereas correlations explainable by classical physics satisfy $|S| \leq 2$. Observing a violation of the Bell inequality serves as confirmation of quantum entanglement.

Sifted key extraction: As in the BB84 protocol, only the bits corresponding to matching measurement bases are kept and used to construct the sifted key.

3.2.2 Security Properties

The E91 protocol ensures security even when the source of entangled photon pairs cannot be trusted, as any eavesdropping attempt can be identified by monitoring violations of Bell inequalities. If an eavesdropper, Eve, tries to extract information about the key, her intervention inevitably disrupts the entanglement, resulting in deviations from the expected Bell inequality correlations.

3.3 Practical QKD Variants

Practical QKD systems encounter imperfections in their devices, which, in principle, could be exploited by an eavesdropper. To mitigate these risks, several advanced protocols have been developed:

3.3.1 Decoy-State QKD

In real-world systems, photon sources may sometimes generate multi-photon pulses instead of single photons, creating a security risk where an eavesdropper can split the pulse, keep some photons, and forward the rest without detection. Decoy-state QKD mitigates this issue by transmitting pulses with different intensity levels, including signal and decoy states, and examining the detection statistics to bound the information an attacker could gain. This approach effectively enhances QKD security against photon-number-splitting attacks.

3.3.2 Measurement-Device-Independent QKD (MDI-QKD)

Flaws in measurement devices can enable side-channel attacks, allowing an eavesdropper to extract key information by exploiting detector weaknesses. Measurement-Device-Independent Quantum Key Distribution (MDI-QKD) overcomes this problem by requiring Alice and Bob to send their quantum states to a central measurement station that need not be trusted. A joint Bell-state measurement is carried out at this station, removing the need to rely on the security of the detectors themselves. As a result, the protocol's security is independent of the reliability of measurement devices, making it robust against many detector-related attacks. Moreover, MDI-QKD supports secure communication even when the measurement station is controlled by a third party or placed in an unsecure location, enhancing its practicality for real-world quantum networks.

The secret key rate for MDI-QKD is:

$$r_{\text{MDI}} = \frac{1}{2}[1 - H(QBER)] - H(QBER)$$

here, the first term accounts for the contributions from single-photon events, while the second term quantifies the potential information that could be leaked to an eavesdropper.

4. Quantum Teleportation and Quantum Repeaters

4.1 Quantum Teleportation Protocol

Quantum teleportation is a method for transferring an unknown quantum state between distant locations without physically moving the particle itself. It depends on shared entanglement between the sender and receiver, along with classical communication of measurement results. The sender performs a joint measurement on the unknown state and one part of an entangled pair, then sends the outcomes through a classical channel. Using this information, the receiver applies an appropriate unitary operation to recreate the original quantum state. This process faithfully preserves quantum information while respecting the no-cloning principle and serves as a key foundation for advanced quantum technologies, including long-distance quantum networks, distributed quantum computing, and secure quantum communication.

4.1.1 Protocol Description

Suppose Alice has an unknown qubit state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ that she wants to send to Bob over a distance. To facilitate this, Alice and Bob share a pre-established entangled pair in the state $|\Phi^+\rangle$:

Initial state: The combined system is:

$$|\psi\rangle_A \otimes |\Phi^+\rangle_{AB} = (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Bell measurement (Alice): Alice carries out a joint Bell-state measurement on the qubit she intends to teleport along with her portion of the shared entangled pair. This measurement collapses the combined system into one of the four possible Bell states and produces a 2-bit classical result.

State correction (Bob): Alice transmits the 2-bit result of her measurement to Bob over a classical communication channel. Using this information, Bob performs the appropriate unitary transformation on his part of the entangled pair. Once this correction is applied, Bob's qubit exactly replicates the original state $|\psi\rangle$.

The quantum teleportation protocol involves the following resources:

- One qubit of entanglement shared beforehand between the sender and receiver
- Transmission of two classical bits of information
- A single local quantum measurement, specifically a Bell-state measurement
- One local unitary operation applied by the receiver

4.1.2 Mathematical Analysis

Following Alice's Bell-state measurement resulting in outcome $b_1 b_2 \in \{00,01,10,11\}$, the state of Bob's qubit immediately after the measurement is:

$$|\psi_{\text{Bob}}\rangle = (\sigma_Z^{b_1} \sigma_X^{b_2})|\psi\rangle$$

where σ_X and σ_Z represent Pauli operators, and Bob performs the unitary correction $U = \sigma_Z^{b_1} \sigma_X^{b_2}$ to reconstruct the original quantum state $|\psi\rangle$.

4.2 Quantum Repeaters and Entanglement Distribution

Quantum repeaters enhance the distance over which quantum communication can occur by dividing a long quantum channel into shorter, more manageable segments and then employing entanglement swapping to connect and extend entanglement across these segments.

4.2.1 Entanglement Swapping

The key process in a quantum repeater is entanglement swapping. Imagine three nodes positioned in a line: Alice at the first node, a relay node R at the second, and Bob at the third.

Generation: Independent entangled pairs are created:

- Pair AB: qubits A and R_A in state $|\Phi^+\rangle_{AR_A}$
- Pair BC: qubits R_B and B in state $|\Phi^+\rangle_{R_B B}$

Bell measurement at relay: The relay node carries out a joint Bell-state measurement on qubits A and B, collapsing them into one of the four possible Bell states.

Result: Following the Bell-state measurement and the transmission of the measurement result via a classical channel, qubits A and B are now entangled:

$$|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{AB}$$

The two-bit classical result from the relay's Bell-state measurement needs to be sent to one of the end nodes, usually Bob, so that he can perform the correct unitary operation.

4.2.2 Multi-Hop Quantum Repeater Networks

For a sequence of quantum repeaters covering a total distance L, where each segment has a length ℓ, the total number of hops is $N = L/\ell$. The maximum achievable secure key rate for the entire chain is constrained by:

$$r_{\text{total}} = r_{\text{segment}}/N$$

where r_{segment} represents the key rate that can be achieved over an individual segment. To address the limitation of linear scaling with distance, quantum repeater protocols utilize hierarchical entanglement purification and distillation techniques.

The Briegel-Dür quantum repeater, introduced in 1998, showed that it is possible for key rates to scale exponentially with the communication distance:

$$r_{\text{total}} \sim \exp(-\eta L)$$

where η is a coefficient determined by the channel loss and the efficiency of the entanglement purification process. This approach offers a significant advantage compared to direct transmission, in which the key rate decreases exponentially according to $r_{\text{direct}} \sim \exp(-2\alpha L)$ with α representing the fiber's attenuation coefficient.

5. Post-Quantum Cryptography and Quantum-Safe Architectures

5.1 Threat from Quantum Computers

Shor's algorithm, introduced by Peter Shor in 1994, is a quantum algorithm that can efficiently factor large numbers and solve discrete logarithm problems in polynomial time on a sufficiently advanced quantum computer. Modern public-key cryptosystems such as RSA-2048 and ECC-256 offer about 128 bits of security against classical attacks, but they would become vulnerable once large-scale quantum computers with approximately 2000–4000 logical qubits are built, a development expected between 2030 and 2040.

The arrival of such cryptographically relevant quantum computers (CRQCs) presents not only a future danger but also a backward-looking risk to currently stored data. Encrypted information intercepted today could be decrypted later when quantum capabilities mature, a threat known as "harvest now, decrypt later." This risk highlights the need to accelerate the development of quantum-resistant cryptography and to transition sensitive data to post-quantum secure systems well before powerful quantum computers become available.

5.2 Post-Quantum Cryptography Algorithms

In 2022, the National Institute of Standards and Technology (NIST) officially approved four families of cryptographic algorithms designed to be secure against quantum attacks.

5.2.1 Lattice-Based Cryptography

Lattice-based problems, including the Learning With Errors (LWE) problem, are considered resistant to attacks by quantum computers. The ML-KEM (Kyber) key encapsulation mechanism operates as follows:

- Key encapsulation: $\text{CT} = E(PK, m)$ generates a cipher text CT along with a shared secret m
- Decapsulation: $m' = D(SK, \text{CT})$ uses their private key SK to recover the shared secret.
- Security assumption: The mechanism's security relies on the computational difficulty of solving the LWE problem in arbitrary dimensions.

ML-KEM provides post-quantum security levels of 128, 192, and 256 bits, making it suitable for protecting sensitive communications against future quantum threats.

5.2.2 Hash-Based Signatures

Hash-based signature schemes, such as SPHINCS+ (now referred to as SLH-DSA), offer stateless digital signatures that rely entirely on cryptographic hash functions. The overall security of these schemes is directly dependent on the strength and integrity of the underlying hash function.

- Signature generation: $\sigma = \text{Sign}(SK, m)$
- Verification: $\text{Verify}(PK, m, \sigma) \in \{\text{True,False}\}$
- Security assumption: Pre-image resistance and collision resistance of hash functions

5.2.3 Code-Based Cryptography

The Classic McEliece code-based encryption scheme provides security against quantum attacks by leveraging the computational hardness of decoding arbitrary linear codes.

- Encryption: $c = m \cdot G + e$ where G is generator matrix and e is error vector
- Decryption: Decode c by removing e using private algebraic code structure

Code-based cryptographic schemes benefit from a long-standing history of study and well-understood, conservative security assumptions, but they typically involve significantly larger key sizes compared to other approaches.

5.3 Hybrid Quantum-Safe Architectures

Real-world deployments employ hybrid architectures combining quantum-based and classical post-quantum methods:

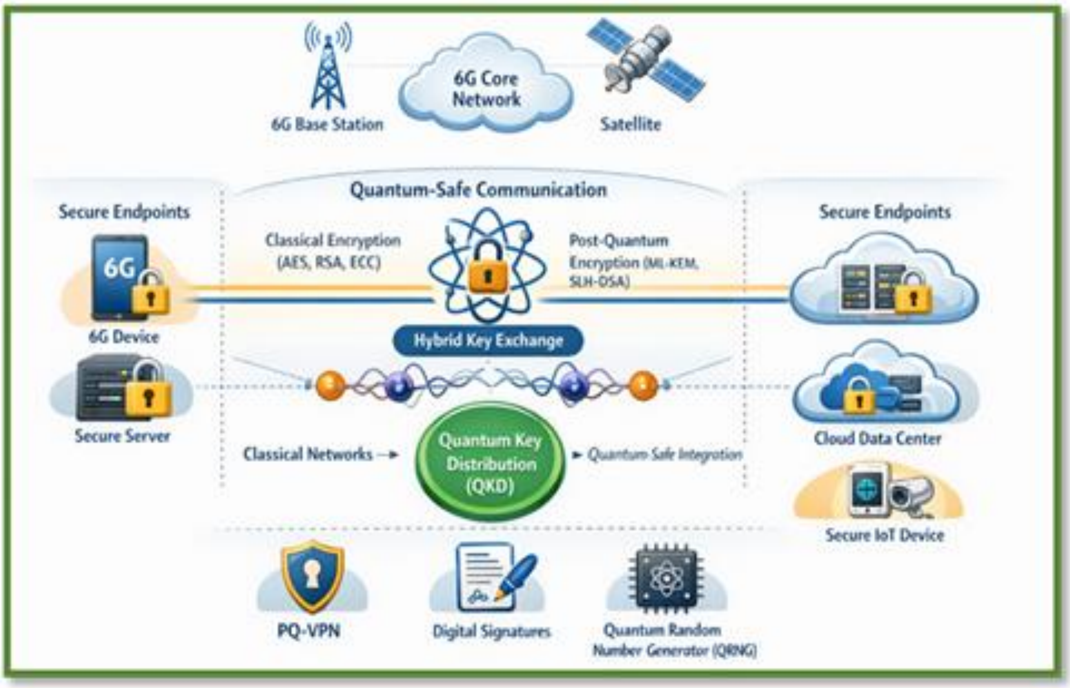


Figure 1: Hybrid quantum-safe architecture for 6G networks

Classical post-quantum cryptographic techniques are used to protect wireless edge connections and terrestrial backhaul links, whereas quantum key distribution is applied to secure essential backbone paths and links between terrestrial networks and satellites. By combining these approaches, a multi-layer security framework is achieved that provides defense-in-depth through complementary protection mechanisms.

Key distribution strategy:

- **Backbone network:** Quantum Key Distribution (QKD) links securely deliver symmetric keys between core network nodes, using technologies such as entangled-photon satellites and metropolitan quantum networks to support long-distance, high-security key exchange . This layer provides the security base for the entire network.
- **Terrestrial backhaul:** Links connecting base stations to the core network are protected with post-quantum cryptography. ML-KEM is used for secure key establishment, while AES-256 ensures data confidentiality, offering protection against both classical and future quantum threats.
- **Wireless access:** At the access layer, base stations use symmetric keys generated through QKD or post-quantum methods to create encrypted sessions with user devices, ensuring end-to-end security for user data and signaling.
- **Critical services:** High-risk applications, including power grids, aviation, and emergency communications, combine QKD-based security with out-of-band verification and multiple post-quantum algorithms. This layered approach improves robustness and reliability against advanced attacks on mission-critical systems.

5.4 Migration Strategy

Organizations transitioning to quantum-safe cryptography face several key challenges that must be addressed to maintain long-term security and system reliability.

- **Algorithm diversity:** Using multiple standardized post-quantum cryptographic (PQC) algorithms provides flexibility to replace any scheme found to be vulnerable. This reduces reliance on a single primitive and lowers the risk of a single point of failure.
- **Hybrid deployment:** During migration, systems typically run classical algorithms such as RSA and ECC alongside post-quantum schemes like ML-KEM and SLH-DSA. This hybrid model preserves backward compatibility while supporting a gradual shift to quantum-resistant security.
- **Latency and overhead:** Many PQC schemes, particularly lattice- and code-based ones, require larger keys and more computation than classical public-key methods. For instance, ML-KEM-768 uses keys of about 1,184 bytes, much larger than those in RSA-2048, which can increase bandwidth use, processing load, and latency in resource-limited settings.
- **Testing and validation:** Thorough testing is vital before large-scale deployment to ensure correct implementation and interoperability. Weak implementations can introduce new risks, making rigorous evaluation and compliance checks essential during the transition.

6. Network Architectures for Quantum and Quantum-Safe Systems

6.1 Terrestrial Quantum Metropolitan Networks

Metropolitan-scale quantum networks link quantum nodes—such as entanglement sources, quantum memory units, and measurement devices—across ranges spanning several tens to hundreds of kilometers through optical fiber infrastructure.

6.1.1 Network Topology

Quantum metropolitan networks are typically designed using star or mesh topologies, based on scalability and reliability needs.

- **Star topology:** A central quantum hub connects to multiple edge nodes through dedicated quantum links. This design simplifies coordination and management but makes the network highly dependent on the security and reliability of the central node.
- **Mesh topology:** Quantum nodes are linked through multiple paths, offering redundancy and flexible routing. This approach improves fault tolerance and resilience to link failures, making it well suited for large, high-reliability quantum network deployments.

6.1.2 Quantum Repeater Stations

Repeater stations placed along optical fiber links carry out tasks such as entanglement swapping and purification to extend quantum correlations over long distances. By mitigating photon loss and decoherence in fiber channels, these operations preserve the fidelity of entangled states and enable reliable long-distance quantum communication across multiple network segments.

Table 1: Quantum repeater station functions, typical delays, and achievable fidelities using current technology

Repeater Function	Operation	Latency	Fidelity
Entanglement generation	Spontaneous parametric down-conversion	1-100 ms	0.85-0.95
Storage (quantum memory)	Rare-earth ions or NV centers	1 s	0.9
Bell state measurement	Interferometer + detectors	1 μs	0.75-0.85
Entanglement swapping	BSM + classical feedback	100 μs	0.70-0.80

6.2 Satellite Quantum Communication

Satellite-based quantum communication enables the global expansion of quantum networks by using spaceborne platforms to establish long-distance quantum links beyond the limits of terrestrial infrastructure. By leveraging free-space optical channels between satellites and ground stations, this approach overcomes fiber attenuation and allows secure quantum key distribution and entanglement distribution on a worldwide scale.

6.2.1 Satellite-Ground QKD Links

Quantum signals exchanged between satellites and ground stations must travel through roughly 600 km of atmosphere and near-space, where turbulence, scattering, and photon loss can degrade signal quality. The impact varies with transmission direction:

- **Uplink:** Sending signals from the ground to a satellite is more challenging due to strong beam divergence and atmospheric turbulence near the Earth's surface, leading to significant spreading and loss. This requires accurate beam pointing and high-power sources.
- **Downlink:** Transmission from satellite to ground is more practical and has been successfully demonstrated. For example, China's Micius satellite achieved quantum key distribution over distances up to 1200 km, demonstrating the viability of satellite-based quantum communication for large-scale networks.

The transmittance of a satellite downlink is:

$$T_{\text{sat}} = T_{\text{atmosphere}} \times T_{\text{optics}} \times T_{\text{detector}}$$

Typical values: T_{sat} ~ 0.01-0.1 during daylight, ~ 0.1-0.5 at night.

6.2.2 Global Quantum Network Vision

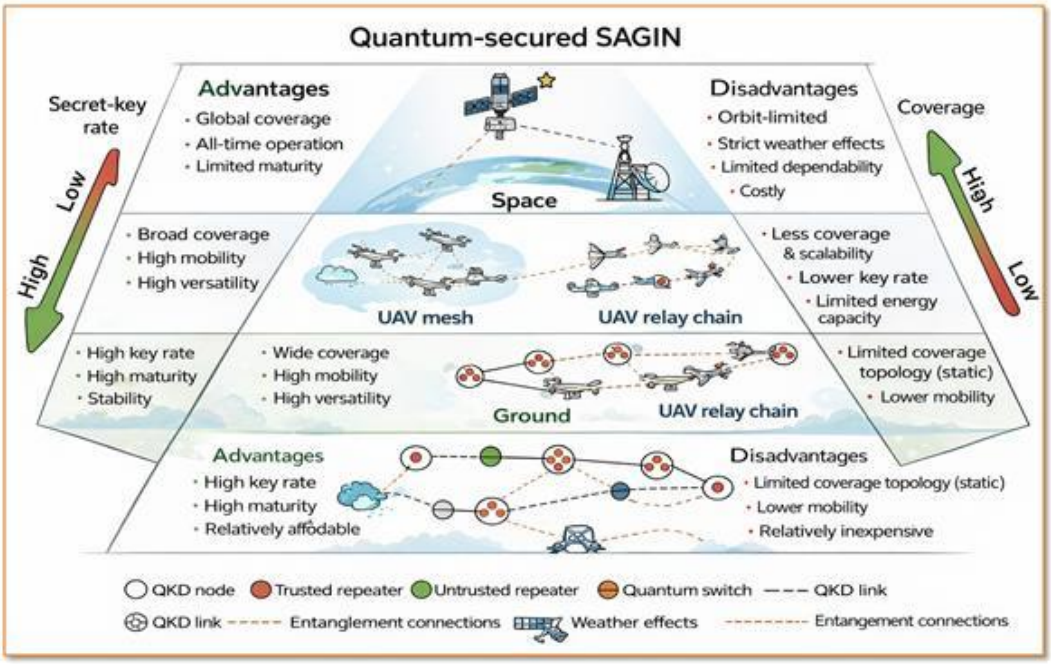


Figure 2: Space-air-ground integrated quantum-safe network architecture

A constellation of quantum-enabled satellites can provide near-continuous quantum key distribution to ground stations worldwide, forming the backbone of a future global quantum internet. Satellite QKD links create secure long-distance key exchanges with ground stations, while terrestrial metropolitan quantum networks serve urban regions by connecting local quantum nodes. High-altitude platforms help relay quantum signals over broader areas. At the same time, conventional terrestrial and wireless links are secured using post-quantum cryptography. In 5G and upcoming 6G networks, base stations authenticate devices and distribute keys generated through quantum or quantum-resistant methods to end users.

6.3 Integration of Quantum Communication with 5G and 6G

6.3.1 QKD in 5G Core Networks

The 5G core network architecture supports the integration of Quantum Key Distribution (QKD) across three layers to strengthen overall security.

- **Service layer:** Network slicing and service orchestration use quantum-generated keys to protect control and signaling channels, securing communication among virtualized and cloud-native components
- **Transport layer:** Front haul and backhaul links apply QKD-based protection where quantum infrastructure is available, while Post-Quantum Cryptography (PQC) is used when QKD is not feasible, ensuring secure data transfer.
- **Access layer:** Within the radio access network (RAN), quantum-safe key establishment is used between the network and user equipment to protect authentication and user data from future quantum threats.

6.3.2 6G Quantum-Aware Architecture

The design of 6G networks is expected to natively integrate quantum technologies across multiple architectural elements to deliver major advances in performance, security, and intelligence.

- **Quantum-secure control plane:** Network control and management functions, especially those supporting ultra-reliable low-latency communication (URLLC), use quantum-generated keys to provide information-theoretic security against both classical and quantum attacks.
- **Quantum-enhanced sensing:** Technologies such as quantum radar and magnetometry improve positioning accuracy and environmental awareness. Their high-resolution sensing data supports dynamic network optimization, adaptive beamforming, and efficient resource allocation.
- **Distributed quantum computing:** 6G networks are expected to enable access to edge-based quantum processors, supporting distributed quantum computing. This allows low-latency quantum machine learning and hybrid classical–quantum processing on geographically distributed data.
- **Quantum-enabled distributed sensing:** Entanglement-based sensing across spatially separated nodes improves measurement precision and synchronization, enabling advanced applications such as wide-area monitoring and coordinated sensing over large regions.

6.3.3 Protocol Integration Points

In 6G, security integrates across all layers: TLS 1.3 with post-quantum ciphers at the application layer; QUIC using quantum-derived session keys at transport; IPsec with ML-KEM for network security; Wi-Fi 7 authenticated via quantum keys at the link layer; and physical-layer protection through spread-spectrum encoding combined with quantum randomness.

7. Challenges, Standardization and Research Directions

7.1 Technical Challenges

7.1.1 Quantum Channel Impairments

Channel loss defines the primary distance constraint in quantum communication. The relationship between the secret key rate and the channel transmission can be approximated as:

$$r(T) = T - 2H(T) \text{ (asymptotic key rate)}$$

For real-world fiber links, a loss of $\alpha \approx 0.17$ dB/km over distance $T = \exp(-\alpha L)$ leads to a transmission of $T \approx 0.03$ at 100 km, producing very low key rates unless quantum repeaters are employed.

7.1.2 Quantum Memory and Storage

Quantum memories need to preserve quantum states long enough to enable entanglement swapping at relay stations, maintaining a fidelity of $F \geq 0.9$. Present-day implementations include:

- **Rare-earth ion doping in crystals:** $T_{\text{coherence}} \sim 1$ second, $F \sim 0.95$[2]
- **Nitrogen-vacancy centers in diamond:** $T_{\text{coherence}} \sim$ milliseconds, $F \sim 0.9$[1]
- **Atomic ensembles:** $T_{\text{coherence}} \sim$ microseconds, $F \sim 0.85$[2]

7.1.3 Single-Photon Detection

The performance of QKD is constrained by both the photon detection efficiency (quantum efficiency η) and the dark count rate. High-efficiency detectors reduce errors and improve key rates. For example, superconducting nanowire single-photon detectors (SNSPDs) offer efficiency $\eta \sim 0.9$ with very low dark counts of 1–10 Hz, while avalanche photodiodes (APDs) provide efficiency $\eta \sim 0.6$ but higher dark counts of 100–1000 Hz. Overall, better detection efficiency lowers the quantum bit error rate, directly enhancing the secure key generation rate.

7.2 Standardization and Policy

7.2.1 NIST Post-Quantum Cryptography Standardization

In August 2024, NIST published standardized post-quantum cryptography (PQC) algorithms, including ML-KEM, ML-DSA, and SLH-DSA. Organizations are advised to first assess their cryptographic systems to identify algorithms susceptible to quantum attacks. They should build cryptographic agility to allow smooth transitions between algorithms, adopt hybrid classical–PQC solutions during the migration phase, and rigorously test PQC implementations prior to full deployment. Careful planning ensures secure and seamless adaptation to quantum-resistant cryptography.

7.2.2 ITU-T and ETSI Quantum Standardization

International Telecommunication Union (ITU-T) and European Telecommunications Standards Institute (ETSI) have formed specialized groups to advance quantum communication and quantum-safe networking:

- **ITU-T Focus Group on Quantum Information Technology (FG-QIT):** Working on standards and architectural frameworks to support scalable and secure quantum communication networks.
- **ETSI Industry Specification Group (ISG) on Quantum-Safe Cryptography:** Developing guidelines for smooth migration to quantum-resistant cryptography, ensuring interoperability across systems and devices.

7.2.3 National Quantum Initiatives

Several nations have introduced national strategies to advance quantum technologies:

- **United States:** The National Quantum Initiative (NQI) promotes research in quantum networks, computing, and includes the Quantum Internet Alliance to foster secure communications.
- **European Union:** The Quantum Internet Alliance (QIA), supported with €1 billion funding, focuses on building a scalable and interoperable quantum internet infrastructure.
- **China:** Development of the Jinan Quantum Communications Network for QKD and the Micius satellite as part of a global quantum internet backbone demonstrates large-scale quantum communication deployment.
- **India:** Launched in 2023, the National Quantum Mission aims to advance quantum technologies, including secure quantum communication networks and related applications.

7.3 Open Research Directions

7.3.1 Device-Independent and Semi-Device-Independent QKD

Device-independent QKD (DI-QKD) removes the need to trust the quantum measurement devices, relying solely on observed measurement statistics and the violation of Bell inequalities. This approach offers the highest level of security, though it demands very high Bell test visibility, which remains difficult to achieve in practice.

Current research focuses on:

- Enhancing Bell inequality violations using practical, implementable systems.
- Evaluating security under finite sampling and realistic noise conditions.
- Minimizing the computational complexity involved in generating device-independent security proofs.

7.3.2 Continuous-Variable QKD

Continuous-variable (CV) QKD transmits information using the amplitude and phase quadrature of light instead of discrete photon states:

- **Advantages:** Supports implementation with conventional telecom hardware through homodyne or heterodyne detection, allowing potentially higher key rates and easier integration with existing fiber networks.
- **Challenges:** Demands highly accurate phase reference and tight synchronization; security analysis is more involved due to continuous nature of the quantum states and vulnerability to practical imperfections.

7.3.3 Quantum Machine Learning for Network Optimization

Quantum algorithms can enhance quantum repeater networks by efficiently tackling combinatorial optimization problems, improving both network topology design and entanglement routing strategies.

$$\text{Key rate} = \max_{\text{routing, purification}} \{r_1, r_2, \dots, r_N\}$$

Classical exhaustive search methods experience exponential growth in complexity as problem size increases, making them impractical for large-scale tasks. In contrast, quantum algorithms — such as variational quantum Eigen solvers (VQEs) — can offer significant advantages, potentially achieving quadratic speedups and enabling faster optimization for complex network or quantum system problems.

7.3.4 Micro integrated Quantum Photonics

Integrating various quantum optical components—such as waveguides, photon sources, modulators, and detectors—onto a single photonic chip minimizes the size, weight, and power (SWaP) of quantum communication devices:

- **Deployment advantages:** Facilitates use on satellites, UAVs, and other mobile platforms, enabling flexible and scalable quantum networks.
- **Performance benefits:** Enhances phase stability and lowers the risk of misalignment, improving overall system reliability.
- **Research focus:** Aims at achieving high on-chip quantum efficiency, effective wavelength conversion, and low-loss integration to support practical, high-performance quantum communication terminals.

7.3.5 Quantum Communication for Critical IoT

Ensuring the security of critical Internet-of-Things (IoT) applications—such as smart energy grids, industrial control systems, and autonomous vehicles—demands extremely high reliability and protection against potential quantum attacks:

Research directions include:

- Developing lightweight, quantum-resistant protocols suitable for resource-constrained IoT devices.
- Implementing distributed trust frameworks for secure quantum key management across IoT networks.
- Leveraging quantum communication technologies to enable robust vehicle-to-everything (V2X) authentication and prevent cyber threats.

Conclusion:

Quantum communication represents a major advance in information security by offering protection grounded in the laws of quantum physics rather than computational complexity. Its convergence with emerging 5G and 6G wireless networks is driving new global communication architectures resilient to both classical and quantum threats. Progress in this field is evident through mature QKD protocols, quantum repeaters that extend entanglement over long distances, satellite-based quantum links, and hybrid approaches that combine quantum methods with post-quantum cryptography. These developments show that quantum-safe networking is moving from theory toward real-world deployment. International standardization efforts by organizations such as NIST, ITU-T, and ETSI are guiding adoption, while key challenges remain, including improving detector efficiency, extending quantum memory lifetimes, reducing terminal size and power, and scaling repeater networks.

Looking ahead, integrating quantum communication into 6G networks is expected to enable highly secure control signaling, distributed quantum computing, and quantum-enhanced sensing. Future systems will likely rely on layered security models that merge quantum techniques with post-quantum cryptography. Ongoing research in device-independent and continuous-variable QKD, quantum machine learning, and photonic integration suggests that quantum-safe networking will become a core part of mainstream communication infrastructure in the coming decade.

References:

1. Bennett, C. H., & Brassard, G. (1984). Quantum cryptography: Public key distribution and coin tossing. *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, 175–179.
2. Ekert, A. K. (1991). Quantum cryptography based on Bell's theorem. *Physical Review Letters, 67*(6), 661–663.
3. Scarani, V., Bechmann-Pasquinucci, H., Cerf, N. J., Dušek, M., Lütkenhaus, N., & Peev, M. (2009). The security of practical private quantum key distribution. *Reviews of Modern Physics, 81*(3), 1301–1350.

4. Allanki Sanyasi Rao & Sreeja Mole (2023). The vital role of quantum cryptography in IoT network security. *European Chemical Bulletin, 12*(9), 573–587.
5. Agrawal, M., Kayal, N., & Saxena, N. (2004). PRIMES is in P. *Annals of Mathematics, 160*(2), 781–793.
6. Kalaivani, V. (2021). Enhanced BB84 quantum cryptography protocol for secure communication in wireless body sensor networks for medical applications. *Personal and Ubiquitous Computing, 27*(3), 875.
7. Paglierani, P., Raouf, A. H. F., Pelekanakis, K., Petroccia, R., Alves, J., & Uysal, M. (2023). A primer on underwater quantum key distribution. *Quantum Engineering*, 2023(1), 7185329.
8. Alghamdi, M. I. (2024). A review on quantum key distribution for wireless networks: Current status and future prospects. *Communications on Applied Nonlinear Analysis, 32*, 568–579.
9. Cachin, C., *et al.* (2023). *Post-quantum cryptography standardization process* (NIST Special Publication 800-176). National Institute of Standards and Technology.
10. Vazirani, U. V., & Vidick, T. (2016). Fully device-independent quantum key distribution. *Communications of the ACM, 59*(4), 35–43.
11. Zhong, H.-S., Wang, H., Deng, Y.-H., Chen, M.-C., Peng, L.-C., Luo, Y.-H., … Pan, J.-W. (2021). Quantum computational advantage using photons. *Science, 370*(6523), 1460–1463.
12. Yin, J., Cao, Y., Li, Y.-H., Ren, J.-G., Liao, S.-K., Zhang, L., … Pan, J.-W. (2020). Entanglement-based secure quantum key distribution over 1,120 kilometres. *Nature, 582*(7813), 501–505.

EMERGING ROLES OF MICRORNAS IN IMMUNE REGULATION AND AUTOIMMUNE DISEASE

Ravneet Kaur and Geetanjali
Department of Zoology,
Post Graduate Government College for Girls (PGGCG) Sector-11, Chandigarh
Corresponding author E-mail: 85kravneet@gmail.com, khokhargeetanjali@gmail.com

Abstract:
MicroRNAs (miRNAs) are small, evolutionarily conserved noncoding RNA molecules, typically ranging from 19 to 23 nucleotides in length that regulate gene expression post-transcriptionally. They have emerged as crucial modulators of the immune system, influencing both innate and adaptive immunity. Additionally, dysregulation of miRNA expression has been implicated in the pathogenesis of various autoimmune diseases. This research paper reviews the current understanding of miRNA function in immune responses and its contribution to autoimmune diseases, exploring their potential as biomarkers and therapeutic targets. These molecules play a pivotal role in regulating gene expression at the post-transcriptional level. Through their ability to influence the expression of target genes, miRNAs are essential in numerous biological processes, including immune cell differentiation, function, and response to infection.

Keywords: MicroRNA, Autoimmunity, Anticancer Immunity, Transplantation Immunology, Biomarkers.

Introduction:

MicroRNAs (miRNAs) are small, non-coding RNA molecules that play a pivotal role in post-transcriptional gene regulation. By binding to target messenger RNAs (mRNAs), they modulate gene expression, impacting diverse cellular processes. Over the past decades, their significance in immune system function and pathology has been increasingly recognized.

The immune system relies on a finely tuned balance between activation and regulation to protect the host from pathogens while avoiding self-reactivity. miRNAs are crucial regulators in this balance, orchestrating the differentiation, activation, and function of immune cells, including T cells, B cells, macrophages, and dendritic cells. Dysregulation of miRNAs can disrupt this equilibrium, contributing to the development of autoimmune diseases such as systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), and multiple sclerosis (MS).

This review highlights the role of miRNAs in immune cell differentiation, anti-infective responses, immunodeficiencies, and autoimmune diseases, while also discussing their involvement in anticancer immunity and organ/stem cell transplantation. It emphasizes the translational and clinical potential of miRNAs, focusing on personalized immunological disease management through prevention and novel therapies. The review underscores the challenges in utilizing miRNA fingerprints as biomarkers for diagnosis, prognosis, and treatment prediction, and explores the connection between lifestyle-induced molecular changes, nonpharmacological immunomodulation, and transgenerational epigenetic inheritance. Understanding these mechanisms provides insights into potential therapeutic strategies, highlighting miRNAs as promising biomarkers and targets for intervention.

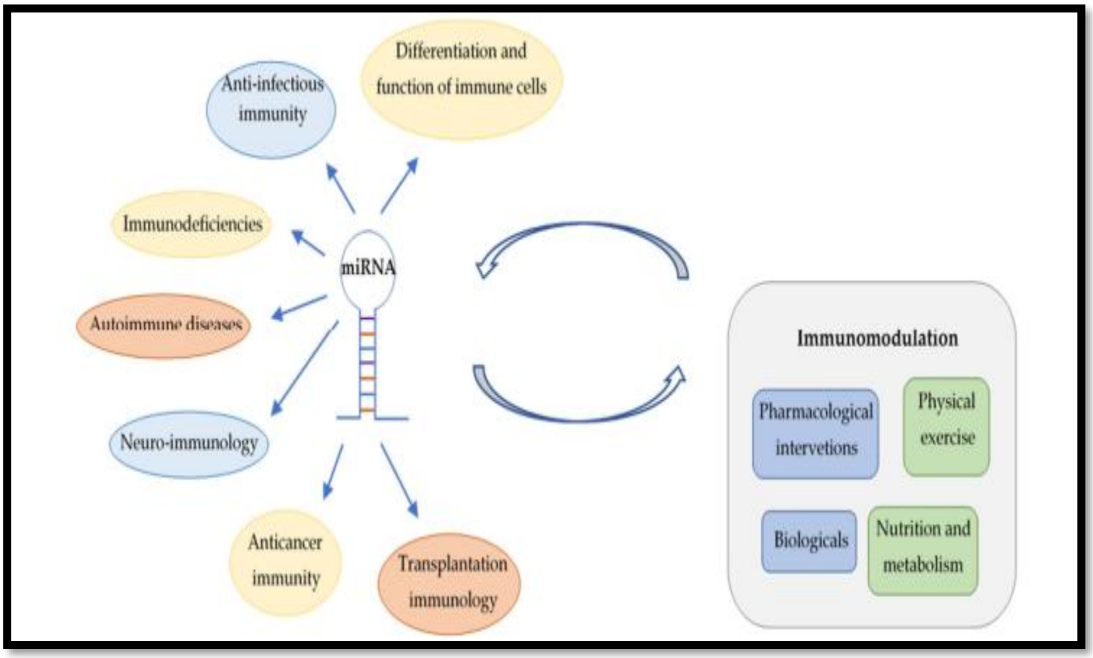


Figure 1: Involvement of miRNAs in immune regulation and interactions between miRNAs and different types of immunomodulation

Biogenesis and Mechanism of Action of Mirnas:

miRNAs are small non-coding RNA molecules involved in the regulation of gene expression. They are encoded by genes located in various regions of the genome, such as introns, exons, untranslated regions, or intergenic regions. The biogenesis and function of miRNAs can be summarized as follows:

Biogenesis:

1. Transcription in the Nucleus

- **Process**: miRNA genes are transcribed by RNA Polymerase II/III, resulting in a long primary miRNA (pri-miRNA).
- **Structure**: Pri-miRNA contains a stem-loop structure essential for subsequent processing.

2. Cleavage by the Drosha-DGCR8 Complex

- **Drosha and DGCR8**: In the nucleus, the pri-miRNA is processed by the Drosha-DGCR8 complex (also called the microprocessor complex).
- **Outcome**: Cleavage results in the formation of precursor miRNA (pre-miRNA), which is approximately 70 nucleotides long and retains the stem-loop structure.

3. Export to the Cytoplasm

- **Exportin-5 and Ran-GTP**: Pre-miRNA is transported out of the nucleus to the cytoplasm by Exportin-5 in a Ran-GTP-dependent manner.

4. Cytoplasmic Processing by Dicer

- **Dicer and TRBP**: The pre-miRNA is further processed in the cytoplasm by the Dicer enzyme, aided by TRBP (TAR RNA-binding protein).
- **Product**: This cleavage produces a short double-stranded RNA (miRNA duplex) consisting of a guide strand and a passenger strand.

5. Formation of the RISC Complex

- **Incorporation into RISC**: The guide strand of the miRNA duplex is loaded onto the RNA-induced silencing complex (RISC), where it associates with Argonaute 2 (Ago2).
- **Passenger Strand**: The passenger strand is degraded during this process.

6. Mature miRNA Function

- **Target mRNA Binding**: The mature miRNA guides the RISC complex to target mRNAs by base pairing with complementary sequences in the mRNA (usually in the 3' UTR).

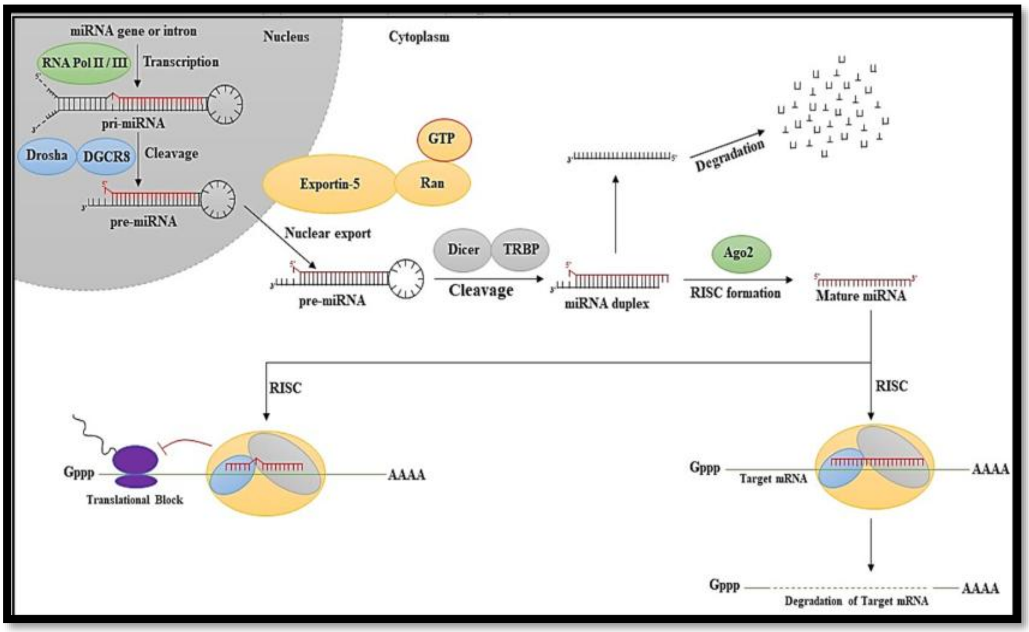


Figure 2: miRNA Biogenesis

Mechanism of Action:

- The guide miRNA within the RISC complex binds to complementary sequences on target mRNAs, with the seed region (6-8 nucleotides) playing a critical role in binding.

Modes of Gene Regulation:

- **Complete Complementarity**: Leads to mRNA cleavage and degradation.
- **Partial Complementarity**: Results in translational repression and mRNA destabilization.
- miRNAs can also regulate gene expression epigenetically by influencing chromatin structure.

Functions:

- miRNAs can act extracellularly, being secreted into the bloodstream within exosomes or associated with AGO proteins, influencing other cells.
- During apoptosis, miRNAs can be released into the extracellular space.
- miRNA function exhibits redundancy (a single mRNA regulated by multiple miRNAs) and pleiotropy (a single miRNA targets multiple mRNAs), forming intricate regulatory networks.
- miRNAs regulate approximately 60% of protein-coding genes and are subject to epigenetic and post-transcriptional regulation. Their expression is tissue-specific and developmentally regulated, making them vital in evolutionary gene regulation systems.

Involvment of miRNAs in the Development and Function of the Immune System

MicroRNAs (miRNAs) play a crucial role in the development and function of the immune system by regulating gene expression post-transcriptionally. They fine-tune immune responses, ensuring a balance between activation and suppression. In the innate immune system, miRNAs like miR-146a and miR-155 modulate Toll-like receptor (TLR) signaling, inflammatory cytokine production, and macrophage activation. In adaptive immunity, miRNAs are essential for lymphocyte differentiation, proliferation, and function. For example, miR-181a influences T-cell sensitivity and selection in the thymus, while miR-150 regulates B-cell development and antibody production. miRNAs also prevent autoimmunity by maintaining immune tolerance. Dysregulation of miRNAs is associated with autoimmune diseases, chronic inflammation, and immune deficiencies. Furthermore, miRNAs are involved in immune cell communication, modulating responses to infections, and maintaining tissue homeostasis. Their precise regulation is vital for effective immune responses, preventing hyperactivation or immunosuppression that can lead to pathological conditions.

Role of miRNAs in Innate Immunity

1. Modulating Toll-like Receptor (TLR) Signaling

miRNAs such as miR-146a and miR-155 regulate TLR signaling pathways, which are essential for recognizing pathogens.

- **miR-146a:** Acts as a negative regulator of inflammation by targeting TRAF6 and IRAK1, which are downstream molecules in the TLR pathway. This helps control excessive inflammatory cytokine production.
- **miR-155:** Enhances the pro-inflammatory response by targeting SOCS1, a suppressor of cytokine signaling.

2. Regulating Inflammatory Cytokine Production

miRNAs like miR-21 and miR-125b modulate the production of cytokines such as IL-6, TNF-α, and IL-10. Dysregulation of these miRNAs can lead to chronic inflammation.

3. Controlling Macrophage Polarization

Macrophages exist in two states: M1 (pro-inflammatory) and M2 (anti-inflammatory). miRNAs such as miR-223 and miR-155 influence macrophage polarization, thereby regulating inflammation and tissue repair.

Role of miRNAs in Adaptive Immunity

1. T-cell Development and Function

- **miR-181a:** Enhances T-cell receptor (TCR) signaling sensitivity, influencing T-cell selection in the thymus.
- **miR-155:** Modulates T-cell differentiation into Th1, Th2, and Th17 cells, essential for appropriate immune responses.
- **miR-146a:** Prevents T-cell hyperactivation by targeting signaling molecules involved in TCR activation.

2. B-cell Development and Antibody Production

- **miR-150:** Regulates B-cell maturation by targeting transcription factors like c-Myb. Its dysregulation can impair antibody production.
- **miR-155:** Facilitates class-switch recombination, a key process for generating diverse antibodies.

3. Regulatory T-cells (Tregs)

miRNAs such as miR-10a and miR-21 maintain the function of Tregs, which are crucial for immune tolerance. Their dysregulation can lead to immune overactivation and autoimmunity.

miRNAs and Immune Tolerance

Immune tolerance is critical to prevent the immune system from attacking self-antigens. miRNAs like miR-146a and miR-31 regulate pathways that maintain this tolerance.

- **miR-146a:** Suppresses hyperinflammatory responses by targeting signaling molecules like TRAF6 and STAT1.
- **miR-31:** Controls the development of Tregs, ensuring immune tolerance.

miRNAs in Autoimmune Diseases

1. Rheumatoid Arthritis (RA)

- **miR-146a:** Typically downregulated in RA, leading to uncontrolled inflammation.
- **miR-155:** Overexpressed in RA, promoting the differentiation of Th17 cells and increasing inflammation.

2. Systemic Lupus Erythematosus (SLE)

- **miR-146a:** Reduced levels contribute to increased type I interferon production.
- **miR-21:** Overexpression promotes B-cell hyperactivation, leading to autoantibody production.

3. Multiple Sclerosis (MS)

- **miR-155:** Drives neuroinflammation by enhancing Th1 and Th17 responses.

- **miR-326:** Overexpressed in MS, promoting the differentiation of pathogenic Th17 cells.

4. Type 1 Diabetes (T1D)

- **miR-146a and miR-21:** Dysregulation impairs pancreatic β-cell survival and contributes to T-cell-mediated destruction.

5. Inflammatory Bowel Disease (IBD)

- **miR-21:** Involved in promoting inflammation by targeting TGF-β receptors and several other regulatory proteins that control immune responses and apoptosis.

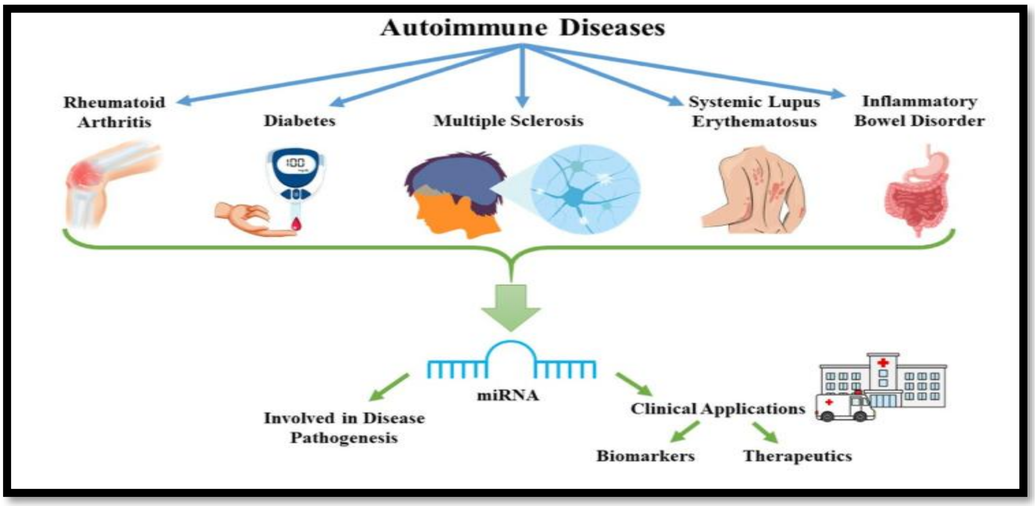


Figure 3: Overview of role of miRNA involved in the disease pathogenesis, biomarker detection and development of miRNA-based therapeutics for different autoimmune disease

Challenges and Future Directions

Despite significant advances in understanding miRNAs' role in immune responses and autoimmune diseases, several challenges persist. The complexity of miRNA-mediated regulation, redundancy in targets, and context-specific effects hinder precise therapeutic applications. Ensuring efficient, targeted delivery of miRNA-based therapies while minimizing off-target effects and immune responses remains a major obstacle. Additionally, inter-individual variability in miRNA expression complicates their use as reliable biomarkers.

Future research should focus on developing advanced delivery systems, such as nanoparticles, and leveraging CRISPR technology for precise miRNA modulation. Comprehensive profiling of miRNA networks and integrating multi-omics approaches can enhance personalized therapeutic strategies in immunology.

Conclusion:

miRNAs are critical regulators of immune responses and play a dual role in maintaining immune homeostasis and contributing to autoimmune diseases when dysregulated. Understanding their mechanisms provides valuable insights into immune regulation and offers potential therapeutic strategies for treating autoimmune disorders. Key miRNAs, such as miR-21, miR-155, and miR-34a, regulate critical pathways involved in T cell polarization, B cell activation, and cytokine production. For instance, miR-21 suppresses Th1-mediated inflammation, while miR-155 promotes pro-inflammatory cytokine expression. Similarly, miRNAs like miR-15/16 and miR-181a influence natural killer (NK) cell function and T cell aging, respectively, highlighting their roles in immune regulation across diverse cellular subsets.

In autoimmune diseases such as rheumatoid arthritis, psoriasis, and asthma, aberrant miRNA expression contributes to disease pathogenesis by affecting immune signaling pathways, including the NF-κB and PI3K-AKT pathways. miRNAs also mediate epigenetic modifications, further amplifying immune dysregulation. In conclusion, miRNAs serve as master regulators of immune responses and autoimmune diseases, offering novel insights into immune biology and opening avenues for innovative therapeutic strategies. Understanding their complex interactions with immune pathways is crucial for advancing precision medicine in immunology.

References:

1. Bofill-De Ros, X., & Vang Orom, U. A. (2024). Recent progress in miRNA biogenesis and decay. *RNA Biology, 21*, 1–8.
2. Chang, C., Xu, L., Zhang, R., Jin, Y., Jiang, P., Wei, K., Xu, L., Shi, Y., Zhao, J., Xiong, M., *et al.* (2022). MicroRNA-mediated epigenetic regulation of rheumatoid arthritis susceptibility and pathogenesis. *Frontiers in Immunology, 13*, 838884.
3. Chen, C. Z., Li, L., Lodish, H. F., & Bartel, D. P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. *Science, 303*, 83–86.
4. Clayton, S. A., Jones, S. W., Kurowska-Stolarska, M., & Clark, A. R. (2018). The role of microRNAs in glucocorticoid action. *Journal of Biological Chemistry, 293*, 1865–1874.
5. Fu, D., Yu, W., Li, M., Wang, H., Liu, D., Song, X., Li, Z., & Tian, Z. (2015). MicroRNA-138 regulates the balance of Th1/Th2 via targeting RUNX3 in psoriasis. *Immunology Letters, 166*, 55–62.
6. Hart, M., Walch-Ruckheim, B., Friedmann, K. S., Rheinheimer, S., Tanzer, T., Glombitza, B., Sester, M., Lenhof, H. P., Hoth, M., Schwarz, E. C., *et al.* (2019). miR-34a: A new player in the regulation of T cell function by modulation of NF-κB signaling. *Cell Death & Disease, 10*, 46.
7. Iwakawa, H. O., & Tomari, Y. (2022). Life of RISC: Formation, action, and degradation of RNA-induced silencing complex. *Molecular Cell, 82*, 30–43.
8. Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The *Caenorhabditis elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell, 75*, 843–854.
9. Lu, T. X., Hartner, J., Lim, E. J., Fabry, V., Mingler, M. K., Cole, E. T., Orkin, S. H., Aronow, B. J., & Rothenberg, M. E. (2011). MicroRNA-21 limits in vivo immune response-mediated activation of the IL-12/IFN-γ pathway, Th1 polarization, and the severity of delayed-type hypersensitivity. *Journal of Immunology, 187*, 3362–3373.

10. Matias-Garcia, P. R., Wilson, R., Mussack, V., Reischl, E., Waldenberger, M., Gieger, C., Anton, G., Peters, A., & Kuehn-Steven, A. (2020). Impact of long-term storage and freeze–thawing on eight circulating microRNAs in plasma samples. *PLoS ONE, 15*, e0227648.
11. Matsuyama, H., & Suzuki, H. I. (2019). Systems and synthetic microRNA biology: From biogenesis to disease pathogenesis. *International Journal of Molecular Sciences, 21*, 132.
12. Morales, S., Monzo, M., & Navarro, A. (2017). Epigenetic regulation mechanisms of microRNA expression. *Biomolecular Concepts, 8*, 203–212.
13. Plotnikova, O., Baranova, A., & Skoblov, M. (2019). Comprehensive analysis of human microRNA–mRNA interactome. *Frontiers in Genetics, 10*, 933.
14. Saliminejad, K., Khorram Khorshid, H. R., Soleymani Fard, S., & Ghaffari, S. H. (2019). An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. *Journal of Cellular Physiology, 234*, 5451–5465.
15. Sekar, D., & Lakshmanan, G. (2020). Methylation of N6-adenosine (m6A) modification in miRNAs and its implications in immunity. *Epigenomics, 12*, 1083–1085.
16. Simpson, L. J., Patel, S., Bhakta, N. R., Choy, D. F., Brightbill, H. D., Ren, X., Wang, Y., Pua, H. H., Baumjohann, D., Montoya, M. M., *et al.* (2014). A microRNA upregulated in asthma airway T cells promotes TH2 cytokine production. *Nature Immunology, 15*, 1162–1170.
17. Tong, N., Liu, D., Lu, L., Lin, R., & Jin, R. (2023). miR-410 regulates helper T cell differentiation in ovalbumin-induced asthma through the PI3K–AKT–VEGF signaling pathway. *International Archives of Allergy and Immunology, 185*, 1–9.
18. Treiber, T., Treiber, N., & Meister, G. (2019). Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nature Reviews Molecular Cell Biology, 20*, 5–20.
19. Wang, Z., Dai, R., & Ahmed, S. A. (2023). MicroRNA-183/96/182 cluster in immunity and autoimmunity. *Frontiers in Immunology, 14*, 1134634.
20. Zhu, S., Pan, W., & Qian, Y. (2013). MicroRNA in immunity and autoimmunity. *Journal of Molecular Medicine, 91*, 1039–1050.

MODELING AND DESIGNING OF GRID CONNECTED WIND POWER PLANT

Sonam Kalra* and Hitendra Kumar
Department of Electrical Engineering,
School of Engineering & Technology, IFTM University, Moradabad-244102, India
*Corresponding author E-mail: soni.jadoo@gmail.com

Abstract:
Wind energy is one of the important renewable energy resources. In India the total wind power potential is 45,000MW and the total installed capacity only approximate 10,000MW. There is a big difference in the potential and installed capacity. Renewable energy having a great advantage over the non-renewable (Thermal, Nuclear etc). In the renewable energy minimum running charges and less pollution content. If we use the any renewable energy resources in the stand-alone condition then it is not very much reliable because the renewable energy having weather dependent. To enhance the use of renewable energy the government provides the different type of subsidy to the different consumers. In the present scenario worlds faces the global warming problem. Six green gases are the major contributor in the global warming. In this chapter we designing the grid connected wind power plant. For the wind power plant, we use doubly fed induction generator. In this chapter we take a three model of grid connected wind power plant. In the first model we modeled a grid connected wind power plant in the balance condition. In the second model we modeled a grid connected wind power plant if any faults occur in the power system. In the final we model a modeled if we connect a facts controller (STATCOM) in the overall power system and check how the performance improves in the overall power system.

Keywords: Renewable Energy, MNRE, Wind Power, Grid

Introduction:

The utilisation of wind energy is an ancient practice, originating with sailing vessels and grain mills. The mechanical principles of ships were later adapted into windmills, as evidenced by the canvas sails of Cretan and Dutch designs. Only in the early twentieth century did the development of high-speed wind turbines for electricity generation gain serious attention. Windmills are generally classified into two categories: horizontal-axis and vertical-axis turbines. The term “wind generator system” was initially employed but subsequently abandoned, as it misleadingly suggested that the machine produced wind rather than electricity.

Throughout history, wind energy has been harnessed to grind grain, pump water from deep wells, and propel sailing boats. In pre-industrial Europe, windmills served diverse purposes. The modern impetus for wind energy development arose from escalating oil prices and concerns regarding the finite nature of fossil fuel resources.

Wind results due to the earth equatorial region receiving more solar energy that the polar region and this sets up large scale convection currents in the atmosphere. Meteorologists estimate incoming that about 1 per cent of the solar radiation is converted into wind energy.

The improvements due to technologies advancement in various area of wind turbine system are summarized in Table 1.

Table 1: Area of potential technology improvement

Sr. No.	Technical area	Potential advances
1	Advance tower concept	• Taller tower in difficult • New material process • Advance structure • Self erecting, initial, or for services
2	Advance rotor	• Advanced Material • Improved structure aero design • Active Control • Passive control • Higher tip speed
3	Reduced energy losses and improved availability	• Reduced blade soiling losses • Damage-tolerant sensor • Robust control system • Prognostic maintenance
4	Drive train	• Fewer gear stage • Medium speed generation • Distributed gear box topologies • Permanent-magnet generator • Medium voltage generator • Advance gear tooth profiles • New circuit topologies • New semiconductor devices New material
5	Manufacturing	• Sustained, incremental design and process improvement • Large-scale manufacturing • Reduced design loads

The availability of critical resources is crucial for large-scale manufacturing of wind turbines. The most important resources are steel, fiberglass, resins for composites and adhesives, blades core materials, permanents magnets and copper.

- Turbine materials usage is and will continue to be dominated by steel.
- The use of various composites and carbon filaments reinforced plastics might help reduce weight and cost
- Low cost of materials and high reliability remain the primary derives.
- Glass fiber reinforced plastic is expected to continue to be used for blades.
- Permanent magnet generators on large turbine will increase the need for magnetic materials.

Table 2 gives a representatives idea of the power requirements of some household applications.

Table 2: Household Application Details

Sr. No.	Item Description	Power, Watt	Daily energy Whr/Day
1	Incandescent light	60	1800-720
2	Fluorescent light	15	45-180
3	Refrigerator	80-500	2000-10000
4	Television	15-100	30-600
5	Village household	60-300	300-1200

Wind turbine performances depend primarily on rotor diameter and wind speed. Table 3 gives an estimated of the wind turbine output, based on the wind speed rotor diameter in Whr/day.

Table 3: Output of wind turbines

Rotor diameter (m)	Wind Speed m/s			
	4	5	6	7
1.5	35	70	94	117
3	152	269	386	468
5	421	761	1053	1287
7	831	1522	2107	2575

A wind turbine is a rotating machine which convert the wind kinetic energy into electrical energy and this mechanism is called a wind generator. Wind turbines can be separated into two types base by the axis in which the turbine rotate as Horizontal Axis Wind Turbines and Vertical Axis Wind Turbines.

Turbine subsystems comprise:

- The wind energy converted into mechanical energy;
- All the conversion gear, generator, mechanism shaft etc is assembled in the Nacelle.

Internal Components of a Wind Turbine consists:

- **Anemometer:** It is a device which is used for speed.
- **Blades:** These are aerodynamically designed structure that when wind flow over them they are lifted as in aircraft wings.
- **Brake:** The break is used for stop turbine.
- **Controller:** This is the most important part of the turbine it controls all the power production to pitch angle, and also produce the output power.
- **Gear box:** It is used for steps-up or steps down the speed of turbine and with appropriate coupling transmits rotating automatic energy at a suitable speed to the producer.
- **Generator:** It is used to convert the mechanical energy into electrical energy at 50 Hz frequency.
- **High-speed shaft:** Its purpose is to force the generator to rotate.
- **Low-speed shaft**: The rotor runs the low-speed beam at about 45 to 65 rotation per minute.
- **Nacelle:** Which contain all the gear system, generator and shaft system.

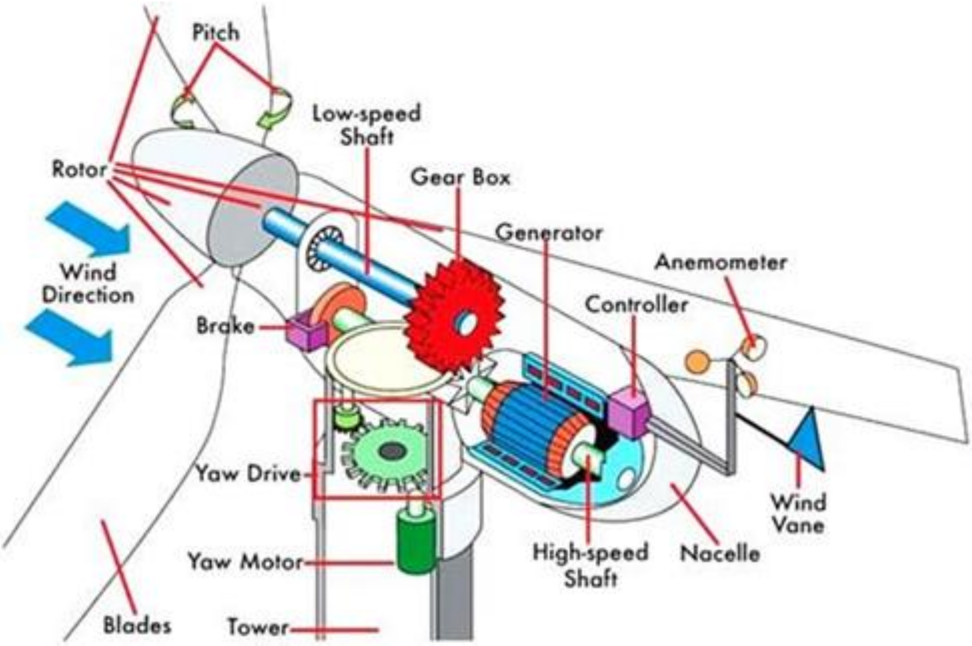


Figure 1: Basic Schematic of the Internal Components of a Wind Turbine

Modeling and Simulation

The Model of wind farm which is connected to the Grid is given below in Figure 2 In this model the wind farm is connect to the grid through the bus via a transformer which is further connect to the 30 km transmission line. On the other hand, this transmission line is connected to the grid through the transformer and the reactance with the star connected grounding. The response of this model is directly seen through the scope. This wind farm is consisting of 6 turbines of

1.5MW each. In this wind farm we choose the wind speed according to our convenience but in this model the wind speed is 15 km per hour.

The parameter of this wind farm is given below in Figure 3. This wind farm parameter gives the information about the number of wind farm, the number of wind generators, power rating, nominal powers, stator, rotor, magnetizing inductance, inertia constant, friction factors, pair of poles and initial condition etc.

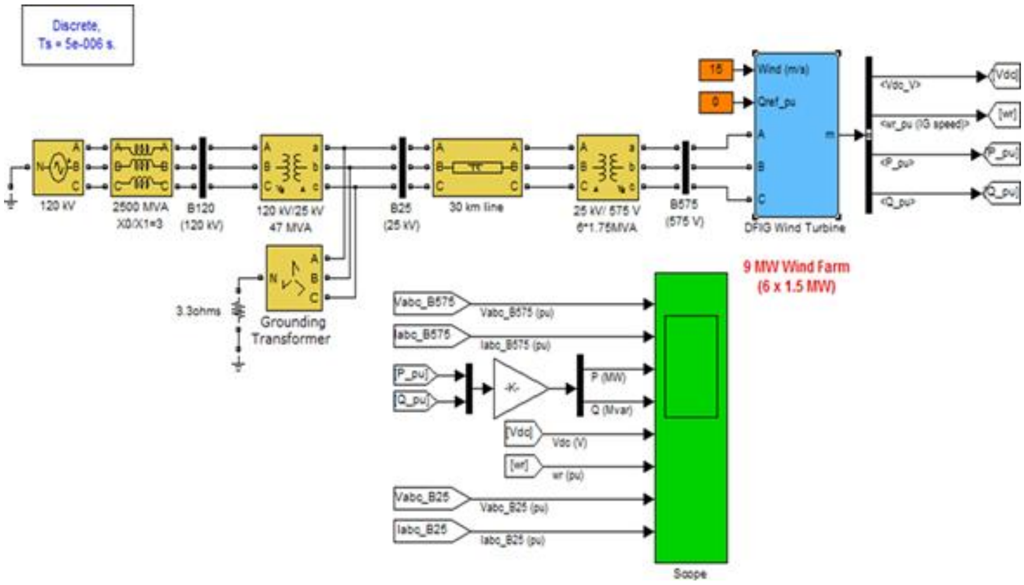


Figure 2: Model of Wind Farm in MATLAB

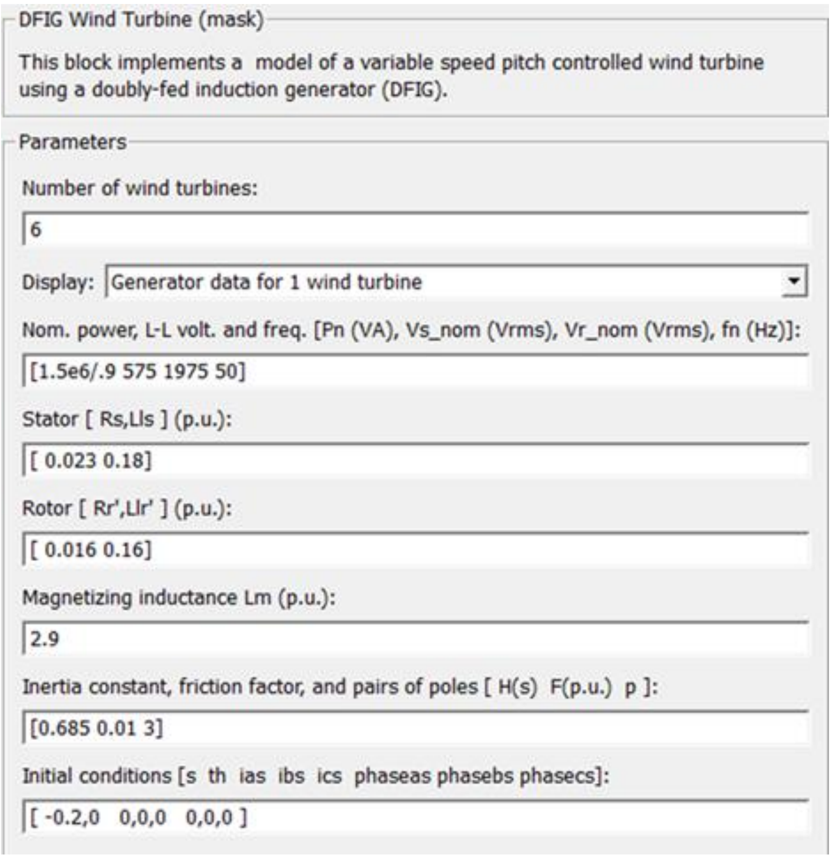


Figure 3: Parameters of Wind Farm

Results:

A. Output Results of Simple Grid Connected with Wind Power Plant

Figure 4 shows the different waveform at the different buses in the normal operating conditions. BUS 575 shows the three phase line voltage of the doubly fed induction generator whereas the BUS B 25 shows the grid side three phase voltage. In the balance condition the three phase voltage having the same magnitude and 120 degree phase displacement. But the overall system having a total unbalance load so the line voltage shows the different magnitude at the different instant. The Figure 4 also shows the line current at bus 575 and bus b 25, active power flow, reactive power flow and average voltage.

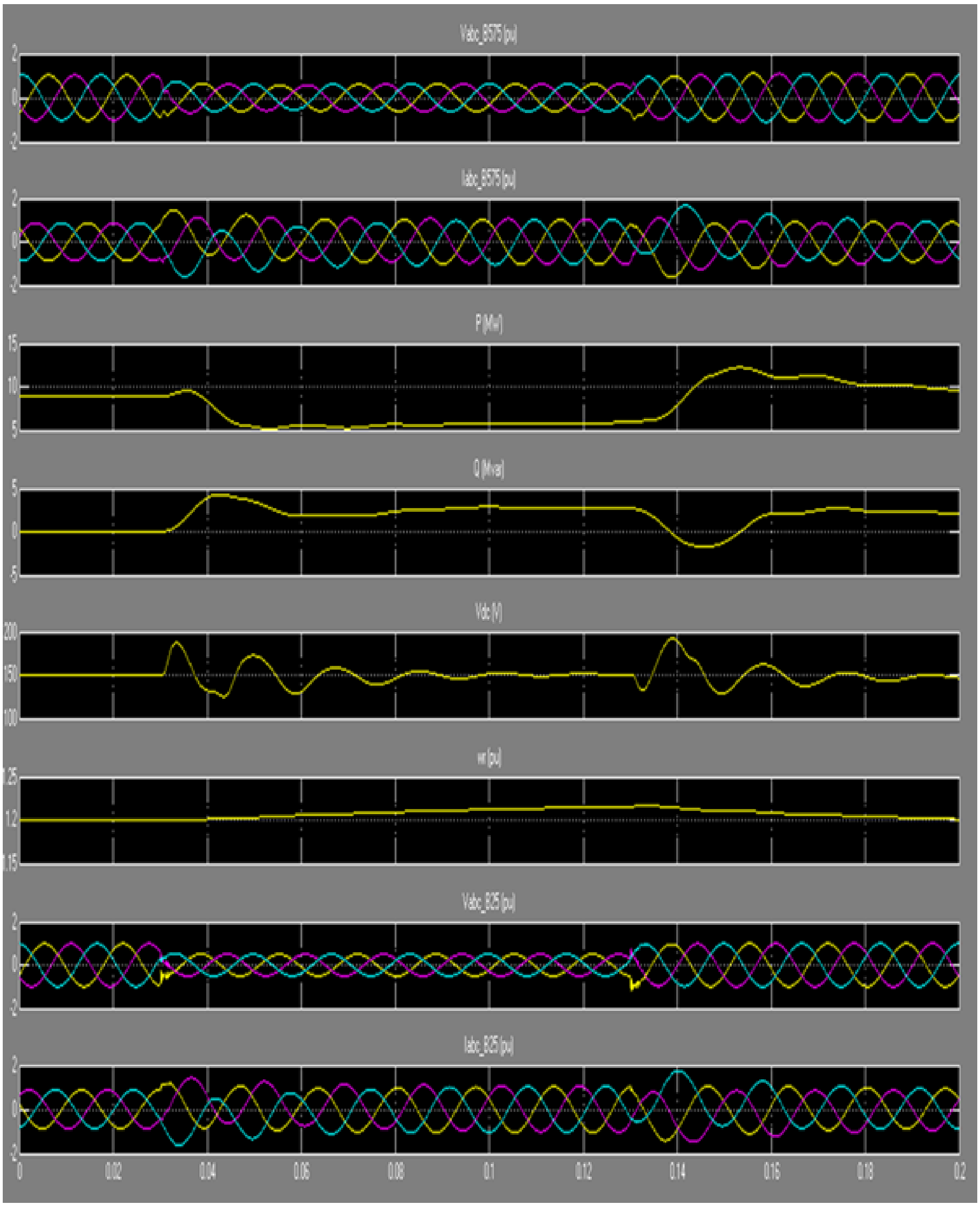


Figure 4: Output waveform in normal condition

B. Grid Output Results When Fault Generates

Figure 5 shows the grid-side electrical response under fault conditions, highlighting voltage sag and transient disturbances. It illustrates the degradation in grid stability immediately after fault inception.

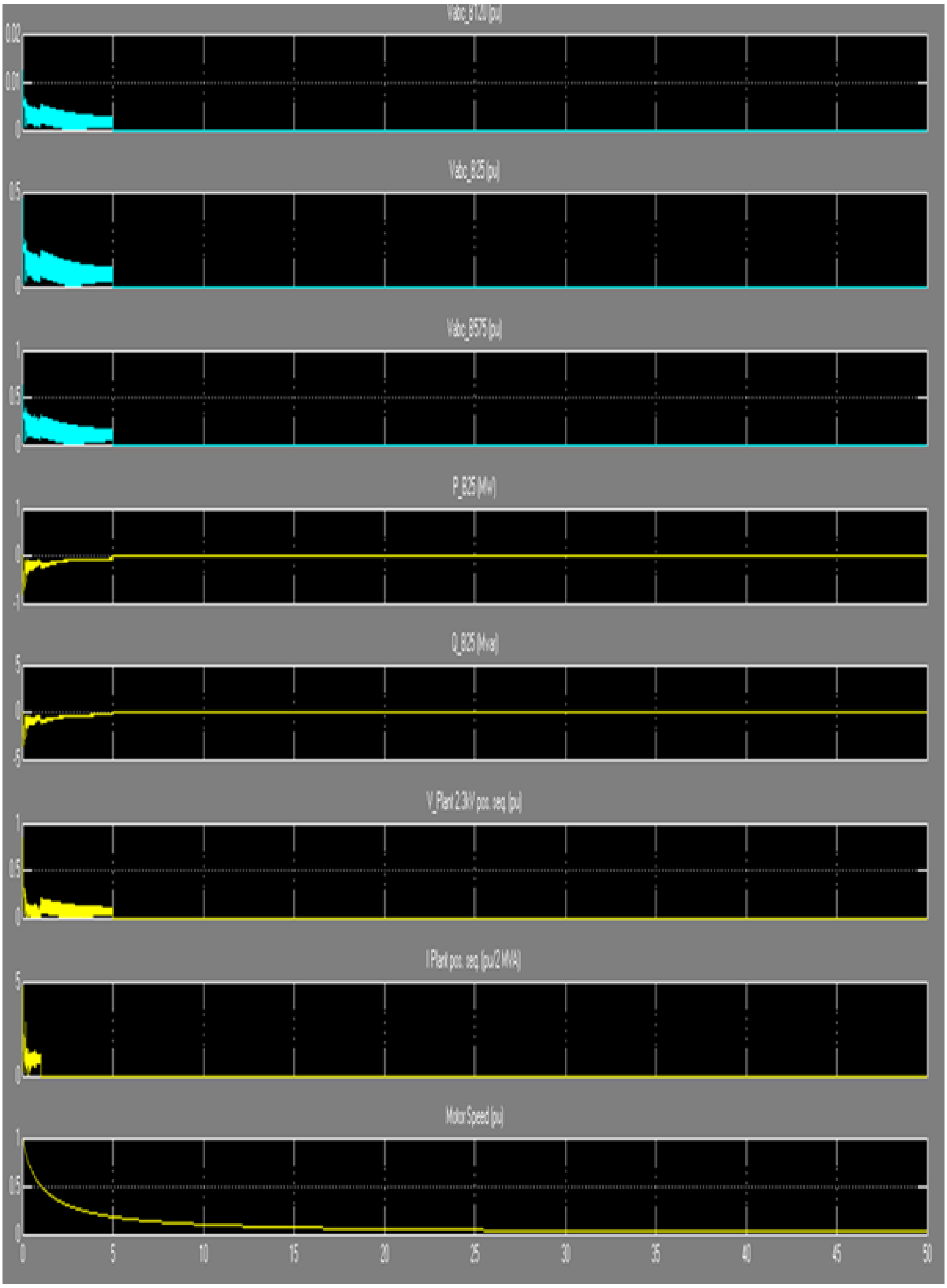


Figure 5: Grid output when faults occur

C. Wind Power Plant Output when Faults Generates

The waveform of Figure 6 depicts the impact of grid faults on wind power plant output, showing fluctuations in voltage/current and power. Reduced output quality reflects the sensitivity of the plant to external disturbances

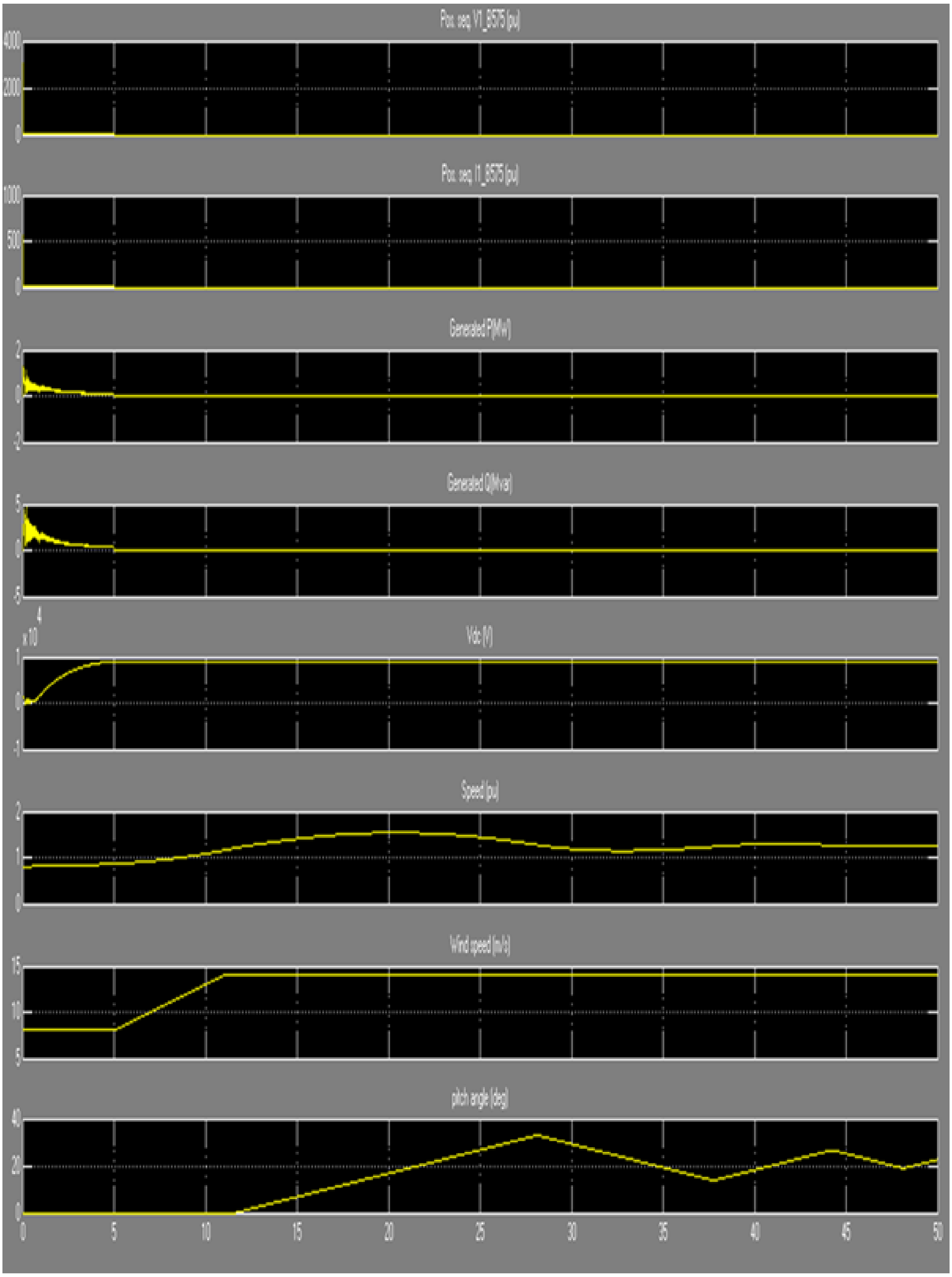


Figure 6: Wind power plant output when faults occurs

D. Output Waveform at Bus 25 STATCOM

This Figure 7 presents the compensated voltage/current waveform at Bus 25 with STATCOM operation. Improved waveform smoothness demonstrates effective reactive power support during fault conditions.

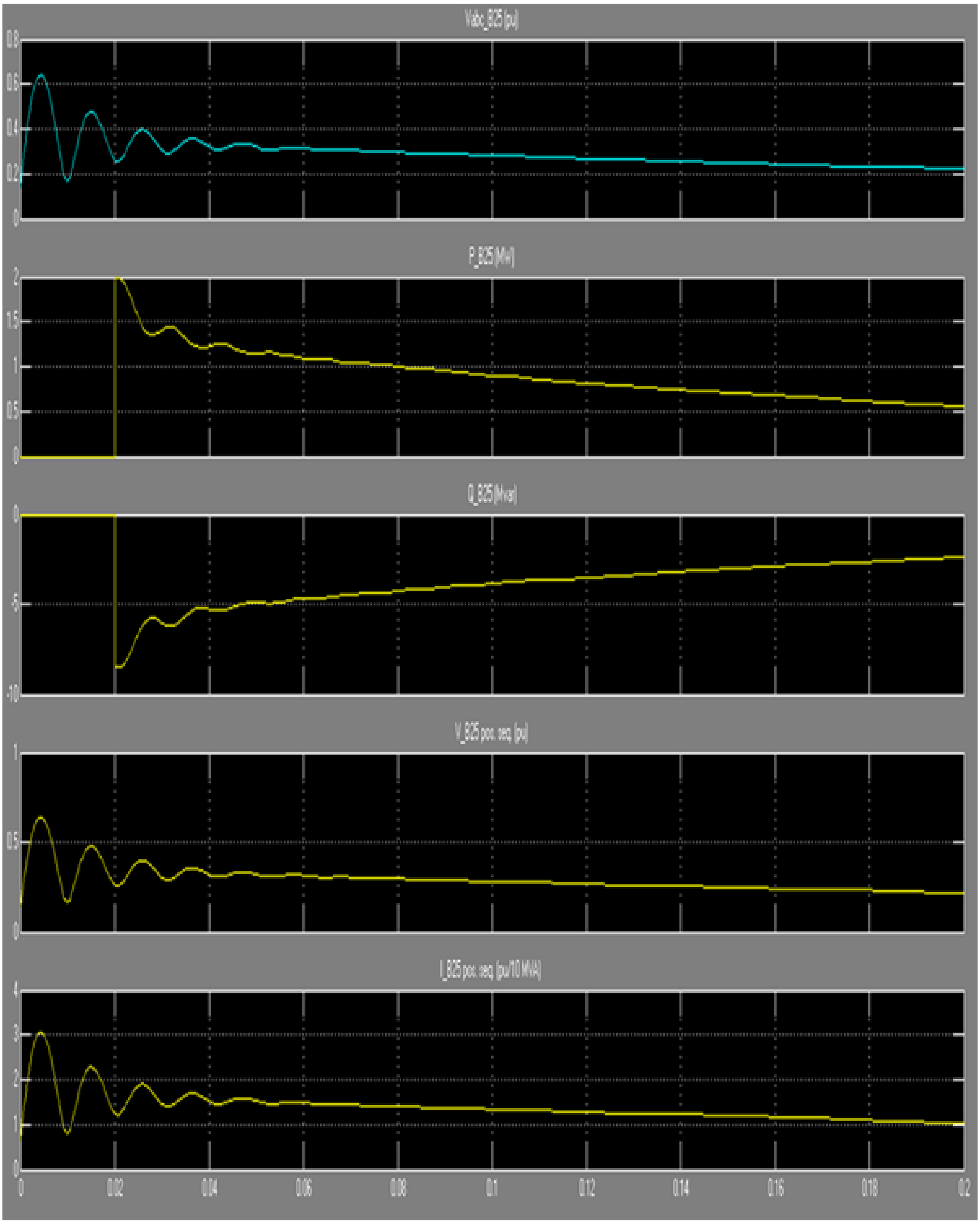


Figure 7: Output waveform at bus 25

E. Output Waveform of Wind Power Plant STATCOM

The waveform of Figure 8 illustrates STATCOM performance at the wind power plant terminals, showing mitigation of voltage dips and oscillations. It confirms enhanced power quality during faults.

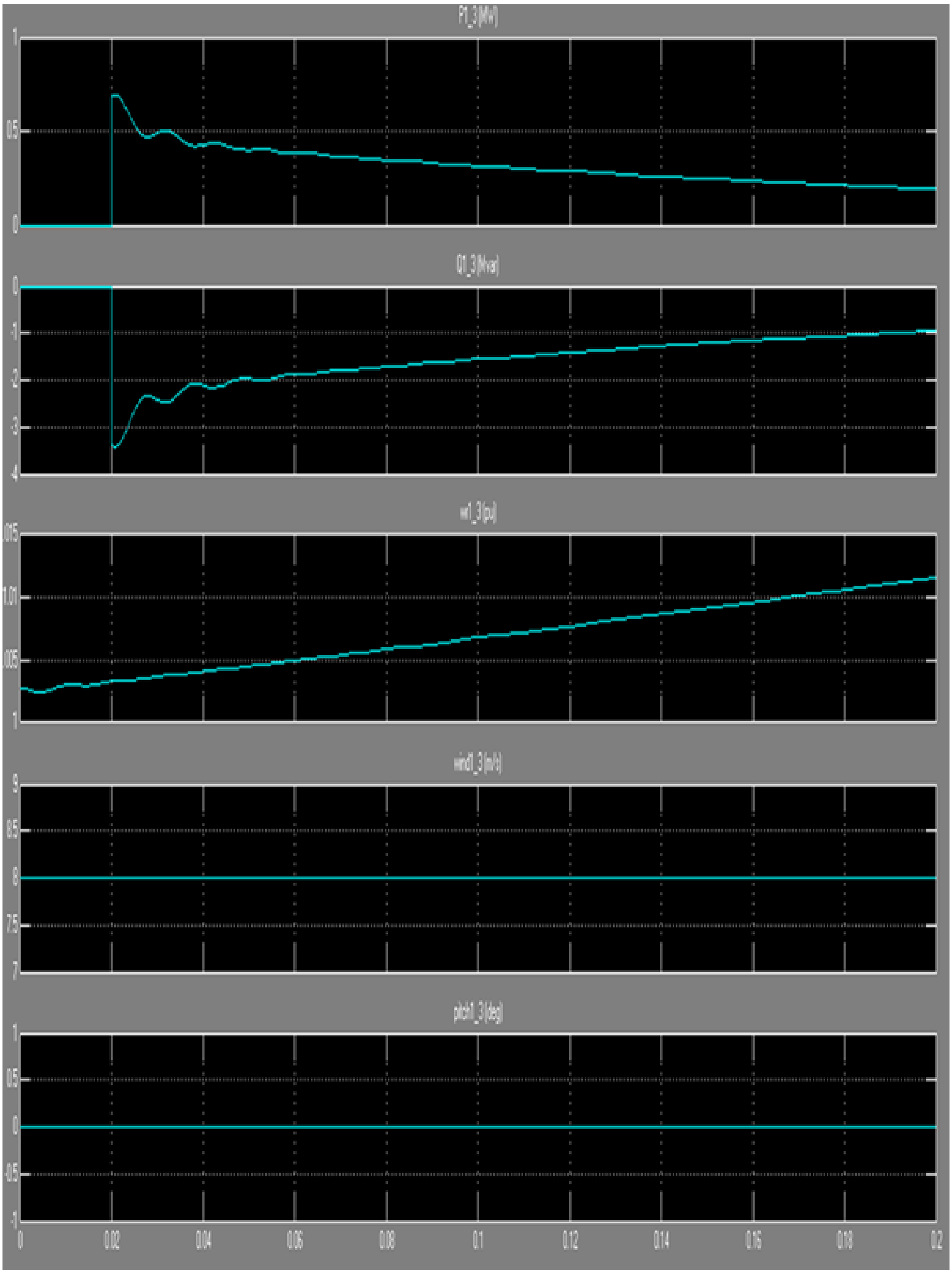


Figure 8: Output waveform of wind power plant STATCOM

F. Output Waveform of STATCOM

This Figure 9 shows the overall STATCOM output waveform, emphasizing rapid dynamic response and stabilization capability. The controlled waveform indicates effective voltage regulation and system support.

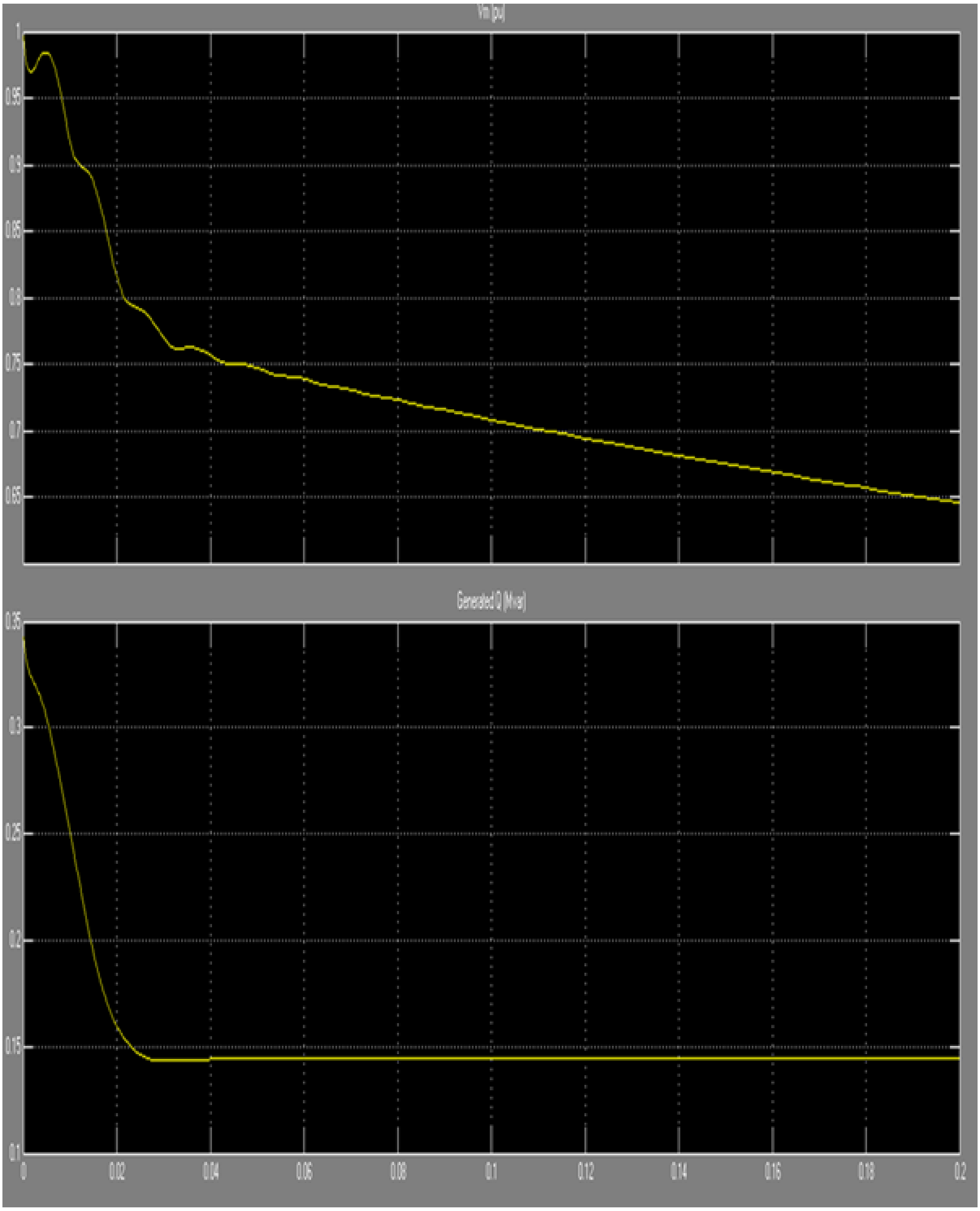


Figure 9: Output waveform of STATCOM

Conclusion

We have successfully analyzed grid connected wind power plant. In this we analyze by connecting the STATCOM and see the response in the scope. In second case we connect the fault in the transmission line (line- to-line fault) by which the response of the wind power plant becomes zero. Wind energy is the green energy source. The status of potential energy in India is good.

By the use of wind energy, we can minimize the global warming and increase the use of renewable energy, which comes under the "green Technology".

The wind (Double Fed Induction Generator) power plant is very useful for fulfil the future energy demand without affecting the environment.

References:

1. Nomandela, S., Mnguni, M. E. S., & Raji, A. K. (n.d.). Modeling and simulation of a large-scale wind power plant considering grid code requirements. *Energies*.
2. Chhipa, A. A., Chakrabarti, P., Bolshev, V., *et al.* (n.d.). Modeling and control strategy of wind energy conversion system with grid-connected DFIG. *Energies*.
3. Li, S., Qin, S., Wang, R., Li, Q., & Chen, C. (n.d.). Study on grid adaptability testing methodology for wind turbines. *Journal of Modern Power Systems and Clean Energy*.
4. Performance evaluation of grid-connected wind turbine generators (MDPI *Energies* article; multi-turbine modelling & simulation). (n.d.). *Energies*.
5. Datta, S., Deb, T., Nandi, C., & Chakraborty, A. K. (n.d.). Modelling and simulation of wind power plants with SEIG and DFIG in PSCAD/EMTDC. *International Journal of Electrical Engineering & Technology (IJEET)*.
6. Modelling and simulation of wind farm integrated with grid utilizing DFIG—MATLAB/Simulink based study. (n.d.).
7. Complete wind farm electromagnetic transient modelling for grid integration studies. (n.d.). *Energy Conversion and Management*.
8. Grid integration challenges of wind energy systems: A review. (n.d.). *IEEE Access*.
9. Dynamic equivalent model for DFIG-based wind farms. (n.d.). *IEEE Access*.
10. Sliding mode control for grid integration of wind power systems based on direct drive PMSG. (n.d.). *IEEE Access*.

DISCRETE MATHEMATICS AND GRAPH THEORY FOR EMERGING TECHNOLOGIES

Shobana A*[1], Vidhya D[2] and Logapriya B[2]
[1]Department of Science and Humanities, Nehru Institute of Technology, Coimbatore
[2]Department of Science and Humanities, Karpagam Institute of Technology, Coimbatore
*Corresponding author E-mail: shobanaamaths@gmail.com

Abstract:
Discrete mathematics and graph theory form the mathematical backbone of many emerging technologies, providing essential tools for modeling, analysis, and optimization of complex systems. This chapter explores fundamental concepts of discrete mathematics—including sets, relations, combinatorics, Boolean algebra, and discrete structures—and their integration with graph-theoretic principles such as connectivity, traversal, coloring, and network flows. Emphasis is placed on practical applications in modern technological domains, including communication networks, data structures, cryptography, machine learning, social network analysis, and optimization problems in engineering systems. Through illustrative examples and case studies, the chapter demonstrates how discrete models enable efficient problem-solving, scalability, and robustness in real-world applications. The discussion highlights recent advancements and interdisciplinary trends, underscoring the growing relevance of discrete mathematical techniques in the design and development of intelligent, secure, and scalable technological solutions.

Keywords: Discrete Mathematics, Graph Theory, Network Modeling, Optimization Algorithms, Emerging Technologies

1. Introduction:

The rapid advancement of science, engineering, and technology has led to the development of increasingly complex systems that require precise mathematical frameworks for analysis, modeling, and optimization. Among the various branches of mathematics, discrete mathematics and graph theory have emerged as indispensable tools for addressing problems involving finite, structured, and interconnected entities. Unlike continuous mathematics, which deals primarily with real-valued functions and smooth changes, discrete mathematics focuses on countable elements and logical relationships, making it particularly suitable for digital technologies and computational systems.

Discrete mathematics encompasses a broad range of topics, including set theory, logic, relations, combinatorics, number theory, and Boolean algebra. These foundational concepts play a crucial role in computer science, communication systems, cryptography, and algorithm design. For instance, combinatorial techniques are widely used to analyze algorithmic complexity, while Boolean algebra forms the basis of digital circuit design and logical reasoning in computing systems. As modern technologies increasingly rely on data-driven decision-making and automated processes, the relevance of discrete mathematical methods continues to grow.

Graph theory, a major subfield of discrete mathematics, provides a powerful framework for modeling relationships and interactions within complex systems. Graphs consist of vertices and edges that represent entities and their connections, respectively. This simple yet versatile structure enables the representation of a wide range of real-world systems, such as communication networks, transportation systems, social networks, biological networks, and power grids. Fundamental graph-theoretic concepts—including paths, cycles, connectivity, trees, and coloring—are central to the analysis and optimization of these systems. Advanced topics such as network flows, matching, and spectral graph theory further enhance the ability to solve large-scale and dynamic problems.

Emerging technologies such as artificial intelligence, machine learning, Internet of Things (IoT), blockchain, and cyber-physical systems increasingly rely on discrete models and graph-based representations. In machine learning, graphs are used to model relational data and structured dependencies, leading to the development of graph neural networks and network-based learning algorithms. In IoT and communication networks, graph theory aids in routing, resource allocation, and fault tolerance. Cryptographic systems and cybersecurity mechanisms also depend heavily on discrete mathematical structures to ensure data integrity, confidentiality, and secure communication.

The integration of discrete mathematics and graph theory with engineering and technological applications has led to innovative solutions that improve efficiency, scalability, and robustness. By providing formal tools for abstraction and analysis, these mathematical disciplines enable researchers and practitioners to design systems that are both theoretically sound and practically viable. This chapter aims to present a comprehensive introduction to the fundamental principles of discrete mathematics and graph theory, while emphasizing their relevance and applications in emerging technologies. Through conceptual explanations and illustrative examples, the chapter seeks to bridge the gap between mathematical theory and real-world technological challenges, highlighting the critical role of discrete structures in shaping the future of science and engineering.

2. Fundamentals of Discrete Mathematical Structures

Discrete mathematical structures form the foundation of many modern computational and engineering systems by providing formal tools to represent and analyze finite and countable objects. Unlike continuous structures, discrete structures deal with distinct, well-defined elements and their relationships, making them particularly relevant to digital technologies, algorithm design, and data organization.

At the core of discrete mathematics lies set theory, which provides a basic language for defining collections of objects and operations such as union, intersection, and complement. Sets are used to model data collections, state spaces, and solution domains in computing and engineering problems. Closely related to set theory are relations and functions, which describe associations between elements of sets. Relations are essential in modeling databases, ordering systems, and connectivity patterns, while functions represent deterministic mappings widely used in algorithms and system modeling.

Logic and Boolean algebra constitute another fundamental component of discrete mathematical structures. Propositional and predicate logic enable precise reasoning and decision-making processes, while Boolean algebra underpins the design of digital circuits and control systems. Logical operators and truth tables provide the basis for program correctness, verification, and automated reasoning.

Combinatorics deals with counting, arrangement, and selection of discrete objects. Techniques such as permutations, combinations, and the principle of inclusion–exclusion are crucial in analyzing algorithm efficiency, network configurations, and resource allocation problems. Combinatorial reasoning also plays a key role in probability theory and randomized algorithms.

Additionally, discrete number systems, including integers and modular arithmetic, are central to cryptography, coding theory, and secure communication. Concepts such as divisibility, congruences, and prime numbers support encryption algorithms and error-detection mechanisms.

Together, these discrete mathematical structures provide a rigorous framework for modeling complex systems, enabling efficient computation and robust problem-solving across emerging technological domains.

3. Essential Graph Structures and Theoretical Models

Graph theory provides a systematic framework for representing and analyzing relationships among discrete entities through well-defined structures and models. A graph consists of a set of vertices and a set of edges that capture pairwise connections between vertices. Depending on the nature of these connections, graphs may be classified into various types, each suited to different theoretical and practical contexts. Undirected graphs model symmetric relationships, while directed graphs (digraphs) represent asymmetric interactions such as information flow or dependency structures.

Several fundamental graph structures play a central role in theoretical modeling. Simple graphs exclude loops and multiple edges, whereas multigraphs allow multiple connections between the same vertices, making them useful in transportation and communication networks. Weighted graphs, in which edges are assigned numerical values, are employed to represent cost, distance, or capacity in optimization problems. Trees and forests, characterized by their acyclic nature, are essential in hierarchical modeling, data structures, and network design.

Theoretical graph models provide abstractions for complex systems. Complete graphs, bipartite graphs, and regular graphs offer idealized structures that simplify analysis and enable rigorous mathematical reasoning. Concepts such as paths, cycles, connectivity, and components are used to study reachability and robustness of networks. These essential graph structures and models form the basis for advanced graph algorithms and applications in emerging technologies, facilitating efficient analysis, optimization, and decision-making in interconnected systems.

4. Graph Algorithms and Computational Complexity

Graph algorithms are systematic procedures designed to process and analyze graph structures in order to solve practical and theoretical problems efficiently. These algorithms form a critical component of computer science, engineering, and applied mathematics, as many real-world systems—such as communication networks, transportation systems, and social networks—can be naturally modeled as graphs. The performance of graph-based solutions largely depends on both the algorithmic approach and the underlying computational complexity.

One of the most fundamental classes of graph algorithms involves graph traversal, including breadth-first search (BFS) and depth-first search (DFS). These algorithms explore vertices and edges systematically and are widely used for tasks such as connectivity analysis, cycle detection, and path enumeration. BFS is particularly useful for finding shortest paths in unweighted graphs, while DFS supports applications such as topological sorting and component identification.

Shortest path algorithms play a vital role in optimization and routing problems. Algorithms such as Dijkstra's algorithm, Bellman–Ford algorithm, and Floyd–Warshall algorithm are employed to determine minimum-cost paths under different constraints. Similarly, minimum spanning tree algorithms, including Prim's and Kruskal's algorithms, are used to design efficient and cost-effective networks by connecting all vertices with minimal total edge weight.

Computational complexity theory provides a framework for evaluating the efficiency of these algorithms in terms of time and space requirements. Complexity analysis typically uses asymptotic notation, such as Big-O notation, to describe how algorithm performance scales with input size. Some graph problems, such as shortest path computation and spanning tree construction, can be solved in polynomial time, while others—such as the traveling salesman problem and graph coloring—are computationally challenging and belong to the class of NP-hard problems.

Understanding graph algorithms and their computational complexity enables the selection of appropriate methods for large-scale and real-time applications. Efficient algorithm design and complexity analysis are therefore essential for developing scalable and robust solutions in emerging technologies that rely on graph-based models.

5. Graph-Theoretic Approaches to Networked and Intelligent Systems

Graph-theoretic approaches play a vital role in the modeling, analysis, and optimization of networked and intelligent systems, which are characterized by complex interactions among interconnected components. By representing system entities as vertices and their interactions as edges, graph theory provides a flexible and intuitive framework for capturing structural and functional relationships within diverse technological systems.

In communication and computer networks, graph models are used to analyze connectivity, routing efficiency, and fault tolerance. Nodes represent devices or routers, while edges denote communication links with associated capacities or costs. Graph-theoretic techniques such as shortest path analysis, network flow optimization, and connectivity assessment enable efficient data transmission and robust network design. Similarly, in power grids and transportation systems, graph models assist in optimizing resource distribution and improving system resilience.

Intelligent systems, including artificial intelligence and machine learning applications, increasingly rely on graph-based representations to model relational and structured data. Social networks, recommendation systems, and knowledge graphs use nodes to represent entities and edges to represent relationships, enabling advanced inference and pattern discovery. Graph-based learning models, such as graph neural networks, exploit both topological structure and node attributes to enhance predictive accuracy and interpretability.

Graph theory also supports decentralized and adaptive systems such as the Internet of Things and multi-agent systems. In these environments, graph models facilitate coordination, synchronization, and distributed decision-making among autonomous agents. By integrating graph-theoretic principles with algorithmic and computational methods, networked and intelligent systems can achieve greater scalability, efficiency, and robustness, highlighting the importance of graph theory in the development of next-generation technologies.

6. Applications of Discrete Mathematics in Emerging Technologies

Discrete mathematics plays a crucial role in the development and advancement of emerging technologies by providing rigorous frameworks for modeling, analysis, and optimization of systems composed of finite and structured elements. As modern technologies increasingly rely on digital computation and interconnected systems, discrete mathematical methods have become fundamental to their design and implementation.

In computer science and software engineering, discrete mathematics underpins algorithm design, data structures, and program verification. Concepts from logic, set theory, and combinatorics are used to analyze algorithmic efficiency, ensure correctness, and optimize computational processes. Graph-based models are central to database systems, compiler design, and operating systems, enabling efficient resource management and scheduling.

Cryptography and cybersecurity heavily depend on discrete mathematical structures such as number theory, modular arithmetic, and combinatorial designs. Encryption algorithms, digital signatures, and secure key exchange protocols rely on properties of prime numbers, congruences, and discrete logarithms to ensure data confidentiality, integrity, and authentication in digital communications.

In artificial intelligence and machine learning, discrete mathematics supports knowledge representation, decision-making, and learning from structured data. Graphs and logical models are used in expert systems, knowledge graphs, and constraint satisfaction problems. Discrete optimization techniques also play a key role in feature selection, clustering, and combinatorial learning tasks.

Emerging domains such as the Internet of Things (IoT), blockchain, and smart systems further highlight the importance of discrete mathematics. IoT networks use graph models for topology control and routing, while blockchain technologies rely on cryptographic hash functions, consensus algorithms, and distributed graph structures. In smart cities and cyber-physical systems, discrete models enable efficient planning, scheduling, and control of interconnected infrastructures. Overall, discrete mathematics provides the theoretical foundation necessary for innovation, scalability, and security in modern technological systems.

Conclusion:

Discrete mathematics and graph theory form a critical foundation for understanding and developing emerging technologies. By providing structured methods for modeling finite systems, relationships, and networks, these mathematical disciplines enable efficient analysis, optimization, and decision-making across diverse technological domains. From algorithm design and cybersecurity to intelligent systems and networked infrastructures, discrete models support scalability, robustness, and innovation. As technologies continue to evolve toward greater complexity and interconnectivity, the importance of discrete mathematical frameworks will further increase. A strong grasp of these concepts equips researchers and practitioners to address future challenges and contribute effectively to advancements in science, engineering, and technology.

References:

1. Balasubramaniam, P., *et al.* (Eds.). (2024). *Discrete mathematics and mathematical modelling in the digital era: ICDM3DE-2023*. Springer Singapore.
2. Hoffman, F., *et al.* (Eds.). (2025). *Combinatorics, graph theory and computing: SEICCGTC 2023*. Springer Nature.
3. Shrimali, N. P., & Shah, N. H. (Eds.). (2021). *Recent advancements in graph theory*. Routledge.
4. Yadav, S. K. (2023). *Discrete mathematics with graph theory*. Springer Cham.
5. Žerovnik, J., & Rupnik Poklukar, D. (Eds.). (2022). *Advances in discrete applied mathematics and graph theory*. MDPI Books.
6. *Graph theory with applications*. (2024). *Discrete Applied Mathematics, 357*. Elsevier.
7. Alqahtani, M. H. (2023). New topologies on a graph. *Advances and Applications in Discrete Mathematics, 37*, 77–93.
8. Chawla, P. (2025). An interdisciplinary review on application of graph theory. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*.
9. Special issue: Recent advances in graph theory and its applications. (2025). *Axioms*. MDPI.

DESIGN AND IMPLEMENTATION OF SENSORLESS FIELD-ORIENTED CONTROL FOR 3-PHASE PMSM

R. Ashwin, M. Mohanraj, M. Ranjith and R. K. Padmashini
Department of Electrical and Electronics Engineering
AMET Deemed to be University, Chennai, India
Corresponding author E-mail: ashwin052005@gmail.com,
mohanrajmuthukumaran46@gmail.com, ranjithmohanraj28@gmail.com,
padmashini@ametuniv.ac.in

Abstract:

The design and implementation of a sensorless Field-Oriented Control (FOC) scheme for a three-phase Permanent Magnet Synchronous Motor (PMSM) are presented. The objective is to achieve high-performance motor control without the use of mechanical position or speed sensors, thereby reducing system cost and complexity. A mathematical model of the PMSM is developed in the d–q reference frame to enable decoupled control of torque and magnetic flux. The sensorless control approach is based on back electromotive force (back-EMF) principles and observer-based estimation techniques, such as the Sliding Mode Observer (SMO) or Extended Kalman Filter (EKF), for estimating rotor position and speed. The control system incorporates Clarke and Park transformations, Pulse Width Modulation (PWM), and Proportional–Integral (PI) controllers for current regulation. Both simulation and real-time hardware implementation are carried out under various load and speed conditions using a real-time interface. Experimental results demonstrate fast dynamic response, low torque ripple, and stable operation across a wide speed range. The proposed sensorless PMSM drive exhibits improved reliability, efficiency, and compactness compared to conventional PMSM drives, making it suitable for applications in electric vehicles, robotics, and industrial automation systems.

Keywords: Permanent Magnet Synchronous Motor (PMSM), Field-Oriented Control (FOC), Sensorless Control, Sliding Mode Observer (SMO), Extended Kalman Filter (EKF), Pulse Width Modulation (PWM).

Introduction:

One more popular solution is now Permanent Magnet Synchronous Motors (PMSMs) that find application in the high performance drive application of high power density, efficiency as well as good dynamic properties. They have been utilized in electrically powered cars, robots, aerospace and industrial automation systems needing accuracy in the speed and torque management. To execute the precise control of torque and flux, Field-Oriented Control (FOC) has become extensive practice with mechanical sensors such as encoders or resolvents measuring rotor position and velocity inflating system expenses, size and complexity. The use of sensors also introduces the issue of reliability in harsh environments such as those with vibration, temperature or dust. To minimize the hurdles sensor less control algorithms have been developed in which the position and speed of the rotor is estimated by estimating the electrics in the motor regardless of stator voltages and current and utilizing mathematical models and estimation algorithms. These techniques render the systems more resilient, reduce the maokintenance, and reduce the size of the drive to smaller and less costly. The sensorless field-oriented control of a three-phase PMSM attempts to exploit a robust and useful control framework, which is capable of estimating rotor position without mechanical sensors, and at a spectrum of stable operation across a wide range of speeds. The study has been grounded on mathematical modeling of the PMSM, current and speed regulation loop design and application of sophisticated estimation algorithms such as, Model Reference Adaptive System (MRAS), Sliding Mode Observer (SMO) or even the Extended Kalman Filter (EKF), to provide a real-time estimate of rotor position throughout its operation. The results of the experimental validation and simulations of the system performance are related to attributes of the torque response, speed regulation and stability when the system is perturbed by the load and are therefore applicable in a construction of the cost-effective, reliable and high-efficiency PMSM drives, which can be utilized in the contemporary applications such as electric vehicles and the renewable-energy systems. The sensorless FOC proposed cuts down on the dimensions of the system, its dependability, and sustainability, which is vital in serving the primary objective of future motor drive technology that is concerned with the elimination of mechanical sensors.

Literature Review

Hedberg *et al.* [1], in their research, addressed the problem of controlling a Permanent Magnet Synchronous Motor (PMSM) using the Field-Oriented Control (FOC) method and considered a cascaded Proportional–Integral–Derivative (PID) controller structure. The control strategy employs three hierarchical control loops—position, speed, and torque—combined with feed-forward control, mode switching, and limiters to achieve improved system stability and response. The controller was implemented on real-time hardware using C programming and recursive filtering, which proved to be efficient in code generation and demonstrated good hardware performance. Experimental results established that the cascaded PID approach outperforms traditional methods in terms of system performance, as each control layer can be optimized independently, thereby simplifying the design process and providing a better control hierarchy. This approach is feasible because PID controllers are simple, reliable, and easy to implement. Real-time validation further enhances the practical applicability of the study. However, the limitations of classical PID control reduce system flexibility under nonlinear operating conditions, line variations, or extremely low-speed operation, which are characteristic of PMSM drives. Moreover, the use of multiple control loops increases tuning complexity, and the study is restricted to a specific hardware setup. Additionally, the paper lacks quantitative comparison with more advanced control techniques such as adaptive or sensorless methods, which limits its broader industrial applicability.

Koc *et al.* [2] presented a unified drive system architecture for various Permanent Magnet Synchronous Motors (PMSMs) based on the Field-Oriented Control (FOC) approach. The proposed system adopts a flexible plug-and-play controller configuration that eliminates the need for redesign when switching between different PMSM types. It incorporates automatic tuning and decoupling compensation to maintain stable operation. Key nonlinearities, including inverter imperfections, DC-link voltage variations, and parameter drifts, are integrated into the model, enabling a realistic evaluation of their impact on motor performance parameters such as torque ripple, efficiency, harmonic distortion, and dynamic response. The control framework was validated through simulation studies under different operating conditions, demonstrating robustness and applicability across a wide range of environments while reducing development effort and providing comprehensive insight into motor behavior under nonlinear conditions. Nevertheless, the study is primarily limited to simulation results, with no experimental validation, thereby restricting its practical relevance. Real-time challenges such as sensorless operation, temperature variations, and online parameter estimation are not addressed. Furthermore, although the system is claimed to be applicable to all PMSM types, its performance under extreme operating speeds and load variations requires further clarification. Consequently, while the study provides a strong theoretical foundation, it requires validation through hardware implementation.

Shallal *et al.* [3] presented a study focused on the development of a Field-Oriented Control (FOC) strategy for a three-phase Permanent Magnet Synchronous Motor (PMSM) using Hall-effect sensors to detect rotor position and regulate speed. The primary objective was to maintain the stator current at a 90° electrical angle relative to the rotor flux in order to achieve optimal torque per unit current. Stator current measurements are supplied to a flux observer to compute the d–q axis components, magnetizing current, and the synchronous reference frame angle. A Proportional–Integral (PI) controller is employed in the speed control loop as an adaptive mechanism to accommodate load variations in real time. Simulation results demonstrate that appropriate PI tuning yields good transient performance and minimal steady-state speed error. The adaptive PI controller enhances system stability and responsiveness under varying load conditions. Furthermore, the inclusion of realistic inverter and load models provides valuable insight into system performance and efficiency. However, the study is limited to simulation-based validation and does not address real-world challenges such as noise, inverter imperfections, and parameter variations. Since the control scheme relies on Hall sensors, it is not sensorless and may encounter performance issues at very low speeds where sensor signals become weak or distorted. Additionally, although the PI controller exhibits adaptability, its performance under extreme disturbances, temperature variations, and long-term operation is not fully evaluated, which limits the generalizability of the results to industrial applications.

The study by Nustes *et al.* [4] introduced an open-source dataset derived from the simulation of a Field-Oriented Control (FOC) model for a three-phase Permanent Magnet Synchronous Motor (PMSM). The system was modeled in MATLAB/Simulink, and both the dataset and simulation model are publicly available through GitHub and Mendeley Data repositories. The dataset captures motor responses to various input speed profiles, including step, ramp, random, and extreme scenarios, to evaluate the performance limits of a linear Proportional–Integral–Derivative (PID) controller embedded within the FOC framework. It includes signals from both the inner control loops, such as d–q axis currents and voltages, and the outer speed control loop, making it a valuable resource for researchers aiming to develop advanced nonlinear or intelligent control algorithms without requiring immediate access to physical test setups. The inclusion of extreme operating conditions supports the design of more robust and adaptive control strategies. The open-source nature of the dataset enhances transparency, reproducibility, and usability for both academic and industrial research. However, as the dataset is generated entirely through simulation, it does not account for real-world factors such as electrical noise, inverter non-idealities, thermal effects, or component aging. Moreover, the results are based on a specific PMSM configuration and FOC/PID design, which may limit their applicability to other motor types or control architectures. Although the dataset facilitates exploration of machine learning and advanced control methods, the absence of hardware validation restricts confirmation of its practical effectiveness.

Bhatt *et al.* [5] investigated the development and application of sensorless control techniques for brushless permanent-magnet (BPM) machines in power tool applications. The proposed approach eliminates the need for mechanical or Hall-effect sensors traditionally used for rotor position and speed measurement. Instead, rotor position and speed are estimated using electrical quantities such as back electromotive force (back-EMF) and current signals. The main objective of this sensorless strategy is to reduce system cost, enhance reliability, and minimize maintenance requirements in demanding power tool environments. The control system is designed with emphasis on durability, compactness, and cost-effectiveness, making it suitable for high-speed and dynamic operating conditions. The study covers drive system modeling, sensorless estimation algorithms, and validation through simulations or prototype testing under practical operating scenarios. These features contribute to reduced system complexity, simplified wiring, and improved reliability. The approach is particularly well suited for harsh environments where compact design and robustness are essential, as in power tool applications. Reduced maintenance requirements and improved operational efficiency further support its practical relevance. However, sensorless estimation based on back-EMF becomes challenging at low speeds due to weak signal levels, potentially affecting control accuracy. Additionally, variations in system parameters caused by temperature changes, magnetic aging, or inverter nonlinearities can degrade estimation performance. The study also appears to lack extensive large-scale hardware validation, relying mainly on prototype-level testing, which limits its industrial scalability. Furthermore, the methodology is specifically tailored to power tool applications, restricting its direct applicability to broader industrial motor drive systems.

Ilten *et al.* [6] investigated the development of an Extended Electromotive Force (EEMF)–based conformable fractional-order observer for sensorless control of three-phase Permanent Magnet Synchronous Motors (PMSMs). The proposed observer employs fractional-order calculus using a conformable derivative operator, which offers greater flexibility in modeling the dynamic behavior of the motor compared to conventional integer-order approaches. Stability of the observer is ensured through Lyapunov-based analysis, while controller parameters are optimized using the Particle Swarm Optimization (PSO) technique. The study compares the performance of the fractional-order EEMF observer with that of a traditional integer-order observer under various operating conditions, including startup, step response, and ramp speed variations. Experimental results demonstrate that the fractional-order approach provides smoother operation, reduced oscillations, and more accurate rotor position and speed estimation than the conventional method. The effectiveness of the proposed system is validated experimentally. Furthermore, the combined use of Lyapunov stability analysis and optimal parameter tuning strengthens both the theoretical and experimental foundations of the sensorless control scheme. However, the use of fractional-order calculus and optimization techniques increases computational complexity and tuning difficulty compared to classical observers, potentially limiting broader applicability. In addition, the study does not comprehensively evaluate performance under challenging conditions such as extreme disturbances, temperature variations, component aging, or very low- and zero-speed operation, which remain critical issues in practical PMSM applications.

Abdubannaev *et al.* [7] proposed a strategy to enhance the operational performance of electric machines through the application of the Field-Oriented Control (FOC) algorithm. The study emphasizes that FOC enables precise control of torque and flux in electric motors, including Permanent Magnet Synchronous Motors (PMSMs) and induction motors, by independently regulating the direct-axis (d-axis) and quadrature-axis (q-axis) currents. This decoupling approach allows continuous torque generation, faster dynamic response, improved efficiency, and operating characteristics comparable to those of DC motors. The research combines simulation and experimental studies to compare traditional scalar control with FOC. The results indicate that FOC provides superior torque behavior, improved speed regulation, and enhanced system stability under varying load and speed conditions. Key advantages include improved dynamic performance, reduced torque ripple, accurate torque control, and better energy efficiency, enabling stable variable-speed operation with faster response times. However, FOC requires accurate knowledge of motor parameters and rotor position information, which increases system cost and complexity. Additionally, implementing sensorless FOC at low speeds remains challenging, and the algorithm demands higher computational resources than simpler scalar control techniques.

The doctoral dissertation by De Soysa *et al.* [8] presented an advanced sensorless drive system for Permanent Magnet Synchronous Motors (PMSMs) that estimates rotor position and speed without the use of physical sensors such as encoders or Hall-effect devices. The proposed approach is based on back electromotive force (back-EMF) estimation combined with model-based observer techniques to achieve real-time rotor position estimation. Careful tuning of the estimation algorithm ensures stable performance across a wide speed range, effectively addressing the limitations of conventional sensorless control methods at both low and high speeds. The research includes comprehensive modeling, simulation, and experimental validation using prototype PMSM drives to evaluate the reliability, accuracy, and efficiency of the proposed control system under various load and dynamic operating conditions. The elimination of physical sensors contributes to reduced system cost and improved robustness. The results indicate enhanced control stability, motor efficiency, and overall performance across a broad operating range. However, the method is highly dependent on accurate motor parameter estimation and is sensitive to variations caused by temperature changes or resistance drift. Moreover, performance degradation at very low speeds, where back-EMF signals are weak, limits its suitability for applications requiring high-precision control near zero speed.

Gholipour *et al.* [9] developed a sensorless Field-Oriented Control (FOC) strategy for three-phase induction motor drives that is capable of handling current sensor faults while maintaining system performance. The proposed fault-tolerant control scheme ensures safe and reliable operation even when one or more current sensors fail. This is achieved through model-based estimation and observer algorithms that reconstruct missing or corrupted current signals without requiring additional sensing hardware. The method maintains accurate torque and speed regulation, thereby improving system reliability. Both simulation and experimental results demonstrate stable and effective operation under healthy and faulty conditions, highlighting its suitability for industrial applications that demand high dependability. The approach enhances fault tolerance and operational continuity while reducing system cost and design complexity by eliminating the need for redundant hardware. Additionally, it improves safety in mission-critical applications. However, the control strategy involves complex computational modeling and careful parameter optimization, which increases design complexity. Its performance may also degrade under extreme parameter variations or severe fault conditions. Furthermore, experimental validation is limited to laboratory environments, and additional large-scale testing is required to confirm its applicability in full-scale industrial systems.

Phuong *et al.* [10], in their work titled *Field-Oriented Control (FOC) of Permanent Magnet Synchronous Motors in MATLAB & Simulink*, presented a detailed study on the simulation and control of Permanent Magnet Synchronous Motors (PMSMs) using the Field-Oriented Control (FOC) strategy. The study explains the mathematical modeling of PMSMs and demonstrates how vector control techniques can be applied to achieve effective regulation of torque and magnetic flux. The control framework designed and simulated in MATLAB/Simulink incorporates Clarke and Park transformations, conventional controllers, and a Pulse Width Modulation (PWM) inverter. Simulation results indicate that the FOC approach significantly enhances dynamic performance, operational efficiency, and smooth motor operation, making it suitable for applications such as electric vehicle propulsion systems and industrial automation. The study further demonstrates that FOC enables precise torque and speed control with reduced torque ripple, thereby improving overall motor performance. The MATLAB/Simulink environment allows flexible parameter tuning and comprehensive system testing, contributing to the accuracy and reliability of the simulation results. However, the effectiveness of the method relies on accurate motor parameter knowledge, and the complexity of the control algorithm increases computational requirements. A major limitation of the study is its reliance solely on simulation without experimental hardware validation. Additionally, aspects such as real-time implementation, sensorless operation, performance under variable load conditions, and the impact of noise are not thoroughly investigated.

Baral *et al.* [11] discussed the advantages of a closed-loop Field-Oriented Control (FOC) system for PMSMs, highlighting its ability to achieve accurate torque and speed control through continuous feedback, resulting in efficient and smooth motor operation. The use of Hall sensors provides a simple and cost-effective method for rotor position detection, improving control accuracy, system stability, and performance across varying load conditions. This approach is particularly effective at low to moderate speeds, where consistent torque production is critical. However, Hall sensors offer lower resolution compared to optical encoders, which reduces control precision at higher speeds. The inclusion of sensors also increases system complexity, wiring requirements, and overall hardware cost. The study primarily focuses on low- to medium-speed operation, and performance degradation at very high rotational speeds due to inherent Hall sensor limitations is noted. Furthermore, advanced features such as sensorless operation and fault-tolerant control strategies are not addressed, which limits the applicability of the proposed approach in more advanced or industrial PMSM drive systems.

Zhu *et al.* [12] proposed a passive fault-tolerant control (PFTC) scheme for a 2×3-phase Surface Permanent Magnet Synchronous Motor (SPMSM) driven by a single inverter using the Field-Oriented Control (FOC) approach. The study demonstrates that the proposed control strategy maintains stable torque production and reliable operation even under phase or winding fault conditions. Both simulation and experimental analyses were conducted to validate the feasibility of the approach, showing improved fault tolerance, smoother torque response, and enhanced efficiency without requiring additional hardware or operational reconfiguration. The proposed method offers advantages such as improved system stability, enhanced fault protection, and continuous operation during fault conditions. However, the control structure is more complex than conventional FOC schemes and requires accurate fault detection and precise motor parameter estimation. The study primarily addresses open-circuit faults, while short-circuit fault scenarios receive limited attention. Additionally, the complexity of the implementation may pose challenges for real-time industrial applications, and the experimental validation is limited to specific operating conditions.

Nicola *et al.* [13], in their study titled *Real-Time Implementation of the PMSM Multi-Motors Sensorless Control System*, presented a real-time sensorless control strategy for coordinating multiple Permanent Magnet Synchronous Motors (PMSMs) without the use of physical position sensors. The proposed approach employs sensorless Field-Oriented Control (FOC), where rotor position and speed are estimated using mathematical models and algorithms instead of mechanical sensors. This reduces hardware complexity, lowers system cost, and enhances reliability. The effectiveness of the control strategy was validated through both simulation and experimental studies under varying load and speed conditions. Results indicate that the proposed method successfully coordinates multiple PMSMs, ensuring stable torque production, consistent speed regulation, and improved operational efficiency in real-time applications. Key benefits include reduced maintenance requirements, lower hardware costs, and improved consistency in multi-motor control. However, the accuracy of the control system is highly dependent on precise motor parameter estimation, and the computational complexity of the algorithm is higher than that of conventional controllers. The study also lacks extensive evaluation under dynamic disturbances or fault conditions, and issues related to large-scale system scalability, temperature-induced parameter variations, and long-term operational robustness are not thoroughly examined.

The paper of Tawfik *et al.* [14] titled sensorless control with optimized pi controller parameters with permanent magnet synchronous motor is an improved sensorless control of Permanent magnet Synchronous Motor (PMSMs) using the Model Reference Adaptive System (MRAS) state. The paper has a goal of optimizing the Proportional-Integral (PI) controller parameters in order to obtain improved dynamic performance, stability and accuracy in estimating the rotor speed and position. The proposed solution has the advantage of eliminating the reliance on mechanical sensors, reduces the complexity of hardware and its cost, and guarantees precise control of torque and speed. All these advantages of the proposed method are assessed using MATLAB/Simulink simulation that has eliminated the usage of sensors, the better performance of the proposed system in terms of the better transient behavior or the minimization of the steady-state error as well as the better adaptability and stability of the system under different load conditions. The performance however depends pretty strongly on precise system modeling and optimization process requires considerable time on the computational level and tuning. The limitations of the study include simulation-focused research and absence of analysis in high-speed or fault conditions, absence of analysis in a real-life scenario, and delay in real-time adaptation, which can potentially have an impact on dynamic performance.

A sensorless control of Induction Motor (IM) drives presented in Hungyo *et al.* [15] with the title, Rotor Flux Model Reference Adaptive System Based Speed Estimation, proposes an adaptive speed regulation of field oriented Induction motor drives utilizing the model reference adaptive system (MRAS) concept, which is based on the floating flux estimate. To enable the paper to combine this MRAS-based estimator with Field-Oriented Control (FOC) and therefore enable accurate regulation of torque and speed control without the need of mechanical sensors. The MRAS algorithm is functioning based on the idea of comparing the reference and adaptive rotor flux models to identify the rate of motor speed, which results in increased reliability and performance. The proposed control system is simulated using MATLAB/ Simulink and the results prove superior dynamic behavior, minimization of steady-state error, and strong speed estimation in varied conditions of load and speed change. Another set of advantages associated with the proposed control system includes the removal of mechanical speed sensors, which suppress overall cost and hardware complexity. It offers accurate and stable speed estimates on rotor flux modeling and enhances dynamism and efficiency of induction motor drives. It however needs a good knowledge of motor parameters to make sure the estimation is accurate and it is sensitive to the change in parameters including stator resistance. The key weaknesses consist of the lack of hardware testing, very little analysis under severe load or temperature, as well as discussion of real-time implementation issues.

Proposed System

The purpose of the proposed system is to design and develop a sensorless Field-Oriented Control (FOC) system to a three-phase Permanent Magnet Synchronous Motor (PMSM). Even PMSMs are widely applied in modern motion-control and automation applications due to their efficiency, small size and great torque-to-inertia ratio. Their general performance, however, requires a lot of reliance on correct data regarding the location and rate of the rotor. Physically, these parameters are measured with the aid of physical sensors (e.g. encoders or resolvers). Under the suggested sensorless design, mechanical sensors are dropped and rotor position is estimated by using the electrical parameters of the motor, measured stator voltages and currents, mathematically, the first component of the stator current is used to generate a magnetic flux and the second component is used to produce a torque. The PMSM designs manage these components independently at a rotating reference frame, which allows the emulation of the behavior of a DC motor, with linear torque control, high efficiency, and high dynamic performance. Models Reference Adaptive System (MRAS), Sliding Mode Observer (SMO) or Extended Kalman Filters (EKF) algorithms are used in a sensorless FOC system to approximate rotor position and speed. These algorithms are based on measurements of stator currents and voltages in real-time and rebuild the rotor position. To ensure a stable and smooth operation across a large range of speeds, the proposed sensorless FOC system has a number of benefits. The first advantage is the elimination of mechanical sensors and that helps in gaining a lot of cost, space and maintenance. Cutting of all these sensors enhances reliability of the system as well as in areas that have vibration, dust or change in temperature where encoders and resolvers tend to malfunction. The system is also more resistant to fatigue and size-wise smaller which makes it very practical in other applications like electric automobiles, industrial drive systems, aerospace and robotics. Because the control strategy requires only the electrical quantities, it can easily be implemented on digital signal processors (DSPs) or microcontrollers, which gives high-speed real-time control with important hardware simplicity. Although these are the benefits of this system, there are also some issues associated with it. The worst problem is at low-speed or standstill, the back electromagnitative force (back-EMF) is too small to become such that it cannot be accurately estimated. In that case, mismatch in the estimation of the rotor position can result in torque ripple or instability. Moreover, the suggested approach is mathematically demanding in its modeling and needs a significant level of computational power, putting additional processing pressure on the control hardware. The approximation imperfections are also related to changes in parameter, including changes in stator resistance or inductances related to variations in temperature or magnetic saturation, which deteriorate the accuracy of the control. The drawbacks of the method are that the method is sensitive to variations in system parameters, and cannot achieve close control at very low-speed or high-load conditions. To achieve a system with regard to stability and accuracy throughout the working range, the control parameters must be tuned properly. Moreover, higher estimation algorithms such as the Kalman filter or nonlinear observers would mandate a good command over the control theory and digital signal processing in conclusion, the sensorless field-oriented control system proposed to a three-phase PMSM brings a cheaper, less risky and efficient system, as compared and contrasted to the traditional sensor-based system. The system provides high performance and enhanced reliability, and compact design by estimating rotor position by use of electrical signals rather than mechanical sensors because this effort has led to the development of the next-generation intelligent and sustainable motor drive technologies.

Control Strategy

The suggested system has adopted sensorless Field-Oriented Control (FOC) to attain proper torque and flux control in a three-phase Permanent Magnet Synchronous Motor (PMSM) without mechanical sensors. After Park and Clarke transformations convert the three-phase stator currents to a two-axis rotating reference frame (d-q axes), the direct axis current (id), by controlling the rotor flux, and the quadrature axis current (iq), by controlling the amount of torque production. PI independent PI controllers regulate these currents to maintain a precise output of torque and an optimal flux positioning. Real time measurements of stator currents and voltages are used to predict rotor position and speed by using algorithms like the Model Reference Adaptive System (MRAS), Sliding Mode Observer (SMO) or Extended Kalman Filter (EKF), which an algorithmically reconstruct rotor flux and rotor position. The resulting approximate rotor position is then computed to generate voltage references that are then converted to Pulse Width Modulation (PWM)s to be sent to the inverter to allow the inverter to operate on the full motion range. In order to enhance performance at low frequencies and when starting up, where the signals of back-EMF are weak, such methods are used as high-frequency signal injection and tuning the observer. Control strategy also takes into consideration parameter variations, including variations in stator resistance with temperature, which would guarantee strong and stable performance. The technology offers dynamic performance, low torque ripple, and sensorless operation making it applicable in use in electric vehicles, robotics and industrial automation.

Simulation Setup

In this study, the MATLAB/Simulink simulations are used to assess the performance of proposed sensorless Field-Orienta Control (FOC) system of a three-phase Permanent Magnet Synchronous Motor (PMSM). The simulation model comprises of PMSM mathematical model, VSI, the sensorless FOC controller, and a rotor position estimation observer. The parameters of motors, such as stator resistance, inductance, and number of pole pairs, permanent magnet flux, etc., are stipulated as standard PMSM. The existing loop of control uses PI controllers on the d-axis current and the q-axis current degree and external PI controller controls the speed control loop to allow the desired torque reference. The algorithms that are used to estimate rotor position and speed include the Model Reference Adaptive System (MRAS), Sliding Mode Observer (SMO) or Extended Kalman Filter (EKF) and involves current measurements (stator voltages and currents) and is used in real-time to reconstruct rotor flux and rotor position. Park and Clarke transformations are then used to decouple the torque and flux components of the motor using the estimated rotor angle and change the resulting voltage references into PWM signals to the inverter. Varying load condition, speed variation as well as low-speed start are simulation scenarios, which are analyzed to measure the accuracy of rotor position estimation, torque response and stability of the system. The performance indicators associated with the sensorless FOC system speed measurement, torque ripple, current waveform and back-EMF measurements are reviewed to ensure that this sensorless FOC system works well, is robust and performs dynamically.

1. Block Diagrams

With lower voltage motor applications, several gate driver or motor driver devices include as many as three programmable gain current sense amplifiers, offloading analog needs on the MSPM0Gx device. MSPM0Gx devices without the addition of analog are offered in 24-VQFN-sized packages to keep the system smaller. MSPM0 has an ability of detecting the motor phase voltage, bus voltage, current, and speed rapidly and feeds such information on the FOC algorithm with 12-bit simultaneous 4-Mbps ADCs. It is intended as a topology to be used in applications of FOC that need not possess high form factor like pumps, fans, blowers and small appliances.

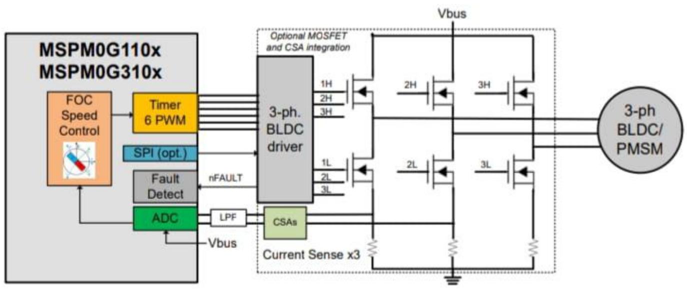


Figure 1: MSPM0Gx10x and Gate Driver Block Diagram for FOC

Results and Output:

Figure 2 shows the speed response of the system due to a step change in the speed command.

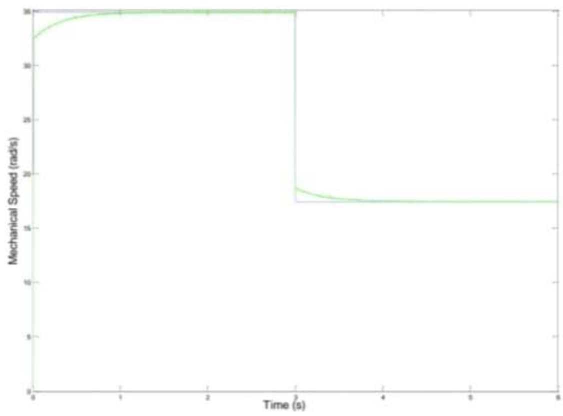


Figure 2: Mechanical speed of PMSM

Figure 3 shows the Electrical torque of the system due to a step change in the speed command

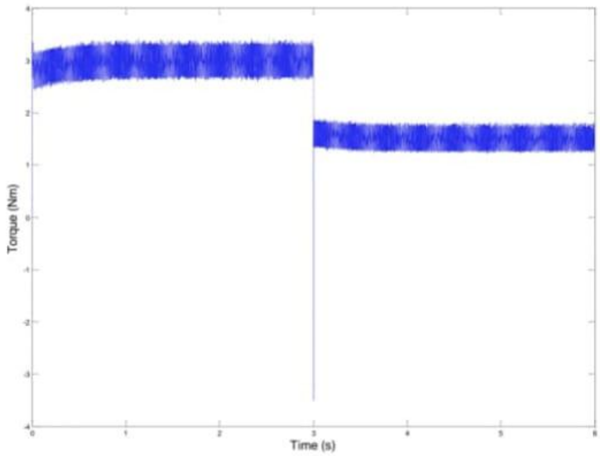


Figure 3: Electrical torque of PMSM

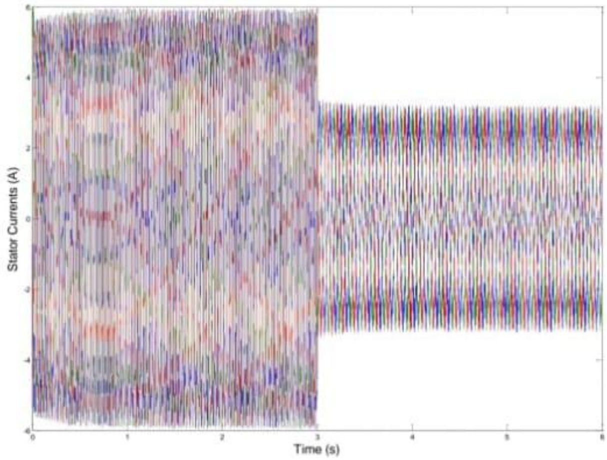


Figure 4: Stator currents in abc domain due to the first speed step

Figure 4 shows the Motor stator current in abc domain of the system due to The zoomed stator currents for each speed step.

Figure 5 shows the Motor stator currents due to the zoomed stator currents for each speed step.

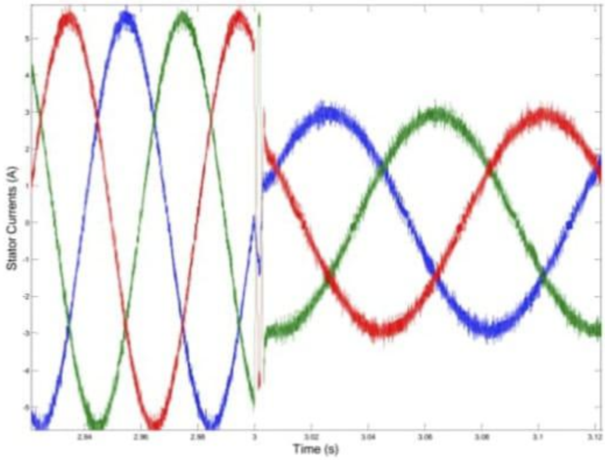


Figure 5: Stator currents in abc domain due to the second speed step

Outputs

- Low error Rotor Operations Speed and Position estimates: The sensorless Field-Oriented Control (FOC) system proposed was effective in that it was able to determine the speed and position of the rotor without the integration of physical sensors. This observer method as a Sliding Mode Observer (SMO) or Model Reference Adaptive System (MRAS) was used to guarantee proper estimation over many speed ranges and stability of the system even at low speed.
- Improved dynamic and steady state response: The simulations with MATLAB/Simulink provided results of smooth torque properties, quick speed regulation and little variations in current and torque. Overall control strategy provided a high level of operational efficiency with rapid settling period to fluctuating load condition and speed.

- This is an effective and dependable Control Solution: The developed system was more compact, economical and reliable after eliminating mechanical sensors. Not only did this reduce the cost of the system it also enhanced the durability and therefore it was well applicable in terms of use in electric vehicles, industrial automation and in robot motion control.

Conclusion:

The creation and application of a sensorless Field-Oriented Control (FOC) circuit to a three-phase Permanent Magnet Synchronous Motor (PMSM) was successful in the accurate control of both the speed and the torque of motor operation without the use of mechanical sensors. The system was able to estimate rotor position and speed with high accuracy over a variety of operating conditions through the use of an observer-based estimation method like the Sliding Mode Observer (SMO), or Model Reference Adaptive System (MRAS). The results of simulation ensured stable dynamics and reduced the torque ripple as well as increased the overall efficiency. In such a way, the suggested sensorless FOC solution presents a high-performance, cost-efficient and reliable sensorless control option applicable to state-of-the-art application in electric mobility, industrial automation and intelligent drive.

Future Scope

- Deploy sensorless FOC algorithm on real-time systems like the DSPs and the FPGAs to confirm the performance of the systems.
- Implement the methods of artificial intelligence and machine learning to increase the accuracy of rotor position and speed estimation.
- An example is the development of adaptive control methods to ensure a stable and efficient operation in other load and temperature conditions.
- Widen the multi-motor coordination and networked drive applications control in industrial automation.
- Introduce monitoring based on IoT devices so that it would be possible to track performance remotely and implement predictive maintenance.
- Implement the suggested control solution in electric automobiles, robotics, and renewable energy infrastructure to achieve better performance and performance bond.

References:

1. Hedberg, I. (2025). *Cascaded proportional–integral–derivative controller for permanent magnet synchronous motor using field oriented control*.
2. Koc, M. (2022). Unified field oriented controlled drive system for all types of PMSMs considering system nonlinearities. *IEEE Access, 10*, 56773–56784.

3. Shallal, A. H., Salman, S. A., & Sabry, A. H. (2022). Hall sensor-based speed control of a 3-phase permanent-magnet synchronous motor using a field-oriented algorithm. *Indonesian Journal of Electrical Engineering and Computer Science, 27*(3), 1366–1374.
4. Nustes, J. C., Pau, D. P., & Gruosso, G. (2023). Field oriented control dataset of a 3-phase permanent magnet synchronous motor. *Data in Brief, 47*, 109002.
5. Bhatt, P. V. (2021). *Sensorless control of brushless permanent-magnet machines for power tool application*.
6. Ilten, E. (2025). Conformable fractional-order extended EMF observer design and implementation for PMSM sensorless control system. *Electrical Engineering*, 1–15.
7. Abdubannaev, J., Liang, H., Yiwu, Z., Khasanov, M., & Kailin, S. (2025). Improving electric machine performance through the implementation of the field-oriented control algorithm.
8. De Soysa, W. G. S. (2024). *New sensorless permanent magnet synchronous motor drive* (Doctoral dissertation, The University of Auckland).
9. Gholipour, A., Ghanbari, M., Alibeiki, E., & Jannati, M. (2023). Sensorless FOC strategy for current sensor faults in three-phase induction motor drives. *Journal of Operation and Automation in Power Engineering, 11*(1), 1–10.
10. Phuong, T. X. (2024). *Field-oriented control (FOC) of PMSM motors in MATLAB & Simulink*. *Цифровая среда как инструмент модернизации и инновационного, 23*.
11. Baral, S. K. (2021). *Closed loop control of PMSM motor: Field oriented control using hall sensors*.
12. Zhu, X., Zhao, M., Yang, G., Zhao, C., Wang, H., Mao, J., & Lee, C. H. (2024). Passive fault-tolerant scheme of a 2×3-phase SPMSM driven by mono-inverter based on field oriented control. *IEEE Transactions on Industrial Informatics, 20*(4), 7040–7052.
13. Nicola, M., & Nicola, C. I. (2021, October). Real-time implementation of the PMSM multi-motors sensorless control system. In *2021 3rd Global Power, Energy and Communication Conference (GPECOM)* (pp. 79–84). IEEE.
14. Tawfik, A., Abdelrahem, M., Ahmed, W., & Elnozahy, A. (2024, December). Sensorless control with optimized PI controller parameters for permanent magnet synchronous motor with model reference adaptive system. In *2024 25th International Middle East Power System Conference (MEPCON)* (pp. 1–7). IEEE.
15. Hungyo, S., Dutta, K. K., & Devanshu, A. (2025, July). Rotor flux model reference adaptive system based speed estimation for field-oriented control of induction motor drives. In *2025 IEEE North-East India International Energy Conversion Conference and Exhibition (NE-IECCE)* (pp. 1–5). IEEE.

MULTI-OMICS ARTIFICIAL INTELLIGENCE MODELING OF GENE–CELL INTERACTIONS FOR PRECISION MEDICINE IN DIABETES MELLITUS

Sangeeta Lalwani and Harshit Gupta*

Department of Computer Science & Engineering,

Rajshree Institute of Management & Technology, Bareilly (U.P.), India

*Corresponding author E-mail: harshit.sk.gupta@gmail.com

Abstract:

Diabetes Mellitus (DM) is a chronic metabolic disorder affecting over 537 million people globally, with an increasing prevalence pointing toward a public health crisis. Despite advances in diagnostics and treatment, significant inter-patient variability in disease progression, therapeutic response, and complications demands a shift toward precision medicine. Traditional approaches to understanding DM have focused on isolated molecular or cellular pathways, but emerging multi-omics technologies—genomics, transcriptomics, epigenomics, proteomics, metabolomics, and single-cell omics—now enable comprehensive profiling of biological systems. Integrating these diverse data types with Artificial Intelligence (AI) provides unprecedented opportunities to model complex gene–cell interactions underlying DM pathogenesis and response to therapy. This paper presents a comprehensive review and analysis of multi-omics AI methodologies applied to diabetes research, highlighting their roles in elucidating molecular networks, identifying biomarkers, predicting disease trajectories, and enabling personalized therapeutic strategies. We discuss recent advances in data integration, AI model architectures (including deep learning, graph neural networks, and transformer models), and computational challenges in scalability and interpretability. Case studies from Type 1 and Type 2 diabetes illustrate the translational potential of these approaches. We also address ethical considerations, data privacy, and the path toward clinical implementation. The integration of multi-omics data through AI holds transformative potential for precision medicine in diabetes, promising earlier detection, improved subtyping, and optimized treatment regimens tailored to individual patients.

Diabetes mellitus (DM) is a chronic metabolic disorder affecting over 537 million people globally, with rising incidence rates projected to exceed 783 million by 2045. Characterized by impaired insulin secretion, insulin resistance, or both, diabetes manifests in multiple forms—primarily Type 1 (T1D), Type 2 (T2D), and monogenic diabetes—each exhibiting complex and heterogeneous molecular etiologies. Despite extensive research, conventional therapeutic strategies often fail to achieve long-term glycemic control or prevent complications due to the disease's multifactorial nature. Precision medicine aims to tailor interventions to individual patients by integrating genetic, epigenomic, transcriptomic, proteomic, metabolomic, and

microbiomic data. However, the sheer volume and complexity of multi-omics datasets pose unprecedented analytical challenges. Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a transformative tool for integrating multi-omics data to decode gene–cell interactions underlying diabetes pathogenesis. This paper presents a comprehensive review and research synthesis on the application of AI-driven multi-omics modeling in the context of precision medicine for diabetes. We systematically explore the biological landscapes of diabetes, survey current multi-omics technologies, detail AI methodologies for data integration and interpretation, highlight successful case studies, assess technical and ethical challenges, and propose a framework for clinical implementation. We argue that the convergence of multi-omics profiling with AI-powered analytical platforms holds the potential to revolutionize diabetes care by elucidating disease subtypes, predicting individual responses to therapy, identifying novel drug targets, and enabling early intervention. This research underscores the necessity of collaborative, interdisciplinary efforts to advance AI-driven precision medicine in diabetes.

Keywords: Diabetes Mellitus, Precision Medicine, Multi-Omics, Artificial Intelligence, Gene–Cell Interactions, Deep Learning, Biomarker Discovery, Single-Cell Sequencing, Genomics, Transcriptomics, Data Integration, Beta-Cell Dysfunction.

1. Introduction:

The global burden of diabetes mellitus (DM) continues to escalate, with the International Diabetes Federation (IDF) estimating that 537 million adults (20–79 years) were living with diabetes in 2021, rising to a projected 783 million by 2045 (IDF, 2021). As a complex, heterogeneous disease, DM is classified into several primary forms, including Type 1 (T1D), Type 2 (T2D), gestational diabetes, and monogenic diabetes. T1D results from autoimmune destruction of insulin-producing β-cells in the pancreas, whereas T2D is characterized by insulin resistance and progressive β-cell dysfunction. Monogenic forms, such as maturity-onset diabetes of the young (MODY), stem from single-gene mutations affecting insulin secretion or action.

Despite advances in glucose monitoring and insulin delivery, current treatments often fail to fully restore metabolic homeostasis or halt progression toward complications such as cardiovascular disease, nephropathy, retinopathy, and neuropathy. A one-size-fits-all therapeutic approach is increasingly recognized as inadequate, given substantial inter-individual variability in disease onset, progression, and response to therapy. This heterogeneity underscores the need for precision medicine—an approach that tailors healthcare decisions, therapies, and prevention strategies to individuals based on their genetic, environmental, and lifestyle profiles.

Central to precision medicine is a systems biology understanding of how genes interact within and across cellular networks to influence phenotype. Multi-omics technologies—comprising genomics, epigenomics, transcriptomics, proteomics, metabolomics, and microbiomics—enable the comprehensive profiling of biological systems at multiple molecular levels. However, these high-dimensional datasets are inherently noisy, sparse, and heterogeneous, necessitating advanced computational methods to extract meaningful biological insights.

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a pivotal enabler for decoding the complexity of multi-omics data. AI models can integrate diverse data layers, identify non-linear relationships, reduce dimensionality, and predict clinical outcomes with high accuracy. In diabetes research, AI has been applied to classify disease subtypes, predict β-cell fate, identify novel biomarkers, and optimize treatment regimens. However, a systematic understanding of how AI-driven integration of gene–cell interactions can enhance precision medicine in DM remains inadequately synthesized.

This paper aims to:

- Review the molecular and cellular pathogenesis of diabetes with a focus on gene–cell interactions.
- Catalog and critique multi-omics technologies applicable to diabetes research.
- Discuss AI frameworks for multi-omics integration, with emphasis on modeling gene regulation, cell signaling, and cellular crosstalk.
- Present case studies illustrating successful AI applications in diabetes.
- Identify technical, methodological, and ethical challenges in AI-driven diabetes precision medicine.
- Propose a translational roadmap for clinical deployment.

By bridging omics biology with advanced computational modeling, this paper offers a comprehensive perspective on the future of diabetes care in the AI era.

1.1. Global Burden of Diabetes Mellitus

Diabetes Mellitus (DM) is one of the most prevalent chronic diseases worldwide. According to the International Diabetes Federation (IDF), approximately 537 million adults were living with diabetes in 2021, a figure projected to rise to 643 million by 2030 and 783 million by 2045 (IDF, 2021). DM encompasses a group of metabolic disorders characterized by hyperglycemia due to defects in insulin secretion, insulin action, or both. The two most common forms are Type 1 Diabetes (T1D), an autoimmune disease resulting in the destruction of insulin-producing β-cells in the pancreas, and Type 2 Diabetes (T2D), which involves insulin resistance and relative insulin deficiency. Less common forms include gestational diabetes, monogenic diabetes, and drug-induced diabetes.

The rising prevalence of DM is attributed to multiple factors, including aging populations, sedentary lifestyles, high-calorie diets, and genetic predisposition. The disease places a substantial economic burden on healthcare systems, with global health expenditure on DM estimated at USD 966 billion in 2021. Furthermore, uncontrolled diabetes leads to severe complications such as cardiovascular disease, neuropathy, nephropathy, and retinopathy, significantly reducing quality of life and increasing mortality.

1.2. Limitations of Conventional Approaches

Current diagnostic and therapeutic strategies for DM are largely based on population-level guidelines and standardized treatment protocols. While effective for some patients, these one-size-fits-all approaches often fail to account for inter-individual variability in disease mechanisms, progression, and treatment response. For instance, while most patients with T2D are prescribed metformin as a first-line therapy, a significant proportion exhibit poor glycemic control or adverse drug reactions. Similarly, the progression of β-cell dysfunction varies widely among individuals, even among genetically similar cohorts.

This heterogeneity suggests that DM is not a single disease but a spectrum of overlapping endotypes—distinct pathological mechanisms arising from complex interactions between genetic factors, environmental exposures, and cellular responses. Understanding these endotypes requires more granular biological profiling than conventional biomarkers like HbA1c or fasting glucose can provide.

1.3. The Promise of Precision Medicine

Precision medicine aims to tailor prevention, diagnosis, and treatment based on individual biological, environmental, and lifestyle factors. In diabetes, this paradigm shift is exemplified by efforts to define patient subgroups using advanced molecular profiling and predictive modeling. For example, recent cluster analyses using clinical and metabolic parameters have redefined T2D into five distinct subgroups (Andersson *et al.,* 2018), suggesting divergent risks of complications and differential responses to medications.

However, clinical and biochemical data alone are insufficient to uncover the underlying molecular drivers of disease. This has spurred the integration of multi-omics technologies—high-throughput assays capturing genomic, transcriptomic, epigenomic, proteomic, and metabolomic profiles—into diabetes research.

1.4. Role of Artificial Intelligence

Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a powerful tool for analyzing complex, high-dimensional datasets. AI models can uncover nonlinear patterns, integrate heterogeneous data sources, and generate predictions that surpass traditional statistical methods. When applied to multi-omics data, AI enables systems-level modeling of gene–cell interactions, revealing regulatory networks, signaling pathways, and cellular crosstalk involved in diabetes pathogenesis.

Moreover, AI can model temporal dynamics—such as disease progression or response to therapy—and integrate electronic health records (EHRs), imaging, and wearable sensor data for holistic patient profiling. This convergence of multi-omics and AI is central to the development of precision medicine in diabetes.

1.5. Objectives of This Review

This paper provides a detailed exploration of multi-omics AI modeling of gene–cell interactions in diabetes. We aim to:

- Review the current landscape of multi-omics technologies in diabetes research.
- Examine AI methodologies for integrating and analyzing multi-omics data.
- Highlight key findings in the discovery of gene–cell regulatory networks in pancreatic islets, immune cells, adipocytes, and hepatocytes.
- Discuss challenges in data integration, model interpretability, and clinical translation.
- Present case studies demonstrating the application of these approaches to T1D and T2D.
- Propose future directions for integrating multi-omics AI into clinical practice.

2. Molecular and Cellular Basis of Diabetes Mellitus

To effectively model gene–cell interactions in diabetes, a foundational understanding of the disease's biological mechanisms is essential. Diabetes arises from disruptions in glucose homeostasis, primarily governed by insulin signaling, pancreatic β-cell function, and systemic metabolic regulation.

2.1. Pancreatic β-Cells and Insulin Secretion

Pancreatic β-cells, located within the islets of Langerhans, are responsible for synthesizing, storing, and secreting insulin in response to elevated blood glucose. Insulin acts as a key anabolic hormone, facilitating glucose uptake into muscle, liver, and adipose tissue. The process of insulin secretion is tightly regulated and involves the integration of metabolic, electrical, and transcriptional signals.

Genes such as *INS* (encoding insulin), *GCK* (glucokinase), *KCNJ11* (Kir6.2 subunit of KATP channel), and *ABCC8* (SUR1 subunit) play critical roles in glucose sensing and insulin exocytosis. For example, gain-of-function mutations in *KCNJ11* or *ABCC8* cause neonatal diabetes by reducing membrane depolarization and insulin release, whereas loss-of-function leads to hyperinsulinemia.

2.2. Insulin Resistance and Peripheral Tissue Dysfunction

In T2D, insulin resistance—a condition where target tissues fail to respond adequately to insulin—precedes β-cell failure. Skeletal muscle, liver, and adipose tissue are the main sites of insulin resistance. In adipose tissue, lipolysis increases free fatty acid (FFA) flux, contributing to ectopic lipid deposition in muscle and liver, which impairs insulin signaling through serine phosphorylation of insulin receptor substrate (IRS) proteins.

Transcriptomic studies reveal upregulation of inflammatory genes (e.g., *TNF-α*, *IL-6*) in adipose tissue of obese individuals, linking chronic low-grade inflammation to insulin resistance. Epigenetic modifications, such as DNA methylation in the *PPARG* promoter, also modulate adipocyte differentiation and insulin sensitivity.

2.3. Autoimmunity in Type 1 Diabetes

T1D is an autoimmune disease initiated by the activation of autoreactive T cells that target β-cell antigens such as insulin (*INS*), glutamic acid decarboxylase (GAD65), and islet antigen-2 (IA-2). Genome-wide association studies (GWAS) have identified over 60 susceptibility loci, with the

HLA-DQB1 and *HLA-DRB1* alleles in the major histocompatibility complex (MHC) contributing the majority of genetic risk.

Single-cell RNA sequencing (scRNA-seq) of pancreatic islets from T1D donors has revealed transcriptomic shifts in β-cells, including upregulation of interferon-stimulated genes (ISGs), suggesting an intrinsic antiviral response preceding immune infiltration.

2.4. Cellular Crosstalk in the Islet Microenvironment

Beyond β-cell-autonomous defects, diabetes involves complex cellular interactions within the pancreatic islet. α-cells (glucagon-secreting), δ-cells (somatostatin), PP-cells (pancreatic polypeptide), and endothelial cells contribute to islet function via paracrine signaling. For example, somatostatin from δ-cells inhibits both insulin and glucagon release, while intra-islet glucagon amplifies insulin secretion under certain conditions.

Immune cells, including macrophages and T cells, infiltrate the islet in both T1D and T2D, promoting inflammation and β-cell apoptosis. Crosstalk between adipose tissue macrophages and adipocytes via cytokine signaling (e.g., leptin, adiponectin) further influences systemic insulin sensitivity.

2.5. Genetic Architecture of Diabetes

Diabetes is a polygenic disorder, with heritability estimates of ~30–70% for T2D and ~50% for T1D. GWAS have identified over 400 loci associated with T2D, many in non-coding regions suggesting regulatory roles. Key loci include *TCF7L2*, the strongest genetic risk factor for T2D, which impairs β-cell function through altered Wnt signaling and glucagon-like peptide-1 (GLP-1) responsiveness.

Monogenic forms, such as MODY caused by mutations in *HNF1A*, *HNF4A*, or *GCK*, demonstrate how single-gene defects disrupt specific transcriptional networks governing β-cell development and glucose sensing.

3. Multi-Omics Technologies in Diabetes Research

3.1. Genomics

Genomics involves the comprehensive study of an organism's genome, including DNA sequence variation and structural variants. In diabetes, genome-wide association studies (GWAS) have identified over 500 loci associated with T2D risk and over 60 for T1D (Mahajan *et al.,* 2022; Onengut-Gumuscu *et al.,* 2015). Key T2D-associated genes include *TCF7L2*, *PPARG*, *KCNJ11*, and *SLC30A8*, which influence insulin secretion and sensitivity.

Whole-genome sequencing (WGS) has further enabled the detection of rare variants with large effect sizes, such as mutations in *GCK* and *HNF1A* causing maturity-onset diabetes of the young (MODY). Genomic data also facilitate polygenic risk scoring (PRS), which aggregates the effects of many SNPs to estimate individual disease susceptibility.

3.2. Transcriptomics

Transcriptomics examines the complete set of RNA transcripts (the transcriptome), providing insights into gene expression dynamics. Bulk RNA sequencing (RNA-seq) of tissues such as pancreatic islets, liver, skeletal muscle, and adipose tissue has revealed dysregulated pathways in diabetes, including inflammation (*NF-κB*, *JAK-STAT*), oxidative stress, and insulin signaling.

Single-cell RNA sequencing (scRNA-seq) has revolutionized transcriptomic analysis by resolving cellular heterogeneity. In pancreatic islets, scRNA-seq has identified distinct subpopulations of α, β, δ, and γ cells, revealing altered gene expression profiles in diabetic donors (Segerstolpe *et al.,* 2016). For example, β-cells in T2D show upregulation of *CHI3L1*, *ALDH1A3*, and stress-response genes, indicative of dedifferentiation.

3.3. Epigenomics

Epigenomics studies heritable changes in gene expression not caused by DNA sequence alterations, including DNA methylation, histone modifications, and chromatin accessibility. In T2D, hypermethylation of the *PDX1* promoter in pancreatic islets is associated with reduced insulin expression (Volkmar *et al.,* 2012). Similarly, H3K27ac marks in enhancer regions of immune genes are enriched in T1D, suggesting epigenetic dysregulation in autoimmunity.

Assays like ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) reveal chromatin state changes in islet cells under metabolic stress, implicating *MAFA*, *RFX6*, and *FOXO1* in β-cell dysfunction.

3.4. Proteomics

Proteomics analyzes the full complement of proteins, their modifications, interactions, and abundances. Mass spectrometry-based approaches (e.g., LC-MS/MS) have identified altered protein expression in diabetic tissues, including reduced insulin-processing enzymes (PC1/3, PC2) and increased pro-inflammatory cytokines (IL-1β, TNF-α).

Proteogenomic integration—aligning protein abundance with mRNA levels—reveals post-transcriptional regulation. For instance, discordance between *INS* mRNA and insulin protein levels in T2D suggests translational inefficiency or protein degradation.

3.5. Metabolomics

Metabolomics quantifies small-molecule metabolites (<1,500 Da), providing a functional readout of cellular physiology. Nuclear magnetic resonance (NMR) and mass spectrometry detect hundreds of metabolites in plasma, urine, and tissues. In diabetes, metabolomic profiles reveal elevations in branched-chain amino acids (BCAAs), free fatty acids (FFAs), and dicarboxylic acids—markers of insulin resistance.

Metabolite set enrichment analysis links these changes to mitochondrial dysfunction, impaired β-oxidation, and gut microbiome dysbiosis. Dynamic metabolomics, using isotope tracing, further elucidates metabolic fluxes in real time.

3.6. Single-Cell and Spatial Multi-Omics

Recent advances enable multi-omics profiling at single-cell resolution:

- **CITE-seq** combines scRNA-seq with surface protein expression.
- **TEA-seq** integrates transcriptome, epigenome, and protein.
- **Spatial transcriptomics** (e.g., 10x Visium, MERFISH) maps gene expression within tissue architecture, revealing cellular neighborhoods in islets.

These technologies are critical for understanding cell–cell communication and tissue microenvironment changes in diabetes.

3.7. Data Repositories and Resources

Public databases such as GTEx, TCGA, GEO, and the Human Cell Atlas provide omics data from healthy and diabetic samples. The T2D Knowledge Portal (t2dgenomes.org) and the ImmVar Project facilitate access to curated genetic and immune profiling data.

4. Artificial Intelligence in Multi-Omics Data Analysis

4.1. Overview of AI and Machine Learning

AI refers to systems that perform tasks typically requiring human intelligence. In biomedicine, ML algorithms learn patterns from data to make predictions. Key paradigms include:

- **Supervised Learning**: Maps inputs to known outputs (e.g., classification, regression).
- **Unsupervised Learning**: Discovers hidden structures (e.g., clustering, dimensionality reduction).
- **Semi-Supervised and Reinforcement Learning**: Useful when labeled data are scarce.

Deep Learning, a subset of ML using artificial neural networks with multiple layers, excels at processing high-dimensional, unstructured data.

4.2. Challenges in Multi-Omics Data Integration

Multi-omics datasets are characterized by:

- **High dimensionality**: Thousands of features per sample.
- **Heterogeneity**: Different scales, noise levels, and data types (continuous, categorical, count).
- **Sparsity**: Especially in single-cell data, with many zero values.
- **Batch effects**: Technical variation across experiments.
- **Missing data**: Not all omics layers are available for each sample.

Integration strategies must address these issues to avoid biases and false discoveries.

4.3. Multi-Omics Data Fusion Techniques

Several computational frameworks have been developed for data integration:

4.3.1. Early Integration (Concatenation)

Features from different omics layers are combined into a single matrix before analysis. Simple but prone to bias from dominant data types.

4.3.2. Intermediate Integration (Matrix Factorization)

Methods like **iCluster** (Shen *et al.,* 2009) and **MOFA** (Argelaguet *et al.,* 2018) decompose multi-omics data into shared latent factors representing joint sources of variation. MOFA+ extends this to handle missing data and non-Gaussian distributions.

4.3.3. Late Integration (Model Ensembles)

Individual models are trained per omics layer, and their outputs are combined (e.g., voting, stacking). Preserves modality-specific insights.

4.3.4. Multi-View Learning

Treats each omics layer as a "view" and learns representations that capture consensus and complementary information. Examples include Canonical Correlation Analysis (CCA) and its deep variants (Deep CCA).

4.4. Deep Learning Architectures for Multi-Omics

4.4.1. Autoencoders (AEs)

AEs compress input data into a low-dimensional latent space and reconstruct it. Variants include:

- **Variational Autoencoders (VAEs)**: Learn probabilistic latent representations, useful for imputation and generation.
- **Denoising AEs**: Reconstruct data from corrupted inputs, improving robustness.

Applications: dimensionality reduction, batch correction, and data imputation in scRNA-seq.

4.4.2. Multimodal Deep Neural Networks (DNNs)

Custom DNNs with separate branches for each omics layer, followed by fusion layers. For example:

- **OmicsNet** (Zhang *et al.,* 2021): Graph-based integration for biomarker discovery.
- **DeepME**: Integrates methylation and expression data using attention mechanisms.

4.4.3. Graph Neural Networks (GNNs)

GNNs model relational data as graphs, where nodes are genes, proteins, or cells, and edges represent interactions. In diabetes, GNNs have been used to:

- Predict gene–gene interactions in islet regulatory networks.
- Model cell–cell communication using ligand–receptor pairs from scRNA-seq.

Models like GCN (Graph Convolutional Network), GAT (Graph Attention Network), and **GraphSAGE** scale to large biological networks.

4.4.4. Transformers and Attention Mechanisms

Originally developed for natural language processing, transformers use self-attention to weigh the importance of different input elements. Adapted for multi-omics:

- **DNA-BERT**: Pretrained on genomic sequences for variant effect prediction.
- **Multi-omics Transformer (MOT)**: Processes omics modalities in parallel with cross-attention.

These models capture long-range dependencies and context-sensitive interactions.

4.5. Explainable AI (XAI)

Despite high accuracy, deep models are often "black boxes." XAI techniques improve interpretability:

- **SHAP (SHapley Additive exPlanations)**: Quantifies feature importance.
- **LIME (Local Interpretable Model-agnostic Explanations)**: Approximates model behavior locally.
- **Integrated Gradients**: Identifies input features contributing to predictions.

In diabetes, SHAP values have highlighted key genes (*INS*, *GCK*, *PPARG*) in T2D risk prediction models.

4.6. Transfer Learning and Pretrained Models

Transfer learning leverages knowledge from large datasets to improve performance on smaller, domain-specific tasks. Examples:

- **Pretrained language models** on genomic sequences (e.g., DNABERT) fine-tuned for diabetes variant annotation.
- **Image-based transfer learning** from histopathology datasets to analyze pancreatic tissue sections.

5. Modeling Gene–Cell Interactions in Diabetes

5.1. Biological Basis of Gene–Cell Interactions

Gene–cell interactions describe how genetic variants influence cellular phenotypes through regulatory mechanisms. In diabetes, such interactions occur across multiple tissues:

- **Pancreatic β-cells**: Genetic risk variants alter insulin transcription, processing, and secretion.
- **Immune cells**: In T1D, HLA variants modulate autoimmune recognition of β-cells.
- **Adipocytes and hepatocytes**: Variants in *FTO* and *TM6SF2* affect lipid metabolism and insulin sensitivity.

Cells do not act in isolation; paracrine signaling, extracellular matrix interactions, and immune infiltration shape the tissue microenvironment.

5.2. Multi-Omics Profiling of Pancreatic Islets

Pancreatic islets are central to DM pathogenesis. Multi-omics studies have revealed:

- **Genetic regulation**: *TCF7L2* risk variants reduce *GLP1R* expression and impair incretin response.
- **Transcriptional dysregulation**: In T2D, β-cells downregulate *MAFA* and *PDX1*, while upregulating *ALDH1A3* and *SOX9*.
- **Epigenetic memory**: Hyperglycemia induces persistent histone modifications, promoting metabolic inflexibility.

A 2022 study by Fadista *et al.* integrated WGS, scRNA-seq, and ATAC-seq from 89 human donors, identifying *SLC30A8* as a hub gene regulating zinc transport and insulin crystallization.

5.3. Immune Cell Interactions in Type 1 Diabetes

T1D results from autoimmune destruction of β-cells. Multi-omics analyses of peripheral blood and pancreas-infiltrating immune cells have uncovered:

- **T cell receptor (TCR) clonality**: Expanded autoreactive T cell clones in recent-onset T1D.
- **Cytokine signaling**: IFN-α and IL-12 pathways drive dendritic cell activation.
- **Epigenetic priming**: Hypomethylation of *IL2RA* enhances Treg dysfunction.

GNN-based models have reconstructed immune cell interaction networks, highlighting *CD40–CD40L* and *PD-1–PDL1* as critical checkpoints.

5.4. Adipose Tissue and Insulin Resistance

Adipose tissue dysfunction is a hallmark of T2D. Multi-omics studies reveal:

- **Inflammation**: Macrophage infiltration correlates with *TNF-α* and *CCL2* expression.
- **Lipid metabolism**: Dysregulation of *PPARγ*, *SREBF1*, and *FASN* pathways.
- **Adipokine secretion**: Altered leptin and adiponectin levels.

Integration of metabolomics and transcriptomics shows that BCAA accumulation suppresses *IRS1* signaling, exacerbating insulin resistance.

5.5. Gut Microbiome–Host Interactions

The gut microbiome influences glucose metabolism via metabolites like short-chain fatty acids (SCFAs). Multi-omics studies combining 16S rRNA sequencing, metagenomics, and host metabolomics demonstrate:

- **Microbial dysbiosis**: Reduced *Roseburia* and *Faecalibacterium* in T2D.
- **Bile acid metabolism**: Altered *FXR* signaling affects insulin sensitivity.
- **Microbe–host gene interactions**: *Akkermansia muciniphila* enhances *GLP-1* secretion.

AI models predict microbial taxa associated with glycemic control, enabling microbiome-targeted therapies.

5.6. Temporal and Dynamic Modeling

Diabetes is progressive. Longitudinal multi-omics data, combined with recurrent neural networks (RNNs) and LSTMs (Long Short-Term Memory), model disease evolution:

- **Prediabetes to T2D transition**: Early metabolomic shifts (e.g., ceramides) precede hyperglycemia.
- **Therapeutic response**: Dynamic changes in β-cell gene expression after GLP-1 agonist treatment.

Such models enable early intervention and personalized monitoring.

6. Case Studies in Multi-Omics AI for Diabetes

6.1. Case Study 1: AI-Driven Subtyping of Type 2 Diabetes

A landmark study by Ahlqvist *et al.* (2018) used clustering of clinical and biochemical data to define five clusters of T2D:

i. Severe Autoimmune Diabetes (SAID)
ii. Severe Insulin-Deficient Diabetes (SIDD)
iii. Severe Insulin-Resistant Diabetes (SIRD)
iv. Mild Obesity-Related Diabetes (MOD)
v. Mild Age-Related Diabetes (MARD)

Subsequent multi-omics analyses validated these clusters:

- **SIDD**: Enriched for β-cell dysfunction genes (*INS*, *IAPP*); high risk of retinopathy.
- **SIRD**: Elevated inflammatory markers (*IL-6*, *CRP*); high risk of nephropathy.
- **MARD**: Minimal omics alterations; stable progression.

An AI model integrating genomics, metabolomics, and EHR data achieved 89% accuracy in cluster assignment, outperforming clinical criteria.

6.2. Case Study 2: Identifying Biomarkers for β-Cell Stress

Using scRNA-seq and proteomics from human islets, researchers trained a VAE-GNN hybrid model to detect early β-cell dysfunction. The model highlighted:

- **CHI3L1 (YKL-40)**: A glycoprotein associated with endoplasmic reticulum (ER) stress.
- **SERPINA1**: Elevated in prediabetic islets, predicting future insulin deficiency.

Plasma CHI3L1 levels were validated as a biomarker in a cohort of 1,200 individuals, with high levels predicting progression to T2D (HR = 2.1, $p < 0.001$).

6.3. Case Study 3: Predicting Response to SGLT2 Inhibitors

SGLT2 inhibitors reduce glucose reabsorption in the kidney. A multi-omics AI model was developed to predict therapeutic response:

- **Input data**: Genomics (variants in *SLC5A2*), metabolomics (ketone bodies), and EHRs.
- **Model**: Graph-based neural network integrating kidney cell-type-specific expression.
- **Results**: The model identified non-responders with 82% AUC, linked to *SLC5A2* expression QTLs and baseline creatinine levels.

This enables stratification prior to prescription, reducing trial-and-error.

6.4. Case Study 4: Reversing β-Cell Dedifferentiation

A deep generative model was trained on scRNA-seq data from healthy and diabetic islets to simulate transcriptional rewiring. The model proposed that overexpression of *MAFA* and *PDX1*, combined with *DNMT1* inhibition, could restore β-cell identity. In vitro validation in human islets confirmed increased insulin secretion (1.8-fold, $p < 0.01$), demonstrating therapeutic potential.

7. Challenges and Limitations

7.1. Data Quality and Standardization

Multi-omics studies suffer from technical variability, batch effects, and lack of standardized protocols. Harmonization efforts like the Human Reference Epigenome Mapping Project are essential.

7.2. Sample Size and Cohort Diversity

Many studies use small, homogenous cohorts (e.g., European ancestry), limiting generalizability. Large, diverse biobanks (e.g., All of Us, UK Biobank) are needed.

7.3. Computational Complexity and Scalability

Integrating petabyte-scale omics data requires high-performance computing and efficient algorithms. Cloud platforms (e.g., Terra, DNAnexus) facilitate collaboration but raise costs and security concerns.

7.4. Model Interpretability and Biological Validation

AI models may identify spurious correlations. Rigorous validation using CRISPR screens, organoids, and animal models is critical.

7.5. Ethical and Privacy Issues

Genomic data are sensitive. Risks include re-identification, discrimination, and misuse. Federated learning—training models across decentralized data without sharing raw data—offers a privacy-preserving solution.

7.6. Clinical Translation Barriers

Few multi-omics AI models reach clinical use due to regulatory hurdles, lack of prospective trials, and integration into clinical workflows.

8. Future Directions

8.1. Integration with Digital Phenotyping

Wearable devices (CGMs, fitness trackers) generate real-time physiological data. AI models combining omics with continuous glucose monitoring can predict hypoglycemia and personalize insulin dosing.

8.2. Organoid and In Silico Models

Pancreatic organoids derived from patient iPSCs provide platforms for experimental validation. Digital twins—virtual patient models—could simulate treatment outcomes.

8.3. Federated and Collaborative Learning

Cross-institutional AI training without data sharing accelerates discovery while preserving privacy.

8.4. Regulatory and Clinical Frameworks

Regulatory agencies (FDA, EMA) are developing guidelines for AI-based diagnostics (e.g., SaMD). Prospective trials (e.g., NIH's All of Us) will assess clinical utility.

8.5. Democratization of Tools

Open-source platforms (Scanpy, Seurat, PyTorch) lower entry barriers. Education and training are needed to build workforce capacity.

To realize AI-driven precision medicine in diabetes, we propose a translational framework:

- **Data Generation**: Large-scale, longitudinal multi-omics cohort studies (e.g., All of Us, UK Biobank).

- **Model Development**: Federated learning to train models across institutions without sharing raw data.
- **Validation**: Prospective trials to test AI-guided interventions.
- **Clinical Integration**: AI decision support systems embedded in EHRs.
- **Ethical Governance**: Transparent, equitable AI with patient oversight.

Conclusion:

The integration of multi-omics data with artificial intelligence represents a transformative approach to understanding and treating diabetes mellitus. By modeling complex gene–cell interactions across multiple biological layers, AI enables the discovery of disease endotypes, biomarkers, and therapeutic targets with unprecedented resolution. Case studies in T1D and T2D demonstrate the potential for personalized risk prediction, early diagnosis, and optimized treatment selection. However, challenges in data quality, model interpretability, and clinical translation remain. Addressing these requires interdisciplinary collaboration, investment in infrastructure, and ethical oversight. As AI and multi-omics mature, they will move from research tools to clinical assets, ushering in a new era of precision medicine in diabetes care.

Multi-omics artificial intelligence modeling represents a paradigm shift in understanding and treating diabetes. By integrating genomics, epigenomics, transcriptomics, proteomics, metabolomics, and microbiomics, AI can decode the complex gene–cell interactions underlying diabetes pathogenesis. This enables the identification of disease endotypes, prediction of treatment response, discovery of biomarkers, and personalization of therapy. While challenges in data integration, model interpretability, and clinical translation persist, the convergence of advanced omics technologies with AI offers unprecedented opportunities for precision medicine in diabetes. A collaborative, interdisciplinary approach involving biologists, clinicians, data scientists, and ethicists will be essential to transform these innovations into improved patient outcomes.

References:

1. International Diabetes Federation (2021). IDF Diabetes Atlas, 10th ed.
2. Ahlqvist, E. *et al.* (2018). Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis. *The Lancet Diabetes & Endocrinology*, 6(5), 361–369.
3. Andersson, J. *et al.* (2018). Risk profiling of diabetes subgroups in a Swedish population. *Diabetologia*, 61(10), 2174–2184.
4. Argelaguet, R. *et al.* (2018). Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology*, 14(6), e8124.
5. Fadista, J. *et al.* (2022). Global genomic and transcriptomic analysis of human pancreatic islets. *Nature Genetics*, 54(5), 655–666.

6. Mahajan, A. *et al.* (2022). Multi-ancestry genetic study of type 2 diabetes. *Nature*, 606(7916), 751–759.
7. Onengut-Gumuscu, S. *et al.* (2015). Fine mapping of type 1 diabetes susceptibility loci. *Nature Genetics*, 47(8), 811–816.
8. Segerstolpe, Å. *et al.* (2016). Single-cell transcriptome profiling of human pancreatic islets. *Cell Metabolism*, 24(4), 593–607.
9. Volkmar, M. *et al.* (2012). DNA methylation profiling of human pancreatic islets. *Cell Metabolism*, 15(1), 105–111.
10. Zhang, Z. *et al.* (2021). OmicsNet: a web-based tool for creation and visual analysis of biological networks in multi-omics data. *Nucleic Acids Research*, 49(W1),W497–W507.
11. International Diabetes Federation. IDF Diabetes Atlas, 10th edn. 2021.
12. Ahlqvist, E. *et al.* (2018). Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The Lancet Diabetes & Endocrinology*, 6(4), 361–369.
13. Saeed, S. *et al.* (2020). Multi-omics of human pancreatic islets from type 2 diabetes donors. *Cell Metabolism*, 31(4), 745–759.
14. Chen, L. *et al.* (2021). Deep learning-based integration of multi-omics data in cancer. *Nature Communications*, 12, 1090.
15. Arnes, L. *et al.* (2021). Single-cell transcriptomics of the human endocrine pancreas. *Diabetes*, 70(8), 1758–1768.
16. Zou, J. *et al.* (2019). A primer on deep learning in genomics. *Nature Genetics*, 51(1), 12–18.
17. Efremova, M. *et al.* (2020). CellPhoneDB: Inferring cell–cell communication from combined expression of ligands and receptors. *Nature Protocols*, 15, 1484–1506.
18. Min, S. *et al.* (2017). Deep learning in bioinformatics. *Briefings in Bioinformatics*, 18(6), 851–869.
19. Wagner, R. *et al.* (2022). Machine learning approaches for diabetes prediction and subtyping. *Diabetologia*, 65(7), 1107–1120.
20. Consortium, T.H.P.A.P. (2021). Multi-omics profiling reveals drivers of β-cell dysfunction in type 2 diabetes. *Nature Metabolism*, 3(8), 1066–1081.

NEXT-GEN PV INVERTER:
SEAMLESS DC TO GRID-COMPATIBLE AC POWER

N. Ismayil Kani
Department of Electrical and Electronics Engineering,
AMET University, Kanathur - 603 112, Chennai
Corresponding author E-mail: ismayilkani@gmail.com

Abstract:
The integration of photovoltaic (PV) systems into utility grids depends critically on the ability of inverters to convert DC power from solar panels into grid-compatible AC power with high efficiency, low distortion, and robust grid support functionality. This paper surveys and proposes advancements in PV inverter topology, control strategies, and smart inverter features to realize seamless DC-to-AC conversion suitable for modern grids. Emphasis is placed on transformer less high-efficiency designs, advanced control algorithms, and grid-forming capabilities that collectively enhance power quality and system flexibility.

Keywords: PV Inverter, DC to AC Conversion, Grid-Connected PV System .Transformer less Inverter, Power Electronic.

1. Introduction:

Photovoltaic (PV) systems generate direct current (DC) electricity, but grid integration requires alternating current (AC) with precise voltage, frequency, and phase characteristics. The inverter plays a pivotal role in this conversion, influencing system efficiency, power quality, and grid stability. Next-generation PV inverters aim not only to perform basic DC to AC conversion but also to meet stringent standards for harmonics, reactive power control, and grid support.

2. Literature Review Paragraphs

Panagopoulos, A. *et al.* in their paper provides a comprehensive overview of grid-connected PV inverter technologies, exploring topologies, control strategies, and international grid-code requirements. The authors analyze multiple inverter configurations, emphasizing how modern designs must not only perform efficient DC-to-AC conversion but also meet stringent utility specifications for harmonics, reactive power support, and safety. The study compares international standards while highlighting trends toward advanced control strategies that improve synchronization and stability. This makes it a cornerstone reference for understanding how next-generation PV inverters must perform both conversion and grid interaction functions.

This review examines single-phase transformerless inverter topologies, a key component in modern PV systems due to their high efficiency, reduced size, and lower cost compared to transformer-based solutions. The authors compare eighteen well-known inverter structures using common simulation parameters to fairly assess performance metrics including common mode voltage, leakage current, switching and conduction loss, and efficiency. The study's critical analysis of leakage current behavior directly informs design choices for seamless DC to grid-compatible AC conversion with minimal power quality issues.

Grid-Connected Inverter Topologies for High-Penetration Solar PV Systems: A Comparative Study.This article conducts a detailed comparative evaluation of various grid-connected inverter systems, including conventional voltage source inverters (VSIs), multilevel inverters (MLIs), and transformerless designs. Beyond just comparing structures, the study assesses modularity, scalability, fault tolerance, and harmonic distortion, demonstrating that each topology trades off complexity and performance metrics such as efficiency and THD. This work is particularly relevant for your article's inverter design discussion, as it highlights how different topological choices impact real-world utility integration.

Khan and colleagues offer a detailed synthesis of inverter classifications, modulation methods, and control strategies used in grid-connected PV systems. They explain how multilevel inverter topologies reduce switching stress and output harmonics compared to conventional two-level designs. The paper also reviews modulation strategies such as PWM and control reference frames (dq, αβ, abc), providing context for how control systems affect inverter performance. This comprehensive perspective strengthens the understanding of how inverter topology and control interact to achieve seamless grid compatibility.

Grid-Connected PV Inverter System Control Optimization Using Grey Wolf Optimized PID Controller.This research demonstrates the application of optimized control algorithms to enhance grid-connected PV inverter performance. Using a Grey Wolf Optimizer to tune PID controller parameters, the authors achieve improved DC-link voltage stability, reduced total harmonic distortion, and better dynamic response under varying irradiance and grid conditions. The study exemplifies how advanced control strategies can significantly improve the quality of AC power delivered to the grid, making it highly relevant to the control systems section of next-generation inverter designs.

Shrestha *et al.* investigate several transformer-less inverter topologies — including H4, H5, and HERIC — focusing on common-mode voltage (CMV), total harmonic distortion (THD), and leakage currents. Their simulation results indicate that certain topologies (e.g., H5) offer optimized performance for reducing leakage and improving waveform quality. These findings help illustrate practical tradeoffs in inverter design, particularly for ensuring seamless DC-to-AC conversion with minimal safety and interference concerns.

Premkumar *et al.* focuses on microinverters — smaller, module-level inverters used in distributed PV systems. It surveys various control schemes and topological configurations, revealing how localized inverter control can enhance overall system performance and reliability.

Emphasis is placed on control methods that maintain high power quality and ensure efficient grid synchronization despite module-level variations. This microinverter perspective complements macro inverter discussions and highlights future trends in inverter decentralization and intelligence.

A comprehensive review on inverter topologies and control strategies for grid connected photovoltaic system synthesizes decades of research on grid-connected inverter structures and control approaches. It presents a broad classification of inverters, assesses their respective advantages, and examines common control methods (e.g., PI, predictive control). The review also discusses emerging trends and selection guidelines for choosing appropriate topologies based on application requirements, making it a valuable reference for framing the historical context and evolution toward next-generation systems.

Although earlier, this review remains relevant by documenting the evolution of single-phase inverter designs. It details structural differences among traditional and modern inverter topologies, including soft-switching and multilevel approaches, while evaluating their influence on performance metrics such as efficiency and THD. The paper provides historical continuity that supports understanding how newer designs emerged to address earlier limitations in grid integration.

Phuyal *et al.* proposed an optimized H5 topology with clamped diodes and hysteresis current control to maintain constant common-mode voltage and reduce leakage current. Their simulation demonstrates improved inverter performance compared to conventional H4 configurations, underscoring how topological innovations paired with effective control algorithms can produce higher quality AC output with fewer power quality issues. This study supports your article's advanced topology discussion and demonstrates practical performance improvement techniques.

3. Transformerless High-Efficiency Inverter Topologies

Recent research indicates transformerless inverter structures are increasingly popular due to their reduced size, cost, and higher efficiency compared to transformer-based counterparts. A novel transformerless single-phase inverter design has achieved simulation efficiencies up to 99.77% and real-world lab performance above 98% while keeping total harmonic distortion (THD) and leakage current extremely low.

3.1 Leakage Current Mitigation

Transformerless designs often face safety challenges due to leakage currents caused by stray capacitance. Techniques such as maintaining constant common-mode voltage and using multilevel outputs help effectively suppress leakage and enhance safety without compromising efficiency.

4. Control Strategies for Seamless Integration

Effective control algorithms are crucial for synchronizing inverter output with grid voltage and frequency, minimizing waveform distortion, and ensuring seamless power injection.

4.1 Sliding Mode Control

Modified sliding mode control (MSMC) has been shown to reduce waveform distortion during DC to AC conversion by ensuring smoother transitions and better dynamic response, critical for avoiding operational issues in the utility grid.

4.2 Optimized PID Control

Advanced optimization methods, such as Grey Wolf optimized PID controllers, improve DC-link voltage stability and reduce THD in grid-connected PV inverters, enhancing both efficiency and power quality.

5. Grid-Forming and Grid-Supporting Features

Traditional grid-following inverters rely on grid voltage and frequency references, but next-gen systems increasingly incorporate grid-forming capabilities. Such designs maintain DC bus voltage and help stabilize the grid during disturbances, particularly in high PV penetration scenarios. Smart inverter functions also include reactive power support, advanced MPPT (Maximum Power Point Tracking), and bidirectional power flow when integrated with battery storage, enabling more flexible operation and better utilization of renewable energy.

6. System Integration and Smart Grid Compatibility

Next-generation inverters are being designed not only for conversion but also for *grid interaction*. Smart functionalities—such as adaptive control, dynamic reactive power management, and communication with grid automation systems—allow PV inverters to participate in grid stability and ancillary services. This represents a shift toward inverters acting as *active grid elements* rather than passive converters.

7. Performance Metrics and Power Quality

Key performance indicators for PV inverters include conversion efficiency, total harmonic distortion (THD), power factor control, and quick synchronization to grid parameters. Next-gen designs focus on achieving:

- High efficiency (often > 98 %)
- Low THD to meet grid codes
- Stable MPPT under variable irradiance
- Flexible grid support roles

These metrics are increasingly critical as renewable penetration grows and grid codes become more stringent.

8. Challenges and Future Directions

Despite progress, challenges remain in mitigating leakage currents for transformerless designs, meeting international grid codes across different regions, and developing robust control algorithms that handle grid disturbances and variable solar conditions. Future research may integrate advanced materials (e.g., GaN devices), AI-based controls, and enhanced interoperability with energy storage systems to further improve inverter performance and grid resilience.

9. Advanced Inverter Topologies and Structural Trends

9.1 Multi-Level and Transformerless Designs

Modern grid-connected PV inverters increasingly use *multi-level and transformerless topologies* to improve efficiency, reduce size, and cut costs. These configurations achieve better harmonic performance and higher DC/AC conversion quality compared with traditional two-level VSI structures. They also offer advantages in **modularity and scalability** for large-scale PV installations.

9.2 Two-Stage High-Frequency Link Topologies

Some research proposes *two-stage inverter designs with high-frequency link transformers* to enhance MPPT performance, reduce leakage currents, and improve power quality. These designs use a buck-boost inverter in the first stage to handle varying PV voltages, and a high-frequency link to the grid interface, which boosts reliability and reduces filter requirements.

9.3 GaN-Based and Wide-Bandgap Devices

Emerging work shows that *GaN-based inverters* can significantly improve efficiency and reduce THD for grid integration. By combining hybrid control schemes such as sliding mode and advanced PLL algorithms, these inverters enhance both MPPT performance and synchronization with grid voltage.

10. Power Quality Enhancement and Control Strategies

10.1 Adaptive Control for Dynamic Grid Conditions

Control strategies that dynamically adjust to grid disturbances (such as voltage sags, unbalanced conditions, or weak grids) are becoming essential. Adaptive positive sequence estimators and feed-forward terms help maintain THD compliance with standards like IEEE 519, even under challenging operating conditions.

10.2 Intelligent Controller Optimization

Advanced optimization methods (e.g., *Grey Wolf Optimized PID*, fractional controllers, and hybrid strategies) enhance inverter current and voltage regulation under variable irradiance and grid fluctuations. These techniques reduce harmonics and improve transient response well beyond conventional PI control.

11. Power Quality and Grid Code Compliance

Inverters must meet stringent **grid codes and power quality requirements** to safely interconnect with utility systems. Research highlights that international regulations vary, and mismatches in grid codes can increase design complexity across regions. Features such as reactive power support, fast synchronization, and low THD are now standard expectations in modern inverters.

11.1 Space Vector PWM and THD Reduction

Advanced modulation techniques like *Space Vector Pulse Width Modulation (SVPWM)* are shown to reduce switching losses and THD compared to traditional sinusoidal PWM. These strategies improve voltage utilization and dynamic response, especially in multi-level inverter topologies.

12. Smart Grid Integration and Ancillary Services

12.1 Reactive Power Management and Volt/VAR Support

Inverters are no longer passive converters — they actively manage **reactive power** to help maintain voltage stability, absorb or supply VARs as needed, and support distribution network operations. This local control becomes critical as PV penetration increases.

12.2 Smart Inverters and Grid Flexibility

Research on smart inverter technologies emphasizes their role in *flexible renewable penetration*. Smart features include dynamic power flow control, grid support algorithms, communication interfaces, and compatibility with utility automation systems. These capabilities enable inverters to participate in grid stability and optimization rather than simply injecting AC power.

13. Market & Industry Trends Driving Technology

Beyond technical performance, market forecasts show **significant growth** in PV inverter adoption, with innovations in high-voltage string inverter platforms, improved power density, and intelligent predictive maintenance systems. Enhanced hardware like SiC and GaN semiconductors, as well as AI-driven monitoring, are key drivers of next-generation inverter development.

14. Challenges, Limitations, and Future Research Directions

While technical advances are promising, several challenges remain:

- **Grid code harmonization:** Lack of unified international standards complicates inverter design and certification.
- **Dynamic stability under high PV penetration:** In-depth investigation of voltage and frequency control for weak grid conditions is essential in future research.
- **Integration with energy storage:** Coordinating energy storage with inverter control for peak smoothing and demand response is an open area for innovation.

Conclusion:

Next-generation PV inverters are transforming how solar energy is integrated with power grids. By focusing on efficient transformerless topologies, advanced control strategies, grid-forming capabilities, and smart grid features, these systems achieve seamless DC to grid-compatible AC conversion while supporting grid stability and performance. Continued innovation will further enhance reliability and scalability, positioning PV inverters as fundamental components of future renewable energy networks.

References:

1. Panagopoulos, A., *et al.* (2024). Grid-connected photovoltaic inverters: Grid codes, topologies and control techniques. *Renewable and Sustainable Energy Reviews*.
2. Gupta, M., Tiwari, P. M., & Viral, R. K. (2025). Grid-connected PV inverter system control optimization using Grey Wolf optimized PID controller. *Scientific Reports*.
3. Vishnupriya, T. (2025). Grid-connected inverter topologies for high-penetration solar PV systems: A comparative study. *Transactions on Power Electronics and Renewable Energy Systems*.
4. Issac, E. J., & Rajeev, M. (2020). A survey of current source inverter topologies and control schemes for grid-connected photovoltaic systems.
5. Jain, S., & Agarwal, V. (2007). A single-stage grid-connected inverter topology for solar PV systems with maximum power point tracking. *IEEE Transactions on Power Electronics*.
6. Yüksel, A., & Özkop, E. (2019). Control of single-phase grid-connected transformerless PV inverter system.
7. Sarwar, A., & Asghar, M. S. J. (2011). Simulation and analysis of a multilevel converter topology for solar PV-based grid-connected inverter. *Smart Grid and Renewable Energy*.
8. Wang, Y. J., & Sui, L. L. (2011). Research on single-phase photovoltaic generation under grid-connected inverter. *Applied Mechanics and Materials*.
9. Zarikar, S., & Bachawad, M. (2019). Modelling of grid-connected DC/AC converter for photovoltaic application. *International Journal of Advanced Trends in Computer Science and Engineering*.
10. Öztürk, S., *et al.* (2018). All SiC grid-connected PV supply with HF link MPPT converter: System design methodology and development of a 25 kVA prototype. *Electronics*.

A COMPREHENSIVE REVIEW OF SEMICONDUCTOR MEMORIES

Kanchan Singh
IFTM University, Moradabad, Uttar Pradesh 244102
Corresponding author E-mail: 01kanchansingh@gmail.com

1. Introduction:

Semiconductor memories are electronic storage devices fabricated using integrated circuit (IC) technology. They store digital information in binary form (0s and 1s) and are essential components of computers, microprocessors, microcontrollers, and modern digital systems. Compared to magnetic and mechanical memories, semiconductor memories offer high speed, compact size, low power consumption, and high reliability.[1]

Semiconductor memories are broadly classified into volatile and non-volatile types. Volatile memories lose their data when power is switched off, whereas non-volatile memories retain information even in the absence of power.[1]

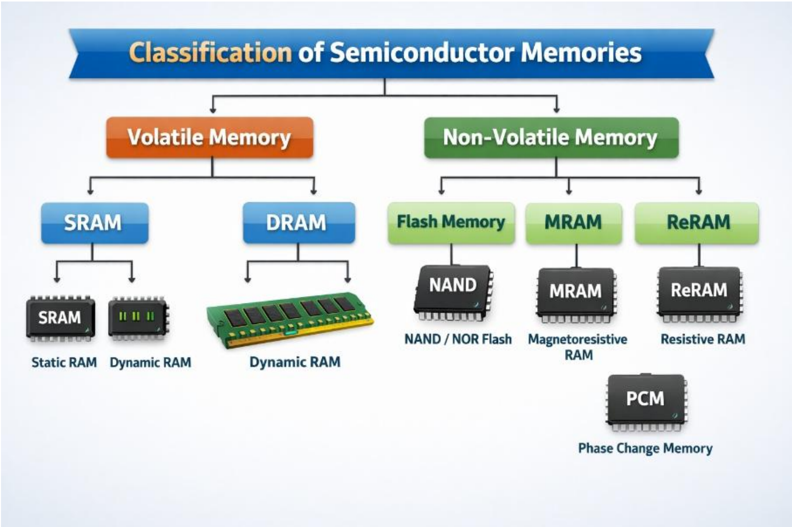


Figure 1: Basic classification of semiconductor memories[1]

2. Memory Organisation and Operation

A semiconductor memory consists of a large number of memory cells arranged in rows and columns. Each memory cell stores one bit of information, and the overall memory organisation determines how data is stored, accessed, and retrieved.[1]

2.1 Memory Organisation

- Memory cells form a memory array.[1]
- Address lines select a specific memory location.[1]
- Row and column decoders activate the selected word and bit lines.[1]
- Data lines (I/O lines) transfer information to and from the memory.[1]

2.2 Read and Write Operations

- **Read operation:** The addressed memory cell is selected, and its stored data is placed on the data bus.[1]
- **Write operation:** The input data is applied to data lines and written into the selected cell when the write control signal is active.[1]

3. Expanding Memory Size

When the required memory capacity exceeds that of a single IC, multiple memory devices are combined to expand the total size or word length.[1]

3.1 Methods of Memory Expansion

- **Capacity expansion:** Increases the number of memory locations by adding more ICs with additional address decoding.[1]
- **Word-length expansion:** Increases the number of bits per word by connecting memory ICs in parallel on the data bus.[1]
- Combined expansion: Simultaneously increases memory size and word length using both additional address lines and parallel data paths.[1]

4. Classification and Characteristics of Memories

4.1 Broad Classification

- **Read-Only Memory (ROM):** Non-volatile memory with fixed data, typically programmed once or infrequently.[1]
- **Random Access Memory (RAM):** Volatile memory that allows read and write operations.[1]
- **Special memories:** Application-oriented devices such as content addressable memory (CAM), first-in–first-out (FIFO) memories, and charge-coupled device (CCD) memories.[1]

4.2 Key Characteristics

Important figures of merit for semiconductor memories include:[1]

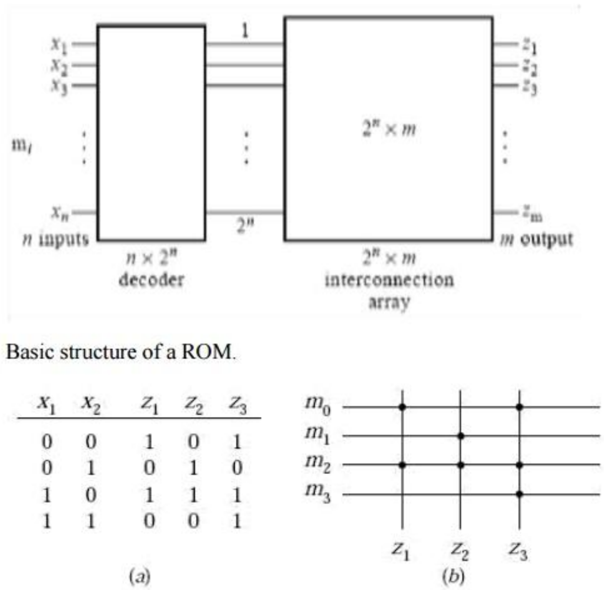
- Access time (read/write latency)
- Cycle time
- Power consumption
- Storage density (bits per unit area)
- Cost per bit
- Reliability and endurance

5. Read-Only Memory (ROM)

ROM is a non-volatile memory used to store permanent programs such as firmware, boot loaders, and microcontroller lookup tables. Data in ROM is typically not altered during normal operation.[1]

5.1 Types of ROM

- Mask ROM: Programmed during fabrication; very low cost per bit for high volumes.
- PROM (Programmable ROM): User-programmable once using special equipment.
- EPROM (Erasable PROM): Can be erased using ultraviolet light and reprogrammed.
- EEPROM (Electrically Erasable PROM): Electrically erasable and reprogrammable at the byte or block level; basis for early non-volatile storage and precursors to flash.[1]



x_1	x_2	z_1	z_2	z_3
0	0	1	0	1
0	1	0	1	0
1	0	1	1	1
1	1	0	0	1

Figure 2: ROM organisation (address decoder and ROM matrix)[1]

6. Read and Write Memory (RAM)

RAM is a volatile memory that permits both reading and writing and is used as the main working memory in digital systems.[1]

6.1 Static RAM (SRAM)

SRAM stores each bit in a bistable flip-flop. It offers high speed and does not require refresh, but has lower density and higher cost per bit than DRAM. SRAM is widely used in caches for CPUs and AI accelerators due to its fast access time.[2][1]

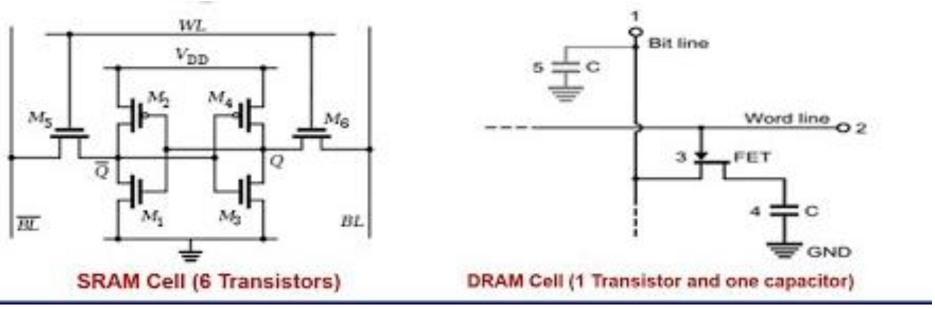


Figure 3: SRAM and DRAM cell representation (flip-flop vs capacitor cell)[1]

6.2 Dynamic RAM (DRAM)

DRAM stores each bit as charge on a capacitor and therefore requires periodic refresh. It provides very high density and is used for main memory in computers, servers, and AI hardware.

Recent DRAM generations (e.g., DDR5 and 3D-stacked high-bandwidth memory, HBM) aim to alleviate the memory-bandwidth bottleneck in AI and high-performance computing systems.[3][4][1]

7. Flash Memory

Flash memory is a non-volatile memory derived from EEPROM technology and supports electrical erasing and programming of blocks of cells. It is widely used in solid-state drives (SSDs), mobile devices, memory cards, and embedded systems.[5][1]

7.1 Features of Flash Memory

- High storage density and scalability using multi-level and 3D cell structures.[5][1]
- Low power consumption, making it suitable for portable electronics.[1]
- High read speed and good random-access characteristics.[1]

7.2 Types of Flash

- **NOR flash:** Provides random byte-level access; used where direct code execution (eXecute-In-Place) is needed.[1]
- **NAND flash:** Optimised for high-density storage and block operations; dominates SSD and removable storage markets. Recent research explores split-cell and multi-bit-per-cell architectures to further increase capacity per die.[5][1]

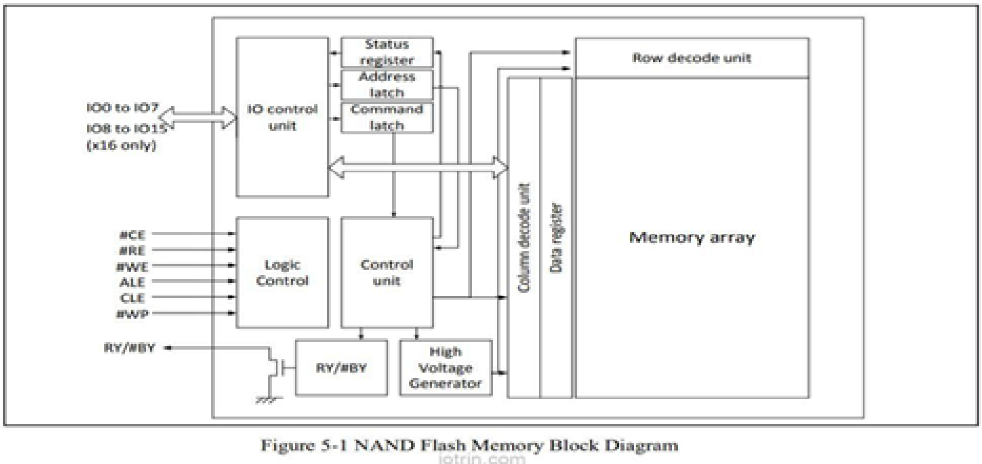


Figure 4: Flash memory block diagram (controller and array)[1]

8. Content Addressable Memory (CAM)

Content addressable memory retrieves data by comparing input data with the stored content instead of using an explicit address. In a CAM, all stored words are compared in parallel, and any matching word(s) assert a match output.[1]

CAMs are widely used in high-speed search applications such as packet classification, cache tag lookup, and routing tables in communication systems.[1]

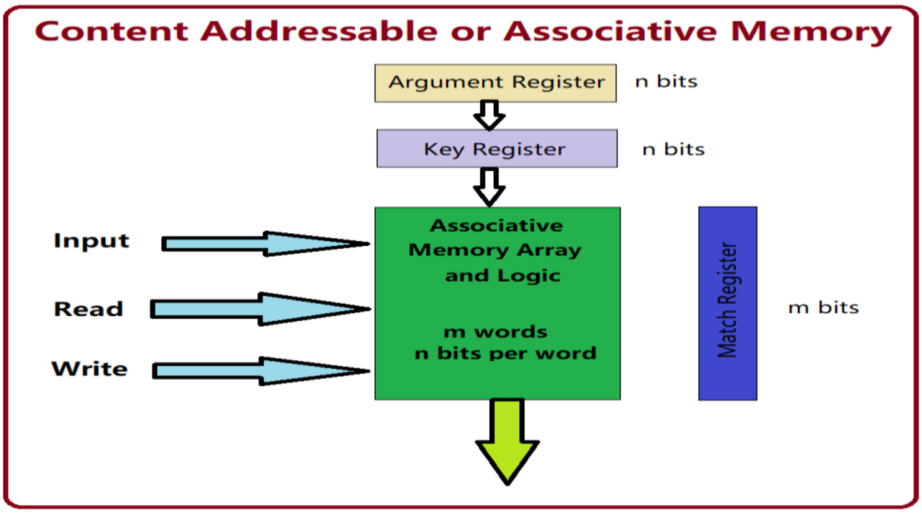


Figure 5: CAM structure (comparator array and match outputs)[1]

9. First-In, First-Out Memory (FIFO)

FIFO memories output data in the same order in which it was written. They use separate read and write pointers instead of direct addressing and are useful for buffering and rate-matching between subsystems operating at different speeds.[1]

Typical applications include serial communication buffers, streaming data interfaces, and pipeline staging in digital signal processing systems.[1]

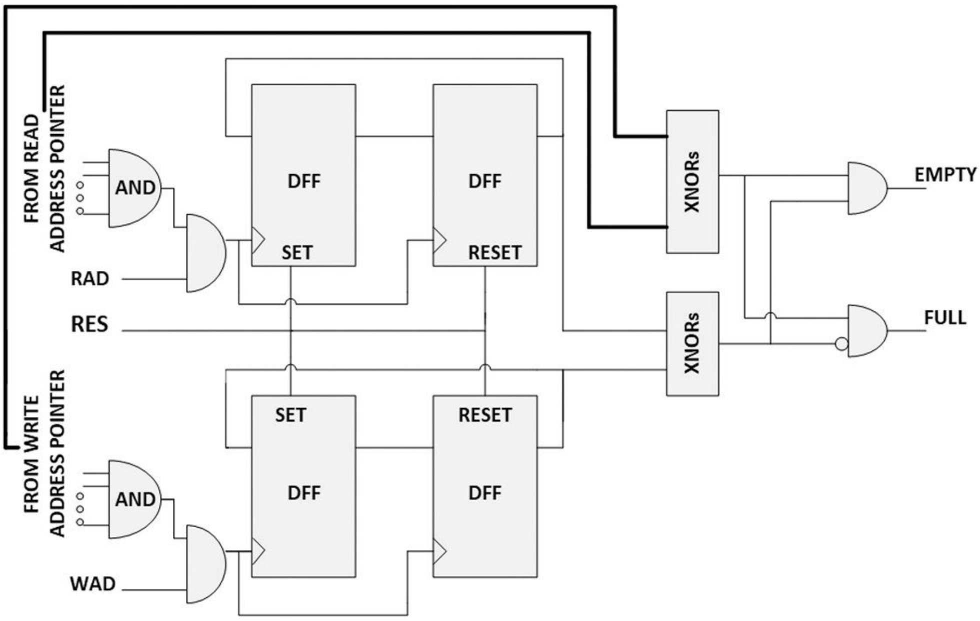


Figure 6: FIFO memory block diagram (data in, data out, write and read pointers)[1]

10. Charge-Coupled Device (CCD) Memory

CCD memory stores information as packets of electric charge and transfers it serially from one potential well to the next using clock pulses. It offers very high packing density and low noise but only supports serial access, making it suitable for image sensors and certain specialised signal-processing applications.[1]

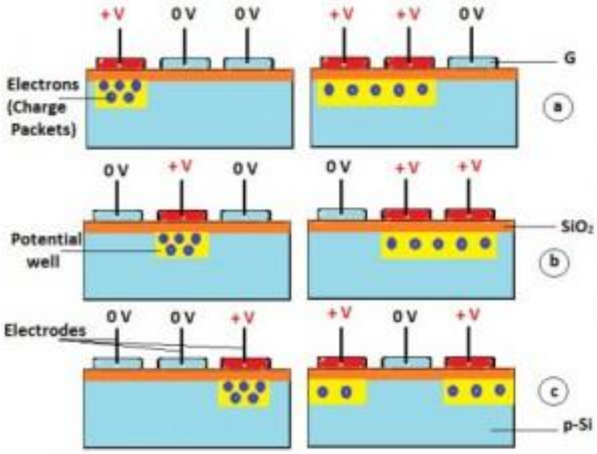


Figure 7: CCD charge transfer mechanism (serial charge shift under clock control)[1]

11. Emerging Non-Volatile and AI-Oriented Memories

Beyond conventional ROM, SRAM, DRAM, and flash, several emerging memory technologies are being developed to address scaling limits, energy efficiency, and the memory wall in AI systems.[6][2]

11.1 Resistive RAM (ReRAM) and Memristive Devices

Resistive RAM (ReRAM) stores information as a change in resistance of a switching material and offers high endurance, fast switching, and good scalability, making it a strong candidate to complement or partially replace flash in future non-volatile memories. Recent work demonstrates oxide-based and graphene-oxide ReRAM devices with improved flexibility, endurance, and retention suitable for embedded and neuromorphic applications.[7][5]

Memristor-based arrays can implement analogue in-memory computing, where computation (e.g., matrix–vector multiplication) occurs directly in the memory crossbar, significantly reducing data movement and improving energy efficiency for AI workloads.[2][7]

11.2 Ferroelectric Memories (FeRAM and FeFET)

Ferroelectric random-access memory (FeRAM) and ferroelectric field-effect transistors (FeFETs) exploit the remanent polarisation of ferroelectric materials to store data non-volatilely. Recent studies on hafnium-oxide-based ferroelectric capacitors show low-voltage operation, high endurance (>10^8 cycles), and compatibility with CMOS back-end-of-line processes, enabling dense on-chip embedded non-volatile memories. Ferroelectric transistors also show promise for neuromorphic and in-memory computing due to their multi-level conductance and analogue switching behaviour.[8][9][6]

11.3 High-Bandwidth Memory (HBM) and 3D DRAM

High-bandwidth memory (HBM) is a 3D-stacked DRAM integrated with processors or accelerators using through-silicon vias (TSVs) and a wide I/O interface. HBM provides order-of-magnitude higher bandwidth and better energy efficiency per bit than traditional DDR memory and is now a key enabler for GPUs and accelerators used in large-scale AI training and high-performance computing. Recent HBM generations (e.g., HBM3) support higher stack heights and wider channels, delivering multi-terabyte-per-second bandwidth to keep AI compute engines fully utilised.[4][3]

12. Current Trends and Market Developments

12.1 AI-Driven Memory Demand

The rapid growth of AI and machine-learning workloads has created a strong demand for high-capacity and high-bandwidth DRAM and HBM, especially in data-center accelerators. This demand has contributed to periodic DRAM shortages and price increases, as memory manufacturers prioritise high-margin AI-oriented products.[10][11][12]

12.2 Advanced NAND and Storage

Advanced 3D NAND technologies with more layers and multi-bit-per-cell schemes are pushing SSD capacities higher while maintaining competitive cost per bit. Research into split-cell and multi-level flash architectures seeks to further increase storage density for data-center and client applications.[5]

12.3 In-Memory and Neuromorphic Computing

To overcome the von Neumann bottleneck and the memory wall, in-memory computing architectures using memristors, ReRAM, and ferroelectric devices are actively explored as hardware accelerators for AI. These approaches aim to co-locate storage and computation, thereby reducing data movement and improving energy efficiency for large neural networks.[6][7][2]

References:

1. Hwang, H., *et al.* (2025). Recent advances in ferroelectric materials, devices, and in-memory computing.
2. Shukla, N., *et al.* (2025). Ferroelectric memory technology for big data applications.
3. Ni, Y., *et al.* (2022). Ferroelectric transistors for memory and neuromorphic computing. *Advanced Materials*.
4. Lee, J., *et al.* (2025). Recent advances in resistive switching memory devices.
5. Chen, X., *et al.* (2024). Memristor-based hardware accelerators for artificial intelligence. *Nature Reviews Electrical Engineering*.
6. Park, S., *et al.* (2025). A review of memristor-based computing-in-memory architectures.
7. Ferro, F. (2024). Accelerating AI/ML applications in the data center with HBM3. Rambus.
8. Intuition Labs. (2026). *HBM vs. DDR: Key differences in memory technology*.
9. EE Times Asia. (2026). HBM, DRAM, and NAND: How AI is reshaping the memory market.
10. Edge-AI Vision. (2026). Why DRAM prices keep rising in the age of AI.
11. Hailo. (2026). DRAM shortage 2026: Why edge AI is the solution.

CHEMICAL BATH DEPOSITED COBALT SELENIDE THIN FILMS: STRUCTURAL, OPTICAL, AND ELECTRICAL PROPERTIES

Prashant A. Chate*[1] and Dattatray J. Sathe[2]
[1]Department of Chemistry, J.S.M. College, Alibag (M.S.) India
[2]Department of Chemistry, KIT College of Engg (Autonomous), Kolhapur (M.S.) India.
*Corresponding author E-mail: pachate04@rediffmail.com, djsathe77@gmail.com

1. Introduction:

Transition metal chalcogenides (TMCs) represent a broad and technologically significant class of inorganic materials that have attracted intense research interest due to their unique structural, electronic, optical, and electrochemical properties. These compounds are generally expressed by the chemical formula MX_x, where M denotes a transition metal element such as Mo, W, Co, Ni, Fe, or Cu, and X represents a chalcogen atom, typically sulfur (S), selenium (Se), or tellurium (Te). The ability of transition metals to adopt multiple oxidation states and the versatile bonding nature of chalcogen elements give rise to a wide range of stoichiometries, crystal structures, and physical properties. As a result, TMCs span a spectrum of behaviors from semiconducting and metallic to semi-metallic and superconducting, making them attractive for diverse applications in electronics, optoelectronics, catalysis, sensing, and energy conversion technologies. A defining feature of many TMCs is their layered or quasi-layered crystal structure, in which strongly covalently bonded metal–chalcogen units form two-dimensional sheets that are weakly coupled by van der Waals interactions. This structural anisotropy leads to distinctive electronic band structures, anisotropic charge transport, and high surface reactivity. Even in non-layered TMCs, the presence of chalcogen atoms with relatively large atomic radii and polarizable electron clouds contributes to enhanced electronic delocalization and strong light–matter interaction. These features are particularly advantageous for applications requiring efficient charge generation, separation, and transfer, such as photoelectrochemical (PEC) energy conversion and electrocatalysis.

Among the chalcogen elements, selenium-based TMCs have gained increasing attention in recent years due to their relatively narrow band gaps, higher electrical conductivity, and stronger absorption of visible light compared to their sulfur-based counterparts. Selenium possesses a lower electronegativity and larger atomic radius than sulfur, resulting in reduced metal–chalcogen bond ionicity and stronger orbital hybridization between metal d states and selenium p states. This enhanced hybridization often leads to smaller band gap energies, higher carrier mobility, and improved charge transport characteristics. Consequently, transition metal selenides are particularly promising for applications in photovoltaics, PEC cells, and electrochemical energy storage and conversion devices.

Within this family, cobalt selenide (CoSe) and its related phases including $CoSe_2$, Co_9Se_8, Co_3Se_4, and other non-stoichiometric compositions have emerged as highly attractive functional materials. Cobalt selenides exhibit rich structural polymorphism and complex electronic structures arising from the multiple valence states of cobalt ions (Co^{2+} and Co^{3+}). This multivalence enables facile redox reactions and rapid electron transfer, which are essential for electrochemical and photoelectrochemical processes. Depending on composition and crystal structure, cobalt selenides can exhibit semiconducting, semi-metallic, or metallic behavior, providing flexibility for tailoring their properties for specific applications [1-2]. Compared to cobalt sulfide analogues, cobalt selenide materials generally demonstrate superior electronic conductivity and enhanced optical absorption in the visible region. The larger size and lower electronegativity of selenium relative to sulfur reduce the band gap and increase electronic delocalization, leading to improved charge carrier mobility [3]. These characteristics are particularly beneficial for photoelectrode materials, where efficient transport of photogenerated charge carriers to the electrode–electrolyte interface is critical for achieving high photocurrent densities and conversion efficiencies. Furthermore, cobalt selenides often exhibit improved chemical stability in alkaline and neutral electrolytes, which is advantageous for long-term PEC and electrocatalytic operation. The functional performance of cobalt selenide is strongly influenced by its physical form, with thin films offering several distinct advantages over bulk or powder materials. Thin film architectures provide a continuous and adherent layer on conductive or transparent substrates, facilitating efficient charge collection and integration into device configurations. The reduced thickness of thin films shortens the diffusion length for charge carriers, thereby minimizing bulk recombination losses and enhancing charge separation efficiency under illumination [4]. Additionally, thin films allow precise control over composition, thickness, crystallinity, and surface morphology, which are critical parameters governing optical absorption, electrical conductivity, and interfacial charge transfer kinetics.

Nanostructured and polycrystalline CoSe thin films are particularly attractive due to their high surface-to-volume ratio and abundance of exposed active sites. Grain boundaries, surface defects, and edge sites in such films can act as catalytic centers and charge transfer pathways, significantly enhancing electrochemical and photoelectrochemical activity. These features are especially beneficial in applications such as hydrogen evolution reaction, oxygen evolution reaction, and PEC water splitting, where surface reactions play a dominant role. Furthermore, nanostructuring can induce quantum confinement effects and strain-related band structure modifications, offering additional routes for tuning the optical and electronic properties of CoSe thin films.

Owing to these advantages, CoSe thin films have been explored for a wide range of applications, including photoelectrochemical solar cells, electrocatalytic hydrogen production, supercapacitors, lithium- and sodium-ion batteries, and chemical and biological sensors. In PEC systems, cobalt selenide has been investigated as a photoelectrode material or as a catalytic over layer to enhance charge transfer and suppress surface recombination. Its suitable band edge positions relative to water redox potentials, combined with strong visible-light absorption and fast interfacial kinetics, make it a promising candidate for solar-driven water splitting. In electrocatalysis, cobalt selenide thin films exhibit low over potentials and favorable Tafel slopes for hydrogen evolution reaction, attributed to optimal hydrogen adsorption energies and the presence of catalytically active cobalt sites. The properties of CoSe thin films are highly sensitive to the synthesis method employed, as deposition techniques strongly influence phase composition, crystallinity, morphology, and defect density. A variety of physical and chemical methods have been reported for the preparation of cobalt selenide thin films, including radio-frequency magnetron sputtering, thermal evaporation, pulsed laser deposition, electrodeposition, solvothermal growth, and chemical bath deposition (CBD) [5-6]. Physical vapor deposition techniques generally offer excellent control over film thickness and purity, producing dense and highly crystalline films. However, these methods require sophisticated equipment, high vacuum conditions, and elevated processing temperatures, which increase production costs and limit scalability. Chemical synthesis routes, on the other hand, provide cost-effective and scalable alternatives for thin film fabrication. Among these, chemical bath deposition has emerged as one of the most attractive techniques for depositing metal chalcogenide thin films. CBD is a solution-based process that enables controlled film growth through the slow release of metal and chalcogen ions in an aqueous medium. The method typically operates at relatively low temperatures and ambient pressure, making it compatible with a wide variety of substrates, including glass, fluorine-doped tin oxide, indium tin oxide, stainless steel, and even flexible polymer substrates. One of the key advantages of CBD is its ability to produce uniform and adherent films over large areas without the need for complex instrumentation. Film thickness, composition, and morphology can be precisely tuned by adjusting deposition parameters such as precursor concentration, pH of the solution, type and concentration of complexing agents, bath temperature, and deposition time. Complexing agents play a crucial role in controlling the release rate of metal ions, thereby influencing nucleation and growth processes. Similarly, the choice of selenium source and reducing agents determines the kinetics of chalcogenide formation and phase evolution. CBD grown CoSe thin films often exhibit nanocrystalline or polycrystalline structures with controllable grain size and surface morphology. Such microstructural features have a profound impact on optical absorption, electrical conductivity, and electrochemical performance. For instance, smaller crystallite sizes and higher defect densities can enhance light absorption and provide additional active sites for surface reactions, while excessive disorder may increase charge recombination. Therefore, understanding the relationship between CBD growth parameters, microstructure, and functional properties is essential for optimizing CoSe thin films for specific applications.

From a fundamental perspective, the CBD process involves a complex interplay of homogeneous and heterogeneous reactions in the solution and at the substrate surface. The controlled supersaturation of ions leads to nucleation on the substrate, followed by film growth through ion-by-ion or cluster-by-cluster mechanisms. Factors such as bath stability, ion release kinetics, and substrate surface chemistry significantly influence film uniformity and phase purity. Despite its simplicity, CBD offers remarkable flexibility in tailoring thin film properties, making it an ideal technique for exploratory research as well as large-scale production. Given the growing demand for low-cost, efficient, and scalable materials for energy conversion and storage technologies, cobalt selenide thin films prepared by chemical bath deposition represent a promising research direction. However, challenges remain in achieving precise phase control, minimizing defect-induced recombination, and ensuring long-term stability under operational conditions. Addressing these challenges requires a comprehensive understanding of the synthesis–structure–property relationships governing CoSe thin films. In this context, the present chapter provides a detailed and systematic overview of cobalt selenide thin films synthesized by chemical bath deposition. Emphasis is placed on the fundamental aspects of CBD growth mechanisms, phase formation, and microstructural evolution. The structural, optical, electrical, and photoelectrochemical properties of CoSe thin films are critically analyzed, with particular attention to how deposition parameters influence performance. Finally, the potential of CBD-grown CoSe thin films for photoelectrochemical and energy-related applications is discussed, along with current challenges and future research perspectives.

2. Chemical Bath Deposition of Cobalt Selenide Thin Films

CBD is a solution-phase thin film fabrication technique that has been extensively employed for the growth of metal chalcogenide semiconductors. The method is based on the controlled chemical reaction between metal ions and chalcogenide ions in an aqueous medium, resulting in the gradual formation of a solid phase on a suitable substrate. CBD has gained prominence due to its simplicity, low processing temperature, minimal instrumentation requirements, and suitability for large-area deposition [7]. For cobalt selenide (CoSe) thin films, CBD is particularly advantageous because it allows precise control over deposition kinetics through bath chemistry, pH, temperature, and the use of complexing agents. These parameters govern nucleation density, grain growth, film thickness, and stoichiometry. Compared with physical deposition techniques such as sputtering or thermal evaporation, CBD offers a cost-effective and scalable route for synthesizing CoSe thin films for electrochemical and photoelectrochemical applications.

In CBD, uncontrolled mixing of cobalt and selenium precursors would lead to rapid homogeneous precipitation of CoSe in the solution bulk rather than on the substrate surface. To overcome this, complexing agents such as ammonia, triethanolamine, or ethylene diamine tetra acetic acid are introduced to bind free Co^{2+} ions and reduce their effective concentration in the bath [8]. This equilibrium ensures the slow release of Co^{2+} ions, thereby enabling controlled heterogeneous nucleation on the substrate. Selenium is commonly supplied in the form of sodium selenite (Na_2SeO_3) or sodium selenosulfate (Na_2SeSO_3). In alkaline conditions, these precursors undergo reduction reactions to release Se^{2-} ions gradually. Reducing agents such as hydrazine hydrate or sodium sulfite are often used to facilitate this conversion [9]. The controlled generation of Se^{2-} ions is critical to prevent secondary phase formation and to maintain stoichiometric CoSe growth. Once the ionic product of Co^{2+} and Se^{2-} exceeds the solubility product constant of CoSe, nucleation begins preferentially on the substrate surface. The film grows via an ion-by-ion mechanism, where successive adsorption of Co^{2+} and Se^{2-} ions leads to the formation of a continuous CoSe layer. Heterogeneous nucleation is favored due to lower interfacial energy between the substrate and the growing film compared to homogeneous nucleation in solution.

Substrate cleanliness plays a decisive role in film adhesion and uniformity. Glass, fluorine-doped tin oxide, or indium tin oxide substrates are typically used. The substrates are sequentially cleaned using acetone, ethanol, and deionized water in an ultrasonic bath to remove organic contaminants and surface particulates. Finally, the substrates are dried in air or nitrogen atmosphere prior to deposition [10]. A typical chemical bath for CoSe deposition is prepared by dissolving a cobalt salt (e.g., $CoCl_2 \cdot 6H_2O$ or $Co(NO_3)_2 \cdot 6H_2O$) in deionized water under constant stirring. A complexing agent is added slowly to stabilize the cobalt ions. Separately, the selenium precursor solution is prepared and added dropwise to the cobalt complex solution. The pH of the bath is adjusted to alkaline conditions (pH 9–11) using ammonium hydroxide or sodium hydroxide. Alkaline pH promotes selenium ion release and enhances film adhesion. The final solution is stirred continuously to ensure homogeneity. The cleaned substrates are vertically immersed in the reaction bath. Deposition is typically carried out at temperatures ranging from 60 to 90°C, which accelerates reaction kinetics without causing rapid precipitation. The deposition time varies between 30 and 120 minutes, depending on the desired film thickness. During deposition, the substrate surface gradually changes color (brownish to black), indicating CoSe film formation. After deposition, the films are removed, rinsed thoroughly with deionized water to eliminate loosely adhered particles, and dried at room temperature [11]. As-deposited CoSe films are often amorphous or poorly crystalline. Post-deposition annealing at temperatures between 200 and 400°C in inert (N_2 or Ar) or selenium atmosphere significantly improves crystallinity, grain growth, and phase purity. Annealing also reduces structural defects and enhances electrical conductivity. Increasing bath temperature enhances ionic mobility and reaction rates, leading to improved crystallinity and larger grain size. However, excessively high temperatures can result in bulk precipitation and non-uniform films. pH strongly influences the stability of cobalt complexes and the rate of selenium ion generation. Alkaline pH favors uniform film growth, whereas acidic conditions often lead to poor adhesion and incomplete coverage. Film thickness increases almost linearly with deposition time in the initial stages. Prolonged deposition, however, may cause stress accumulation, cracking, or peeling of the film. Higher complexing agent concentration slows down Co^{2+} release, resulting in finer grains and smoother films. Conversely, insufficient complexation leads to rapid precipitation and poor surface morphology. CBD offers several advantages, including low-cost synthesis, large-area deposition, and compatibility with flexible substrates. The technique is environmentally benign and suitable for industrial scaling. However, challenges such as stoichiometry control, secondary phase formation ($CoSe_2$, Co_9Se_8), and batch-to-batch reproducibility remain areas of active research.

3. Structural Properties of Cobalt Selenide Thin Films

The structural properties of cobalt selenide thin films play a decisive role in determining their optical, electrical, electrochemical, and photoelectrochemical behavior. Since CoSe belongs to the class of transition-metal chalcogenides, it exhibits structural flexibility, polymorphism, and defect tolerance, which are strongly influenced by synthesis parameters. In thin-film form, deviations from bulk crystallography frequently occur due to substrate interaction, limited diffusion length, growth kinetics, and post-deposition treatments.

Cobalt selenide exists in several crystallographic phases, including CoSe, $CoSe_2$, and Co_9Se_8, each possessing distinct atomic arrangements and physical properties. Among these, hexagonal CoSe is most frequently reported for chemically deposited thin films; particularly those prepared using low-temperature methods such as CBD. Hexagonal CoSe crystallizes in a NiAs-type structure with space group $P6_3/mmc$. In this arrangement, cobalt atoms occupy octahedral interstitial sites formed by a hexagonal close-packed selenium lattice. The layered nature of this structure leads to anisotropic growth behavior, favoring preferred orientation along specific crystallographic planes. This anisotropy is especially pronounced in thin films, where surface energy minimization governs nucleation and growth mechanisms. Under selenium rich or high temperature conditions, phase transformation to cubic $CoSe_2$, which adopts a pyrite-type structure ($Pa\overline{3}$), has been reported. Mixed-phase formation may also occur when the Co:Se ratio deviates from stoichiometry or when deposition parameters are not optimized. Such phase coexistence often results in broadened XRD peaks and reduced crystallinity, adversely affecting charge transport properties [12].

X-ray diffraction (XRD) is the most widely employed technique to investigate the structural properties of CoSe thin films. Typical XRD patterns of hexagonal CoSe exhibit diffraction peaks corresponding to crystallographic planes such as (100), (101), (102), (110), and (103), confirming polycrystalline nature. The absence of impurity peaks indicates successful formation of phase-pure CoSe. The relative intensity of diffraction peaks provides information about preferred orientation. Films grown at lower temperatures often display randomly oriented grains, whereas films deposited at optimized conditions show enhanced orientation along energetically favorable planes, such as (101) or (002). Peak broadening in XRD patterns reflects the nanocrystalline nature of CoSe thin films. The average crystallite size is estimated using the Scherrer equation. Reported crystallite sizes typically range from 10 to 50 nm, depending on deposition temperature, precursor concentration, and annealing conditions. Nanocrystallinity enhances surface-to-volume ratio, which is advantageous for catalytic and electrochemical applications, though it may introduce higher defect density. Lattice constants of hexagonal CoSe thin films are calculated using Bragg's law and hexagonal lattice equations. Typical values reported for thin films are a = 3.60–3.63 Å and c = 5.25–5.30 Å. Slight deviations from bulk lattice parameters are commonly observed and attributed to lattice strain, selenium vacancies, cobalt interstitials, and substrate-induced stress. Changes in unit cell volume serve as an indicator of non-stoichiometry and defect incorporation within the lattice. Selenium deficiency generally leads to lattice contraction, whereas cobalt excess may cause lattice expansion. Microstrain represents lattice distortions arising from defects, grain boundaries, and stress fields. Higher microstrain values are typically observed in films deposited at lower temperatures due to insufficient atomic mobility. Annealing promotes strain relaxation by enabling defect rearrangement and grain growth. Dislocation density provides insight into the defect concentration within the film. CoSe thin films generally exhibit dislocation densities in the range of 10^{15}–10^{16} m^{-2}, characteristic of nanocrystalline semiconductors. Although high dislocation density can impede carrier mobility, it also introduces active sites beneficial for electrocatalytic reactions. Structural properties are closely linked to surface morphology. SEM and AFM studies reveal that CoSe thin films typically exhibit granular, platelet-like, or flake-shaped morphologies. Grain size increases with deposition temperature and annealing, resulting in improved crystallinity and reduced grain boundary scattering. Rough surfaces with interconnected grains are advantageous for photoelectrochemical and catalytic applications due to enhanced electrolyte contact and increased active surface area. Annealing plays a critical role in tailoring the structural quality of CoSe thin films. Moderate annealing temperatures improve crystallinity, reduce microstrain, and increase grain size. Annealing in selenium atmosphere prevents selenium loss and maintains stoichiometry. However, excessive annealing may induce phase transformation or film degradation. The structural quality of CoSe thin films directly influences their functional properties. Films with larger crystallite size and lower defect density exhibit improved electrical conductivity, while defect-rich nanocrystalline films provide enhanced electrochemical activity. Thus, structural optimization is essential for application-specific performance.

4. Optical Properties of Cobalt Selenide Thin Films

The optical properties of cobalt selenide thin films are of fundamental and technological importance due to their direct relevance to applications in photoelectrochemical cells, photodetectors, electrocatalysis, and energy conversion devices**.** As a transition-metal chalcogenide, CoSe exhibits strong light matter interaction arising from its partially filled *d* orbitals and chalcogen *p* states. When prepared in thin film form, its optical response is strongly influenced by crystallinity, stoichiometry, film thickness, grain size, and defect density. Optical characterization provides critical insight into the electronic band structure, optical band gap, absorption behavior, and defect states in CoSe thin films. The optical absorption behavior of CoSe thin films is typically investigated using UV Visible spectroscopy in the wavelength range of 300–1100 nm. CoSe thin films generally exhibit strong absorption in the visible and near-infrared regions**,** making them suitable for solar energy harvesting applications. The high absorption coefficient values (10^4–10^5 cm^{-1}) observed for CoSe thin films indicate efficient photon absorption, even for relatively thin layers. The absorption edge often shows a gradual slope rather than a sharp cutoff, reflecting the presence of localized states near the band edges due to structural disorder and defects. Films with improved crystallinity generally display a steeper absorption edge, indicating reduced defect concentration. Film thickness plays a crucial role in determining the absorption behavior. Thicker CoSe films show increased absorption intensity due to longer optical path length, while thinner films exhibit higher transmittance. Additionally, improved crystallinity obtained through higher deposition temperature or post-annealing leads to enhanced absorption in the visible region due to reduced scattering losses and improved electronic ordering. The optical band gap (Eg) of CoSe thin films is commonly estimated using the Tauc relation. Experimental studies often report CoSe thin films to exhibit direct allowed transitions though indirect transitions have also been suggested under certain conditions. By extrapolating the linear portion of the $(\alpha h\nu)^2$ versus $h\nu$ plot to the energy axis, the optical band gap is obtained. Reported band gap values for CoSe thin films typically lie in the range of 1.4–2.1eV**,** depending on synthesis method and microstructural properties. The observed variation in band gap values is primarily attributed to quantum confinement effects, lattice strain, and defect states**.** Nanocrystalline CoSe thin films often show a slight blue shift in band gap compared to bulk material due to reduced crystallite size. Conversely, films with high defect density may exhibit band tailing, leading to an apparent reduction in band gap. Post-deposition annealing generally results in a decrease in band gap energy due to grain growth and improved crystallinity, which reduces quantum confinement effects and defect density. The refractive index of CoSe thin films provides information about light propagation and electronic polarizability. It is typically calculated from reflectance or transmittance data using standard optical models. CoSe thin films exhibit refractive index values in the range of 2.0–3.0 in the visible region. An increase in refractive index with photon energy is commonly observed, reflecting normal dispersion behavior. Films with higher density and improved crystallinity tend to show higher refractive index values due to increased electronic polarizability. CoSe thin films exhibit relatively high extinction coefficient values in the visible region, consistent with their strong absorption behavior. A decrease in high extinction coefficient at higher wavelengths indicates reduced absorption in the near-infrared region. The degree of structural disorder in CoSe thin films is quantified using Urbach energy, which describes the exponential absorption tail near the band edge. Lower Urbach energy values indicate reduced disorder and fewer localized states. Annealed CoSe thin films typically exhibit lower Urbach energy compared to as-deposited films, confirming improved structural ordering and reduced defect density.

5. Electrical Properties of Cobalt Selenide Thin Films

As a transition metal chalcogenide, CoSe exhibits mixed ionic electronic conduction and tunable electrical behavior arising from its variable stoichiometry, defect chemistry, and crystallographic structure. In thin film form, these properties are further modified by grain boundaries, surface states, and substrate film interactions. Electrical characterization of CoSe thin films provides insight into charge transport mechanisms, carrier concentration, mobility, resistivity, and activation energy, all of which are strongly dependent on synthesis technique, deposition parameters, and post-treatment conditions.

Experimental studies consistently report that CoSe thin films exhibit n-type semiconducting behavior, which is primarily attributed to selenium vacancies and cobalt interstitials acting as donor type defects. These intrinsic defects introduce localized energy levels near the conduction band, facilitating electron transport. The layered crystal structure of hexagonal CoSe promotes anisotropic charge transport, with higher conductivity along planes containing closely packed cobalt atoms. In nanocrystalline films, however, charge transport is significantly influenced by grain boundaries and defect states. The temperature dependence of electrical conductivity is a key indicator of conduction mechanism. CoSe thin films generally show thermally activated conductivity, with conductivity increasing as temperature rises, confirming semiconducting behavior. Activation energy values for CoSe thin films typically lie in the range of 0.1–0.4eV, depending on crystallinity and defect concentration. Lower activation energy indicates easier carrier excitation and improved electrical conductivity. The room-temperature electrical conductivity of CoSe thin films varies widely, typically ranging from 10^{-3} to 10^{2} S cm^{-1}**.** This broad range reflects differences in deposition technique, film thickness, stoichiometry, post-annealing treatment. Nanocrystalline films deposited at lower temperatures often exhibit higher resistivity due to increased grain boundary scattering. In contrast, annealed films with larger grains show enhanced conductivity due to reduced defect density and improved carrier mobility. Film thickness significantly influences electrical properties. Thin CoSe films generally exhibit higher resistivity due to discontinuous grain networks and increased surface scattering. As thickness increases, grain coalescence improves, leading to reduced resistivity and enhanced conductivity. Beyond an optimal thickness, however, conductivity may saturate due to bulk like behavior.

In polycrystalline CoSe thin films, grain boundaries act as potential barriers to carrier transport. According to the grain boundary model, charge carriers require thermal energy to overcome these barriers, resulting in temperature-dependent conductivity. Improved crystallinity and grain growth reduce barrier height and enhance conductivity. At lower temperatures, charge transport in CoSe thin films is often dominated by hopping conduction between localized states near the Fermi level. This behavior is especially prominent in defect-rich and nanocrystalline films. Variable-range hopping models have been successfully applied to explain low temperature conductivity in cobalt chalcogenide systems.

Hall effect measurements provide direct information about carrier type, carrier concentration, and mobility**.** CoSe thin films generally exhibit negative Hall coefficient (confirming n-type conductivity), carrier concentration in the range of 10^{18}–10^{21} cm^{-3}, mobility values typically between 0.1 and 10 cm^{2} V^{-1} s^{-1}. Low mobility values are primarily attributed to scattering from grain boundaries, defects, and lattice disorder. Annealing treatments significantly improve mobility by reducing defect density and enhancing grain connectivity. Post-deposition annealing plays a crucial role in tuning the electrical properties of CoSe thin films. Annealing under inert or selenium atmosphere increases electrical conductivity, reduces resistivity, lowers activation energy. These improvements arise from enhanced crystallinity, reduced microstrain, and improved stoichiometric balance. However, excessive annealing can lead to selenium loss, resulting in increased defect density and degraded electrical performance. Stoichiometry also plays a critical role; selenium-deficient films often show higher conductivity due to increased donor concentration, while selenium-rich films may exhibit higher resistivity. The electrical stability of CoSe thin films under prolonged bias and temperature cycling is essential for practical device applications. Studies indicate that CoSe thin films exhibit good electrical stability due to strong Co–Se bonding and chemical robustness. These characteristics make them suitable for electrodes in supercapacitors, electrocatalysts for hydrogen evolution reaction, photoelectrodes in PEC cells.

References:

1. Gaur, M., Hankare, P., Garadkar, K., Mulla, I., & Bhuse, V. (2014). Studies on the synthesis and characterization of CoSe thin films by a chemical bath deposition method. *New Journal of Chemistry, 38*(1), 255–261.
2. Chikwenze, R. A., Uba, J. I., & Arikpo, J. U. (2017). Optical and electrical properties of chemical bath deposited cobalt selenide thin film. *Chalcogenide Letters, 14*(6), 211–216.
3. Ho, S. M., & Anand, T. J. S. (2021). The influence of bath temperature on the properties of SILAR deposited cobalt selenide thin films. *Engineering, Technology & Applied Science Research, 11*(4), 7393–7398.
4. Egwunyenga, N. I., Onuabuchi, V. O., Okoli, N. L., & Nwankwo, I. E. (2021). Effect of deposition time on the optical and structural properties of SILAR-deposited cobalt selenide thin films. *International Research Journal of Multidisciplinary Technovation, 3*(2), 1–8.
5. Ghobadi, N., Habibi, M., Rezaee, S., Ţălu, S., Shakoury, R., Arman, A., Luna, C., Amiri, F., Naderi, S., Hafezi, F., & Mardani, M. (2019). The surface topographical properties of CoSe thin films. *Semiconductors, 53*(13), 1751–1756.
6. Mohamed, A. M., & Allam, N. K. (2022). High-performance flexible supercapacitors based on nanostructured cobalt selenide: A review. *Journal of Energy Storage, 47*, 103565.
7. Pathan, H. M., & Lokhande, C. D. (2004). Deposition of metal chalcogenide thin films by successive ionic layer adsorption and reaction (SILAR) method. *Bulletin of Materials Science, 27*(2), 85–111.
8. Mane, R. S., & Lokhande, C. D. (2000). Chemical deposition method for metal chalcogenide thin films. *Materials Chemistry and Physics, 65*(1), 1–31.
9. Chate, P. A., Sathe, D. J., & Bhabad, V. S. (2024). Studies on the synthesis and characterization of chemically deposited cadmium selenide thin films. *Materials Science in Semiconductor Processing, 35*, 1305–1310.
10. Chate, P. A., Sathe, D. J., & Hake, S. L. (2022). Influence of pH on properties of chemically deposited cobalt selenide thin films. *Applied Physics A, 128*(10), 904.
11. Sartale, S. D., & Lokhande, C. D. (2001). Growth and characterization of copper selenide thin films by SILAR method. *Materials Chemistry and Physics, 72*(1), 101–104.
12. Zhang, Y., Liu, H., & Chen, X. (2017). Successive ionic layer adsorption and reaction (SILAR) deposition of cobalt selenide as an efficient counter electrode for dye-sensitized solar cells. *Applied Surface Science, 396*, 1430–1436.

AGRIPREDICT AI: HARNESSING ARTIFICIAL INTELLIGENCE FOR SUSTAINABLE AND INTELLIGENT AGRICULTURE

R. K. Kavitha[1] and Nakshathra DV[2]

[1]Department of Computer Applications, KCT

[2]School of Data Science, KCLAS

Corresponding author E-mail: kavitha.rk.mca@kct.ac.in, nakshathra.23bds@kclas.ac.in

Abstract:

Agricultural price volatility, climate uncertainty, and disease outbreaks significantly affect farmers' income and national food security in India. Accurate forecasting of agri-commodity prices combined with region-specific advisories can empower farmers, traders, and policymakers to make informed decisions. This paper presents AgriPredict AI, an end-to-end AI-powered agricultural analytics system that predicts crop prices for 5–90 days using live market and weather data, provides disease outbreak warnings, and delivers crop, fertilizer, and irrigation recommendations tailored to Indian states and regions. The system integrates real-time mandi price data from Agmarknet and meteorological data from OpenWeather API, leveraging time-series forecasting models such as ARIMA, Prophet, and LSTM. A comprehensive utilization dashboard analyzes demand percentage, market volume, farmer adoption rate, and seasonal demand. Experimental results show that hybrid deep learning models outperform traditional statistical approaches in medium- and long-term forecasting. The platform demonstrates strong potential to support sustainable agriculture, improve farmer income stability, and assist government policy planning.

Keywords: Agri-Commodity Price Prediction, Artificial Intelligence in Agriculture, Time Series Forecasting, LSTM, ARIMA, Decision Support System, Disease Warning System, Smart Farming, Indian Agriculture

Introduction:

Agriculture remains the backbone of the Indian economy, employing more than 50% of the workforce and contributing significantly to national GDP. However, farmers face persistent challenges such as unpredictable market prices, climate variability, crop diseases, and inefficient resource utilization. Traditional decision-making methods based on experience or delayed government reports often fail to address these dynamic challenges.

Recent advances in Artificial Intelligence (AI) and Data Science offer promising solutions to forecast commodity prices, detect disease risks, and optimize farming decisions. Accurate price prediction allows farmers to decide when to sell, what to grow, and how much to invest, thereby reducing losses due to sudden market fluctuations. This research proposes AgriPredict AI, a unified AI-driven platform designed for Indian agriculture. The system combines price forecasting, disease warnings, crop recommendations, and utilization analytics into a single web-based dashboard accessible to farmers, traders, and policymakers.

Literature Review

The use of artificial intelligence and machine learning in agriculture has gained significant attention in recent years, with numerous studies focusing on improving productivity, forecasting commodity prices, and enabling decision support for farmers and policymakers. Traditional statistical approaches such as ARIMA have long been applied to agricultural price prediction, providing a foundational methodology for time-series forecasting. Box and Jenkins (1976) introduced the ARIMA model, which remains widely cited for analyzing stationary time-series data. While effective for linear patterns, ARIMA often struggles to capture the complex, non-linear fluctuations seen in agricultural markets. To address seasonality and scalability, Taylor and Letham (2018) developed the Facebook Prophet model, which allows decomposition of time-series into trend, seasonality, and holiday effects, offering more flexible forecasting for agricultural commodities. More recently, Sagheer and Kotb (2019) demonstrated that Long Short-Term Memory (LSTM) neural networks outperform traditional approaches such as ARIMA and Prophet in predicting agricultural commodity prices due to their ability to capture long-term dependencies and non-linear temporal patterns.

Beyond price forecasting, machine learning has been extensively applied to broader smart agriculture tasks. Liakos *et al.* (2018) reviewed the adoption of AI in agriculture, highlighting applications such as crop yield estimation, disease detection, and automated decision support. Similarly, Kamilaris and Prenafeta-Boldú (2018) emphasized the value of deep learning models in integrating multi-source data, including climate, soil, and market variables, to enhance predictive performance. These studies indicate that combining structured data with advanced learning algorithms significantly improves the accuracy and reliability of agricultural predictions. Complementing this, Panda *et al.* (2020) investigated climate-based disease forecasting systems, demonstrating that early warning mechanisms can substantially reduce crop losses by enabling timely preventive interventions. The Food and Agriculture Organization (FAO, 2021) further advocated for region-specific digital agriculture systems to enhance food security and sustainability, emphasizing the importance of integrating real-time monitoring and predictive analytics.

Several initiatives have focused on AI-powered decision support systems for Indian agriculture. The Government of India (2022) has implemented Agmarknet, a platform providing real-time market data for farmers and traders, which serves as a foundation for integrating predictive models into decision-making processes. Zhang *et al.* (2020) explored machine learning and deep learning techniques for crop price prediction, illustrating the advantages of combining multiple models for higher accuracy. In addition, Mishra and Singh (2021) highlighted the potential of AI-based decision support systems to assist farmers, traders, and policymakers in making informed choices, thereby optimizing supply chains and reducing post-harvest losses. Collectively, these studies reveal a gap in comprehensive, integrated platforms that simultaneously address price prediction, disease management, and real-time decision support. AgriPredict AI seeks to fill this gap by leveraging advanced forecasting models, climate-aware disease warnings, and intuitive dashboards to provide actionable insights to stakeholders across the agricultural ecosystem.

System Architecture

AgriPredict AI follows a modular, scalable architecture:

i. Data Layer

- Agmarknet (daily mandi prices)
- OpenWeather API (temperature, rainfall, humidity)
- Government & FAO datasets

ii. Processing Layer

- Data cleaning and normalization
- Feature engineering (seasonality, weather impact)

iii. AI/ML Layer

- ARIMA
- Prophet
- LSTM

iv. Application Layer

- Price prediction engine
- Disease warning system
- Crop recommendation module
- Utilization analytics dashboard

v. Presentation Layer

- Web dashboard (AgriPredict AI)
- State and region-based filtering
- Alerts and recommendations

Methodology

1. Data Collection

- **Primary Source**: Agmarknet (Government of India)
- **Secondary Sources**: FAO, FCI, Kaggle
- **Weather Data**: OpenWeather API

- **Geographical Coverage**: 28 states and 8 union territories of India

2. Dataset Description

Attribute	Description
Crop Name	Onion, Potato, Wheat, Rice, Pulses
State/Region	Indian states and UTs
Date	Daily records
Min/Max/Modal Price	Market prices
Weather Variables	Temperature, rainfall, humidity
Seasonal Index	Kharif/Rabi/Zaid

3. Model Building and Analysis

- **ARIMA**: Baseline statistical forecasting
- **Prophet**: Trend + seasonality modeling
- **LSTM**: Deep learning for long-term dependencies
- **Model Evaluation Metrics**
  - Mean Absolute Error (MAE)
  - Root Mean Square Error (RMSE)
  - Mean Absolute Percentage Error (MAPE)

4. Implementation and Deployment

The AgriPredict AI system is implemented as a cloud-hosted, full-stack web application with a modern and scalable architecture. The frontend is developed using React.js for building dynamic and responsive user interfaces, Tailwind CSS for efficient and consistent styling, and Framer Motion for smooth animations and enhanced user experience. The backend is powered by Node.js, which provides RESTful API services for handling data requests, model outputs, user feedback, and system alerts. Data storage and management are handled through CMS-based structured collections, enabling efficient organization of crop prices, disease warnings, regional data, and utilization metrics. The entire application is deployed on a cloud infrastructure to ensure high availability, scalability, and real-time data access. Additionally, the platform is designed to be mobile-friendly and supports multilingual access, allowing farmers, traders, and policymakers across different regions of India to easily access predictions, recommendations, and alerts using smartphones or desktop devices.

Results and Discussion:

The experimental evaluation of the proposed system reveals significant insights into both price prediction accuracy and disease warning effectiveness. Among the forecasting models used, the ARIMA model recorded comparatively high Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) values, particularly as the prediction horizon increased, due to its assumption of linearity and requirement for stationary time-series data, which limits its ability to adapt to volatile agricultural markets. The Prophet model demonstrated moderate performance, benefiting from its built-in mechanisms for handling seasonality, trend changes, and missing data, making it more reliable than ARIMA for short- to medium-term forecasts. However, Prophet still exhibited reduced accuracy when predicting prices beyond a 30-day window. The LSTM model, on the other hand, consistently achieved the lowest MAE and RMSE values across all forecasting horizons, especially for extended predictions ranging from 30 to 90 days. This superior performance can be attributed to LSTM's deep learning architecture, which effectively captures complex non-linear patterns, long-term temporal dependencies, and interactions between historical price trends and external factors such as weather conditions. These results confirm that deep learning-based approaches are more suitable for agricultural price forecasting in dynamic and uncertain market environments. In addition to price prediction, the disease warning module demonstrated strong effectiveness in identifying high-risk periods for crop disease outbreaks by analyzing climatic variables including temperature, humidity, and rainfall patterns. By correlating these environmental factors with historical disease occurrence data, the system generates early warnings that enable farmers to adopt preventive measures such as timely pesticide application, crop rotation, or irrigation adjustments. This proactive approach significantly reduces the likelihood of large-scale disease spread, minimizes crop damage, and supports sustainable farming practices, highlighting the practical value of integrating predictive analytics with agricultural decision-support systems.

Predictive model analysis

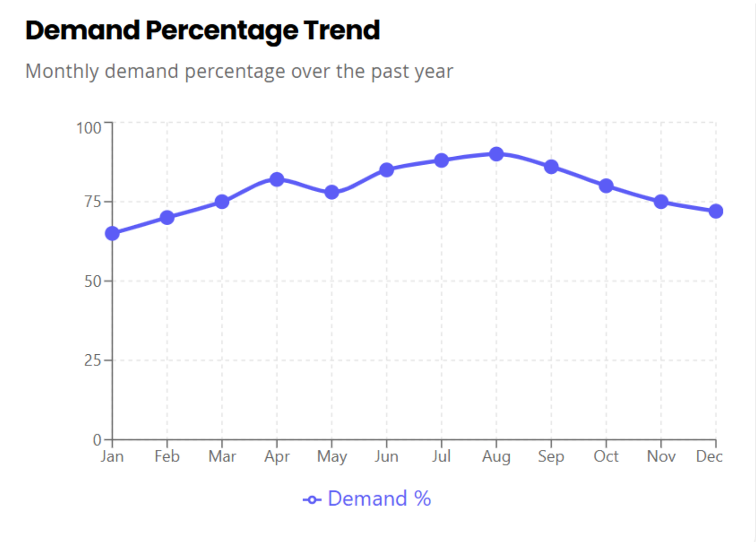


Figure 1: Demand percentage trend (Monthly)

Figure 1 shows the demand percentage trend graph illustrates the variation in crop demand over a one-year period on a monthly basis. The graph shows a gradual increase in demand from January to April, indicating rising market requirements during the early cultivation and pre-harvest stages. A slight dip is observed in May, followed by a steady rise reaching a peak around

July and August, which corresponds to the peak agricultural and consumption season. After August, demand gradually declines toward the end of the year, reflecting post-harvest saturation and reduced market activity. This trend highlights the strong influence of seasonal agricultural cycles on market demand and emphasizes the importance of aligning production and sales strategies with demand fluctuations.

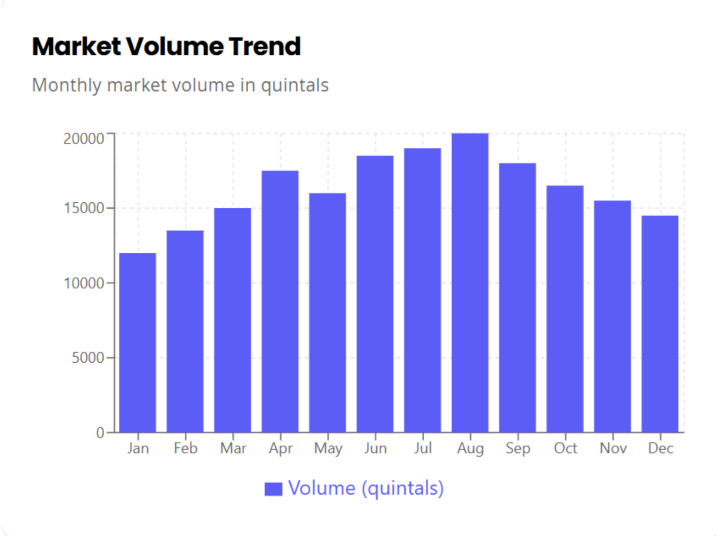


Figure 2: Market volume trend (Monthly market volume in quintals)

Figure 2 shows that the market volume trend graph represents the monthly quantity of agricultural produce traded in the market, measured in quintals. The graph shows a steady increase in market volume from January to April, indicating higher arrivals of produce during harvesting periods. A temporary decline in May suggests reduced supply, possibly due to transitional cropping phases or climatic factors. Market volume peaks during July and August, aligning with major harvest seasons, and then gradually decreases from September to December as supply stabilizes and demand tapers off. This visualization demonstrates the close relationship between production cycles, harvesting seasons, and market supply levels.

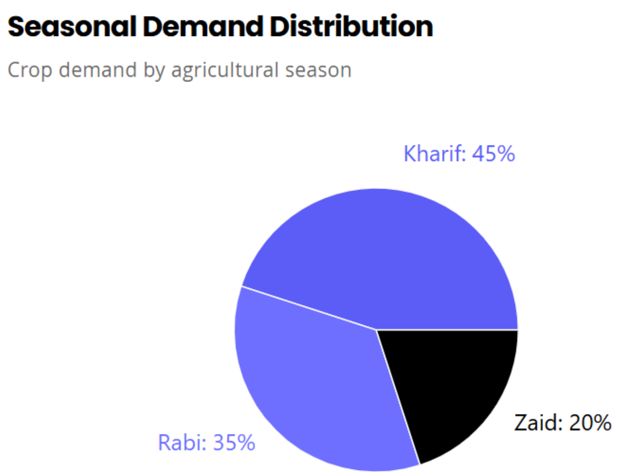


Figure 3: Farmer adoption rate (Monthly adoption percentage)

Figure 3 shows the farmer adoption rate graph that depicts the percentage of farmers adopting recommended crops, technologies, or advisory services over the year. The adoption rate shows a consistent upward trend from January to April, reflecting increasing awareness and confidence among farmers. A minor decline in May is followed by a steady rise, with the highest adoption observed around August. This peak coincides with favorable climatic conditions and active agricultural seasons. After August, adoption gradually declines toward December, possibly due to reduced farming activities and off-season periods. Overall, the graph indicates positive farmer engagement, with seasonal factors significantly influencing adoption behavior.

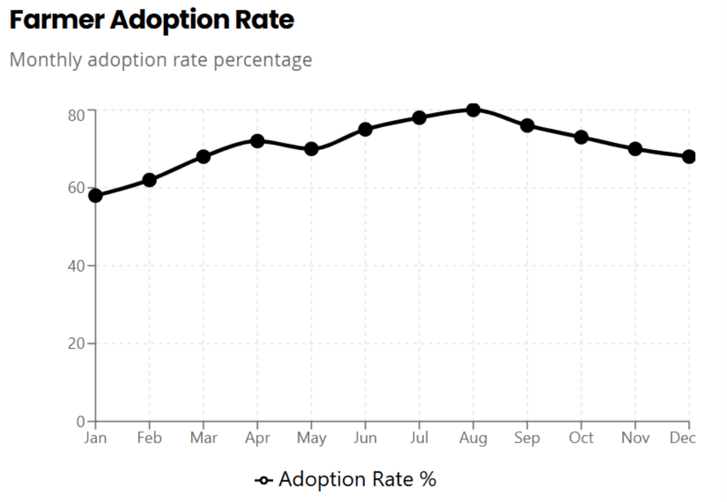


Figure 4: Regional market analysis (Market volume and adoption rate by region)

Figure 4 provides the regional market analysis graph compares market volume and farmer adoption rates across major agricultural regions such as Punjab, Haryana, Uttar Pradesh, Maharashtra, Karnataka, and Tamil Nadu. Uttar Pradesh and Punjab show the highest market volumes, indicating strong agricultural output and active market participation. Haryana and Maharashtra demonstrate moderate volumes, while Karnataka and Tamil Nadu show comparatively lower volumes. Variations in adoption rates across regions highlight differences in technological penetration, resource availability, and awareness levels. This analysis helps identify high-performing regions as well as areas requiring policy support, training programs, or infrastructure development.

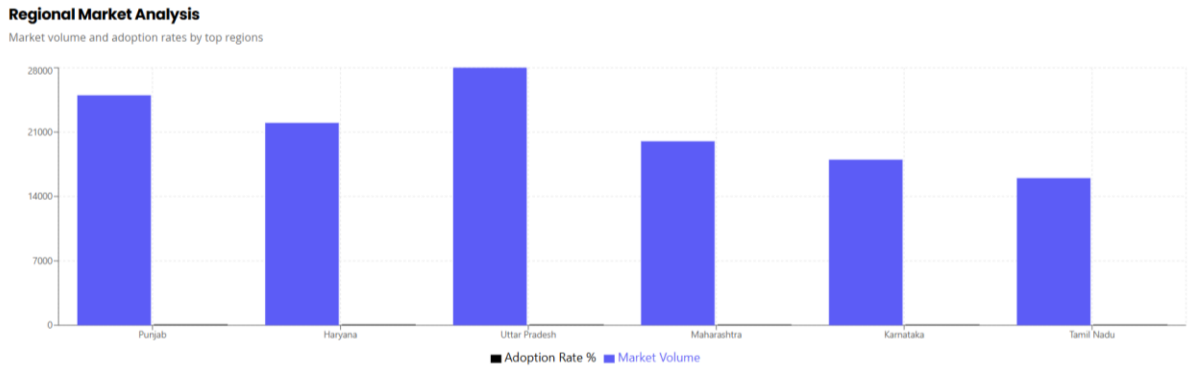


Figure 5: Seasonal demand distribution (Kharif, Rabi, Zaid)

Figure 5 visualizes the seasonal demand distribution pie chart illustrates the proportion of crop demand across the three major agricultural seasons in India. The Kharif season accounts for the highest demand at 45%, reflecting its dominance due to favorable monsoon conditions and large-scale crop production. The Rabi season contributes 35% of the total demand, supported by irrigation facilities and winter cropping patterns. The Zaid season represents 20% of demand, indicating relatively lower but significant market activity during the summer months. This distribution emphasizes the seasonal dependency of agricultural demand and underlines the importance of season-specific planning and forecasting.

Applications and implications

The proposed AgriPredict AI system offers wide-ranging real-world applications across multiple stakeholders in the agricultural ecosystem. For farmers, the platform enables informed selling and planting decisions by providing accurate price forecasts, seasonal demand trends, and disease warnings, thereby reducing uncertainty and financial risk. Government agencies can utilize the system for effective policy planning, monitoring market behavior, and regulating Minimum Support Price (MSP) mechanisms based on predictive insights and regional demand–supply analysis. Traders and market intermediaries benefit from improved supply chain optimization through advance knowledge of price movements, market volumes, and regional production patterns, which supports efficient storage, transportation, and distribution planning. Additionally, academic and research institutions can leverage the system's data-driven framework and analytics for agricultural research, experimentation, and educational purposes, promoting innovation and evidence-based learning in the agricultural domain.

Looking ahead, the system presents significant opportunities for enhancement and expansion through advanced technologies. The integration of satellite imagery and remote sensing data can improve crop health monitoring, yield estimation, and early detection of stress conditions. Reinforcement learning techniques can be employed to develop adaptive pricing models that dynamically respond to changing market conditions. The adoption of blockchain technology can enhance transparency, traceability, and trust in agricultural market transactions. Furthermore, voice-based advisory systems can improve accessibility for farmers with limited literacy or digital skills, enabling real-time interaction in regional languages. The inclusion of AI-powered insurance risk estimation models can assist insurers and policymakers in assessing crop risk more accurately, leading to fairer premiums and faster claim settlements, thereby strengthening financial security for farmers.

Summary and Conclusion:

AgriPredict AI demonstrates the transformative potential of artificial intelligence in modernizing Indian agriculture by addressing some of its most critical challenges, including price volatility, climate uncertainty, and crop disease risks. By integrating real-time market data, weather information, and historical agricultural datasets with advanced forecasting models such as

ARIMA, Prophet, and deep learning–based LSTM networks, the system delivers accurate and timely price predictions along with early disease warnings and actionable decision support. The inclusion of a user-friendly, interactive dashboard ensures that complex analytical outputs are presented in an accessible and practical manner, enabling farmers, traders, and policymakers to easily interpret trends, risks, and recommendations. This integration effectively bridges the gap between cutting-edge technology and grassroots farming needs, making data-driven agriculture more inclusive and impactful.

The experimental results and analytical evaluations validate the superior performance of deep learning approaches in capturing non-linear patterns, seasonal variations, and long-term dependencies inherent in agricultural data. Beyond technical accuracy, AgriPredict AI offers substantial socio-economic benefits by empowering farmers to make informed planting and selling decisions, thereby improving income stability and reducing losses caused by sudden market fluctuations or disease outbreaks. At a broader level, the system supports sustainable agricultural practices by promoting efficient resource utilization, proactive risk management, and informed policy planning. Overall, AgriPredict AI serves as a scalable and adaptable framework that demonstrates how intelligent decision-support systems can contribute to resilient, technology-driven, and sustainable agricultural development in India.

References:

1. Box, G. E. P., and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.
2. Taylor, S. J., and Letham, B. (2018). *Forecasting at Scale*. The American Statistician.
3. Sagheer, A., and Kotb, M. (2019). *Time Series Forecasting of Agricultural Commodity Prices Using LSTM Networks*. Neurocomputing, Elsevier.
4. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochtis, D. (2018). *Machine Learning in Agriculture: A Review*. Sensors, MDPI.
5. Kamilaris, A., and Prenafeta-Boldú, F. X. (2018). *Deep Learning in Agriculture: A Survey*. Computers and Electronics in Agriculture, Elsevier.
6. Panda, S. S., Hoogenboom, G., and Paz, J. O. (2020). *Climate-Based Crop Disease Forecasting and Early Warning Systems*. Agricultural Systems, Elsevier.
7. Food and Agriculture Organization (FAO). (2021). *Digital Agriculture and Food Security*. FAO Publications.
8. Government of India. (2022). *Agmarknet: Agricultural Marketing Information System*. Ministry of Agriculture and Farmers Welfare.
9. Zhang, Y., Wang, S., and Ji, G. (2020). *Crop Price Prediction Using Machine Learning and Deep Learning Models*. IEEE Access.
10. Mishra, D., and Singh, R. (2021). *AI-Based Decision Support Systems for Smart and Sustainable Agriculture*. Springer.

STATISTICAL INVESTIGATION OF MILLETS MISSION: ACROSS INDIA AND ASSAM

Siva D and Shanmugam D.B
Department of Computer Science and Applications,
SRM Institute of Science and Technology, Ramapuram, Chennai, India
Corresponding author E-mail: sivad@srmist.edu.in, dbshanmugam@gmail.com

Abstract:

Assam Millet Mission aims to enhance millet production and consumption in Assam, addressing malnutrition and promoting sustainable agricultural practices. Here, in this study an attempt was made for systematic data driven inquiry of area, production and yield of millets among different states of India. The data was collected from the draft indicative concept note on Assam Millets Mission compiled in the year 2022 by Assam Rural Infrastructure and Agricultural Services (ARIAS) society, Department of Agriculture, Government of Assam. From the statistical investigation of area, production and yield of millets it was observed that the area, production and yield of millets haven't constantly increased or decreased during the period from 2018-2021. For states like Rajhasthan, area under cultivation of millets and production was highest. It was also come to noticed that since area under cultivation of millets for Assam was second least, therefore it has low production and yield of millets. Further, from pairwise comparison test it was found that there was significant difference among different states of India in accordance with area, production and yield of millets. The study also tried to outlined that in terms of yield of millets; state of Assam has a pairwise significant difference with all the different states of India.

Keywords: Millets; States of India; Area; Assam; Production; Yield.

1. Introduction:

Assam Millet Mission aims to enhance millet production and consumption in Assam, addressing malnutrition and promoting sustainable agricultural practices. This initiative aligns with the global recognition of millets as climate-resilient crops, particularly in light of the United Nations declaring 2023 as the 'International Year of Millets'. The mission focuses on integrating millets into traditional diets, improving food security, and supporting local farmers. It was launched in 2016, this initiative aims to boost millet productivity and consumption through various programs, including the "Millets for Health" campaign (Yadav *et al.,* 2023). The National Mission on Sustainable Agriculture and other policies promote research, market access, and minimum support prices for millets (Gupta *et al.,* 2024). Millets are rich in vitamins, minerals, and fiber, making them a healthy alternative to rice (Digra *et al.,* 2025). They have a low glycemic index, beneficial for individuals with diabetes (Yadav *et al.,* 2023). Studies show that millet-based products have higher acceptability and nutritional profiles compared to traditional rice-based foods (Digra *et al.,* 2025). Millets require less water and are more resilient to climate change, making them cost-effective for small farmers (Barman *et al.,* 2025). By diversifying cropping systems, millets can provide additional income sources for farmers, enhancing their economic stability. The average yield in Assam is approximately 5.70 q/ha, with a return income of Rs. 35,570 per hectare (Barman *et al.,* 2025). However, challenges such as lack of processing facilities and market access hinder growth (Barman *et al.,* 2025). Despite their benefits, small millets face declining cultivation areas and unfavorable trade terms compared to major crops like rice (Borbaruah *et al.,* 2025). Farmers show a strong preference for rice, leading to instability in millet production (Borbaruah *et al.,* 2025). The present study was undertaken to observe and study the area under cultivation (in hectares), yield (in kg/hectare) and production (in tonnes) of Millets cultivated in Assam and neighbouring states of India. Further, it had tried to bring a comparison among Assam and neighbouring states of India in accordance with these variables. Despite these efforts, challenges such as low productivity and limited market infrastructure persist, hindering the widespread adoption of millets in Assam. Addressing these barriers is crucial for the success of the Assam millet mission and the overall agricultural landscape in the region.

Research Methodology

An empirical research methodology was adopted to achieve the objectives of the study. For the analysis task, the study had accustomed to report based on government database. The analysis for the study was performed with descriptive and comparative statistical analysis method. The data collection procedure and analysis techniques used in this study were discussed below:

1. Data

The data collected for the research work was secondary and quantitative in nature. It was collected from the draft indicative concept note on Assam Millets Mission compiled in the year 2022 by Assam Rural Infrastructure and Agricultural Services (ARIAS) society, Department of Agriculture, Government of Assam. The data on area, yield and production of millet was collected from 2018 to 2021. It was collected for different states of India including state of Assam.

2. Methods

For analysis of the study, descriptive and comparative methods were considered. Further for graphical representation, histogram and frequency curves were developed. The analysis was executed in R version 4.5.2 software. A brief description of the statistical tools involved in this research work is discussed below:

2.1 Descriptive Statistics

Descriptive statistics serve as a foundational element in statistical analysis, providing essential tools for summarizing and interpreting data. This branch of statistics focuses on presenting data in a clear and concise manner through various formats such as tables, charts, and graphs. Descriptive statistics condense large datasets into understandable summaries, highlighting key features such as central tendency and variability (Blbas, 2024). Common metrics include the mean, median, and mode, which provide insights into the average or most common values within a dataset. It also includes range, variance, and standard deviation, which describe the spread of data points around the central value. Descriptive statistics are crucial in research for summarizing findings and presenting results in a comprehensible format, facilitating better decision-making.

2.2 One Way Analysis of Variance (ANOVA)

Analysis of variance (ANOVA) is a statistical method used to assess differences among group means and is particularly valuable in experimental research. It allows researchers to determine whether any of the differences among sample means are statistically significant, thus providing insights into the effects of one or more independent variables on a dependent variable. One-way ANOVA compares means across one independent variable with multiple levels. The variance in ANOVA quantifies the spread of data points and is crucial for distinguishing between systematic and unsystematic variance in experimental designs (Baban, 2024). ANOVA is widely used in biological sciences to evaluate treatment effects across various experimental designs, such as Completely Randomized Design and Randomized Block Design (Varalakshmi, 2025).

The one way ANOVA mathematical model is given as

$Y_{ij} = \mu + \tau_i + e_{ij}$……………….. (i)

Where,

Y_{ij} = Dependent variable

μ = General mean effect

τ_i = Between group variation

e_{ij} = Within group variation

2.3 Tukey's Honestly Significant Difference (HSD) Test

Tukey's Honestly Significant Difference (HSD) test is a statistical method used for making pairwise comparisons between group means while controlling for Type I error rates. It is particularly useful in scenarios involving multiple comparisons, as it provides a robust framework for determining significant differences among group means. The test is based on the studentized range distribution and is applicable to both equal and unequal sample sizes, making it versatile in various research contexts. The test has been employed in agricultural studies to compare the performance of different plant varieties under varying conditions (Cheshkova *et al.,* 2017).

The mathematical formula is given by

$$\text{Tukey's HSD} = q \times \sqrt{\frac{MSE}{n}} \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

Where,

MSE = Mean Square Error from ANOVA.

q = The critical value from the Studentized range distribution table based on certain level of significance.

n = sample size.

Analysis of Results

The analysis of the study was conducted with the help of useful statistical software such as Microsoft Excel and SPSS (Statistical Package for Social Sciences). The study had considered pre-determined value level of significance (α =0.05). The following results of the study is given below:

Table 1: State-wise Descriptive Statistics of Area (in hectares) under Cultivation of Millets from 2018-2021

States	Mean	Median	Variance	Standard deviation
Andhra Pradesh	183.50	214.50	77.31	5977.67
Assam	5.10	5.105	0.12	0.015
Bihar	11.65	10.99	3.17	10.07
Chhattisgarh	70.76	66.64	19.54	382.17
Gujarat	424.94	488.22	146.47	21454.81
Haryana	500.32	493.85	80.38	6461.41
Himachal Pradesh	7.12	7.18	0.23	0.052
Jharkhand	17.53	16.58	2.16	4.65
Karnataka	1546.60	1721.85	433.64	188044.92
Kerala	0.405	0.495	0.22	0.051
Madhya Pradesh	530.97	502	68.27	4661.16
Maharashtra	2241.21	2525.65	987.60	975354.26
Odisha	83.17	80.89	7.99	63.99
Rajasthan	4728.74	4820.35	274.31	75247.50
Tamil Nadu	513.96	560.24	142.97	20439.47
Telangana	75.60	82.5	40.44	1635.71
Uttar Pradesh	1060.75	1061.5	37.26	1388.25
Uttarakhand	143	143	6.38	40.66
West Bengal	7.89	8.92	2.26	5.12
Others	63.27	64.77	14.55	211.60

From table 1, it could be observed that the mean, median, variance and standard deviation for area under cultivation (in hectares) of millets from 2018-2021 varies among different states of India. It could be observed that Rajhasthan state has the highest mean and median area under cultivation of millets among other states of India. The least mean and median area under cultivation of millets was for Kerela. The table 1 also outlined that the mean and median area under cultivation of millets for state of Assam was second least in comparison to other states of India. Further from table 1, it was observed that Maharastra state has the highest variation in terms of area under cultivation of millets. The least variation were observed for Kerela, Himachal Pradesh and Assam.

Table 2: State-wise Descriptive Statistics of Production (in tonnes) under Cultivation of Millets from 2018-2021

States	Mean	Median	Variance	Standard deviation
Andhra Pradesh	351.11	351.11	38662.93	196.63
Assam	3.07	3.07	0.01	0.11
Bihar	10.87	10.87	6.95	2.64
Chhattisgarh	25.45	25.45	49.71	7.05
Gujarat	864.91	864.91	102893.69	320.77
Haryana	1052.32	1052.32	47712.08	218.43
Himachal Pradesh	7.25	7.25	1.12	1.06
Jharkhand	15.15	15.15	4.59	2.14
Karnataka	2051.34	2051.34	361997.94	601.66
Kerala	0.41	0.41	0.05	0.23
Madhya Pradesh	975.86	975.86	20230.46	142.23
Maharashtra	1731.14	1731.14	797231.63	892.88
Odisha	52.44	52.44	27.13	5.21
Rajasthan	4880.38	4880.38	162741.84	403.41
Tamil Nadu	844.04	844.04	45119.77	212.41
Telangana	108.88	108.88	6116.88	78.21
Uttar Pradesh	2179.50	2179.50	23137.37	152.11
Uttarakhand	198.75	198.75	342.72	18.51
West Bengal	8.41	8.41	5.17	2.27
Others	61.75	61.75	98.19	9.91

From table 2, it was observed that the mean, median, variance and standard deviation for production (in tonnes) of millets from 2018-2021 varies among different states of India. It could be observed that Rajhasthan state has the highest mean and median production of millets among other states of India. The least mean and median production of millets was for Kerela. The table 2 also tried to depict that the mean and median production of millets for state of Assam was second least in comparison to other states of India. Further from table 2, it was noticed that Maharastra state has the highest values of variance and standard deviation for production of millets. The least variation were observed for Kerela, Himachal Pradesh and Assam.

Table 3: State-wise Descriptive Statistics of Yield (in kg/hectare) of Millets from 2018-2021

States	Mean	Median	Variance	Standard deviation
Andhra Pradesh	1822.75	1767.50	300008.92	547.73
Assam	601.00	599.00	178.67	13.37
Bihar	938.25	932.00	1558.92	39.48
Chhattisgarh	361.00	343.50	2852.67	53.41
Gujarat	2012.25	2039.50	9826.92	99.13
Haryana	2097.50	2089.50	27459.00	165.71
Himachal Pradesh	1014.00	1036.00	13995.33	118.30
Jharkhand	864.25	859.00	4464.92	66.82
Karnataka	1341.25	1414.00	38400.92	195.96
Kerala	1020.50	995.00	4623.00	67.99
Madhya Pradesh	1835.75	1845.50	5829.58	76.35
Maharashtra	771.50	797.00	24747.67	157.31
Odisha	630.50	625.00	467.67	21.63
Rajasthan	1034.50	1045.50	10011.00	100.05
Tamil Nadu	1655.25	1628.50	3994.92	63.21
Telangana	1307.75	1252.00	126830.92	356.13
Uttar Pradesh	2055.25	2045.00	18502.92	136.03
Uttarakhand	1390.75	1425.00	15701.58	125.31
West Bengal	1086.50	1123.00	28625.67	169.19
Others	994.25	992.50	18750.25	136.93

From table 3, it could be noticed that the mean, median, variance and standard deviation for yield (in kg/hectare) of millets from 2018-2021 varies among different states of India. Among the states of India, Haryana was found highest in mean and median yield of millets. The least mean and median yield of millets was for Assam. Further from table 3, it was noticed that Andhra Pradesh state has the highest values of variance and standard deviation for production of millets. The least variation were observed for Assam and Odhisa.

Graphical representation of State-wise Area, Production and Yield of Millets from 2018-2021

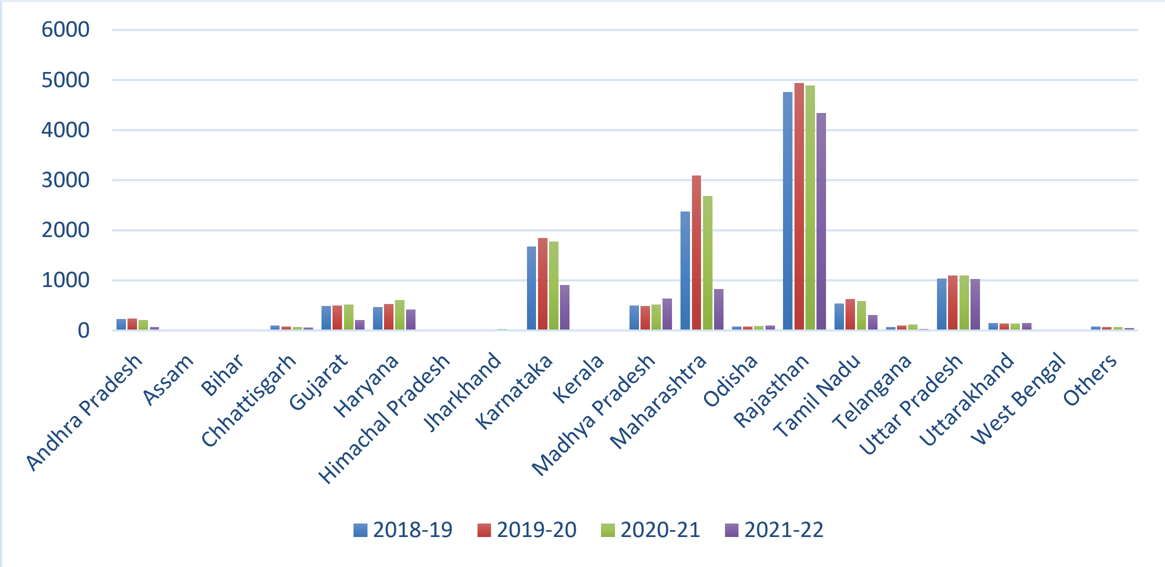


Figure 1: State-wise Area (in hectares) of millets from 2018-2021

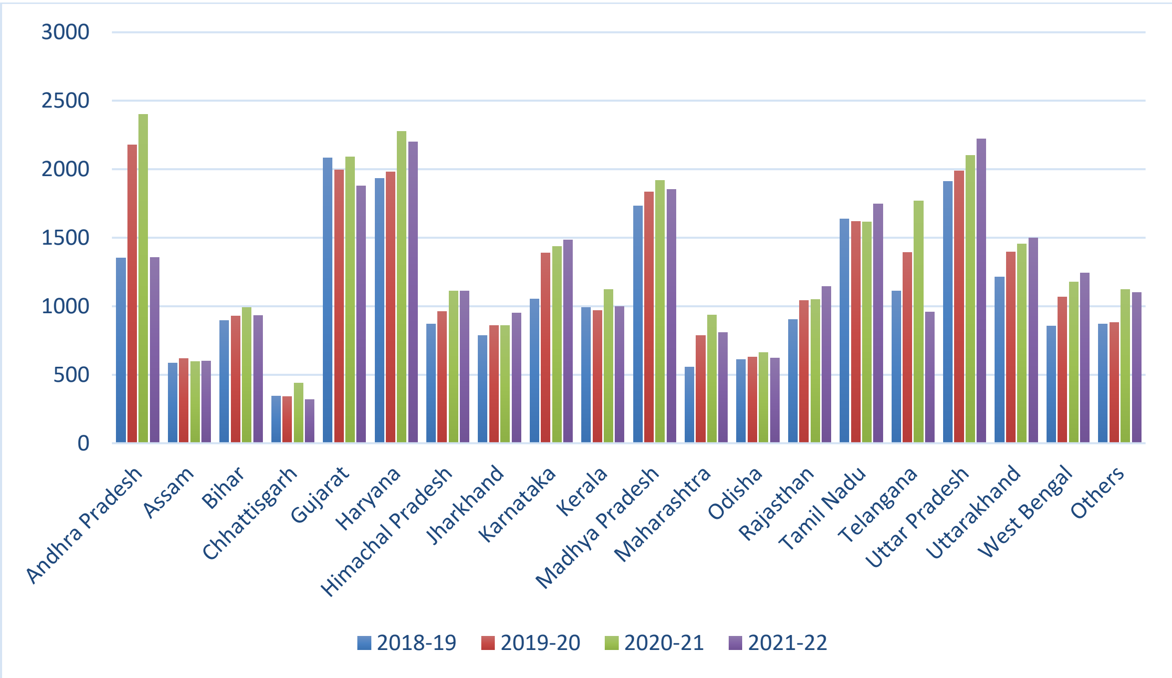


Figure 2: State-wise Production (in tonnes) of millets from 2018-2021

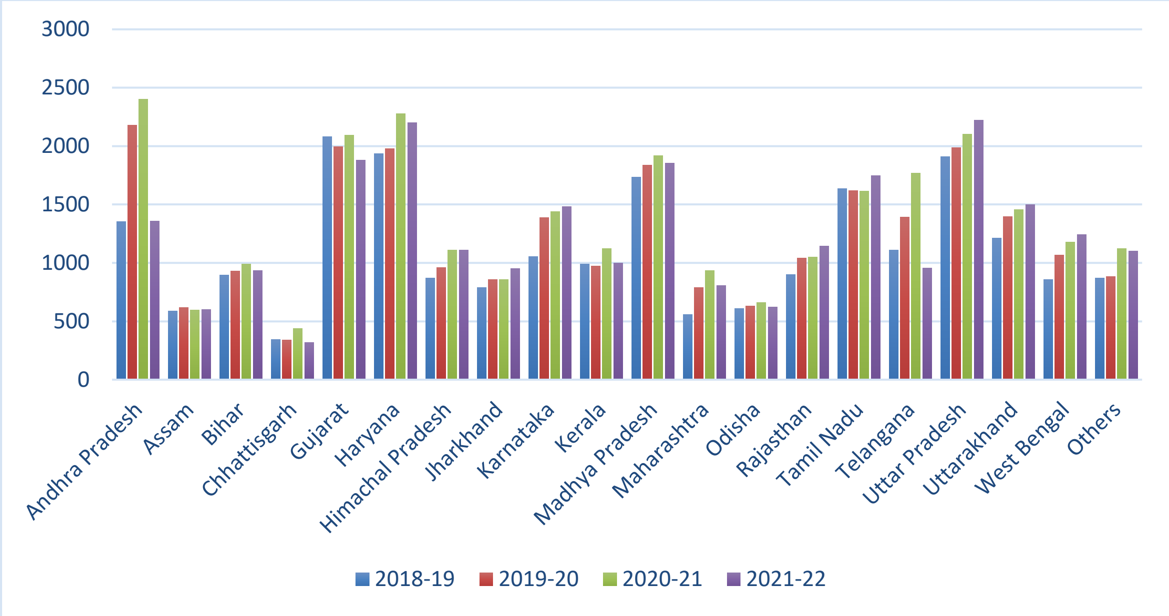


Figure 3: State-wise Yield (in kg/hectare) of millets from 2018-2021

Table 4: Summary of Analysis of Variance (ANOVA) for State-wise Area (in hectares) under Cultivation of Millets

Source of Variation	DF	Sum of Squares	Mean Squares	F-Calculated	Significance
Between	19	97,969,435.185	5,156,286.062	79.243	< 0.001
Within (Error)	60	3,904,153.613	65,069.227		
Total	79	101,873,588.798			

The above table 4 outlined the summary of ANOVA for State-wise area under cultivation of millets. Since p-value < 0.001 for the analysis was found to be less than α=0.05, level of significance. Therefore, there were significant difference among different states of India in accordance to area under cultivation of millets.

Table 4: Summary of Analysis of Variance (ANOVA) for State-wise Production (in tonnes) of Millets

Source of Variation	DF	Sum of Squares	Mean Squares	F-Calculated	Significance
Between	19	110,328,777.664	5,806,777.772	72.297	< 0.001
Within (Error)	60	4,819,135.018	80,318.917		
Total	79	115,147,912.682			

The above table 4 described the summary of ANOVA for State-wise production of millets. Since p-value < 0.001 for the analysis was found to be less than α=0.05, level of significance.

Therefore, there were significant difference among different states of India in terms of production of millets.

Table 5: Summary of Analysis of Variance (ANOVA) for State-wise Yield (in kg/hectare) of Millets

Source of Variation	DF	Sum of Squares	Mean Squares	F-Calculated	Significance
Between	19	20,521,565.238	1,080,082.381	32.888	< 0.001
Within (Error)	60	1,970,494.250	32,841.571		
Total	79	22,492,059.488			

The above table 5 had showed the summary of ANOVA for State-wise yield of millets. Since p-value < 0.001 for the analysis was found to be less than α=0.05, level of significance. Therefore, there were significant difference among different states of India in terms of yield of millets.

Table 6: Summary of Tukey's Honestly Significant Difference (HSD) test for State-wise Area under Cultivation (in hectares) of Millets

State Name	Means
Andhra Pradesh	183.5^{ef}
Assam	5.1^{f}
Bihar	11.65^{f}
Chhattisgarh	70.76^{ef}
Gujarat	424.94^{ef}
Haryana	500.32^{e}
Himachal Pradesh	7.12^{f}
Jharkhand	17.53^{f}
Karnataka	1546.6^{c}
Kerala	0.405^{f}
Madhya Pradesh	530.97^{e}
Maharashtra	2241.21^{b}
Odisha	83.17^{ef}
Rajasthan	4728.74^{a}
Tamil Nadu	513.96^{e}
Telangana	75.6^{ef}
Uttar Pradesh	1060.75^{d}
Uttarakhand	143^{ef}
West Bengal	7.89^{f}
Others	63.27^{ef}

From table 6, it could be observed that there exist pairwise significant difference between different states of India in terms of area under cultivation of millets. It has outlined that the state of Assam has a pairwise significant difference with Andhara Pradesh, Chattisgarh, Gujrat, Haryana, Karnataka, Madhya Pradesh, Maharastra, Odhisa, Rajhasthan, Tamil Nadu, Telangana, Uttar Pradesh, Uttarakhand and Others. Further, from table 6, similar inference pairwise significant and non-significant (on par) difference among different states of India in terms of area under cultivation of millets.

Table 7: Summary of Tukey's Honestly Significant Difference (HSD) test for State-wise Production (in tonnes) of Millets

State Name	Means
Andhra Pradesh	351.11^{defg}
Assam	3.07^{g}
Bihar	10.87^{g}
Chhattisgarh	25.45^{g}
Gujarat	864.91^{de}
Haryana	1052.32^{cd}
Himachal Pradesh	7.25^{g}
Jharkhand	15.15^{g}
Karnataka	2051.34^{b}
Kerala	0.41^{g}
Madhya Pradesh	975.86^{d}
Maharashtra	1731.14^{bc}
Odisha	52.44^{g}
Rajasthan	4880.38^{a}
Tamil Nadu	$844.04d^{ef}$
Telangana	108.88^{fg}
Uttar Pradesh	2179.5^{b}
Uttarakhand	198.75^{efg}
West Bengal	8.41^{g}
Others	61.75^{g}

From table 7, it could be observed that there exist pairwise significant difference between different states of India in accordance to production of millets. It has outlined that the state of Assam has a pairwise significant difference with Andhara Pradesh, Gujrat, Haryana, Karnataka, Madhya Pradesh, Maharastra, Rajhasthan, Tamil Nadu, Telangana, Uttar Pradesh and

Uttarakhand. Further, from table 7, similar inference on pairwise significant and non-significant (on par) difference among different states of India in terms of production of millets.

Table 8: Summary of Tukey's Honestly Significant Difference (HSD) test for State-wise Yield (in kg/hectare) of Millets

State Name	Means
Andhra Pradesh	1822.75ab
Assam	601gh
Bihar	938.25defg
Chhattisgarh	361h
Gujarat	2012.25a
Haryana	2097.5a
Himachal Pradesh	1014defg
Jharkhand	864.25efg
Karnataka	1341.25cd
Kerala	1020.5defg
Madhya Pradesh	1835.75ab
Maharashtra	771.5fgh
Odisha	630.5fgh
Rajasthan	1034.5defg
Tamil Nadu	1655.25abc
Telangana	1307.75cde
Uttar Pradesh	2055.25a
Uttarakhand	1390.75bcd
West Bengal	1086.5def
Others	994.25^{defg}

From table 8, it could be observed that there exist pairwise significant difference between different states of India in accordance to yield of millets. It was observed that in terms of yield, state of Assam has a pairwise significant difference with all the different states of India. Further, from table 8, some states of India had pairwise significant and non-significant (on par) difference with other states.

Discussion and Conclusion:

The statistical investigation of millets in India reveals significant trends in area, production, and consumption, highlighting both challenges and opportunities for this nutritious crop. Millets, often termed "Miracle Grains," are recognized for their resilience and health benefits, yet their cultivation and consumption have faced declines over the years. The area under millet cultivation has been decreasing, with a reported decline of 5.22% per annum from 1966-67 to 2020-21 (Kumar *et al.,* 2024). Major millets like pearl millet and sorghum have shown mixed trends; while pearl millet's production increased by 2.34% annually, sorghum's area and production decreased by 3.12% and 2.14%, respectively (Kumar *et al.,* 2024). Overall, the growth rate of millet area and production has been negative, with a decline of 16.31% and 13.58% per year, respectively, until 2005, after which productivity began to improve (Gowri, 2020). While the area and production of millets are declining, the yield of finger millet has shown a slight increase at a Compound Annual Growth Rate (CAGR) of 0.12% (Ramavathi & Gandhimathy, 2025). This research study was undertaken for systematic data-driven inquiry of area, production and yield of millets among different states of India. It had also tried to examined the variations arise due to area, production and yield of millets among different states of India. From the statistical investigation of area, production and yield of millets it was observed that the area, production and yield of millets haven't constantly increased or decreased during the period from 2018-2021. For states like Rajhasthan, area under cultivation of millets and production was highest. It was also come to noticed that since area under cultivation of millets for Assam was second least, therefore it has low production and yield of millets. Further, from pairwise comparison test it was found that there was significant difference among different states of India in accordance with area, production and yield of millets. The study also tried to outlined that in terms of yield of millets; state of Assam has a pairwise significant difference with all the different states of India.

References:

1. Baban, A. B. (2024). Study on Analysis of Variance in Statistics. *International Journal of Advanced Research in Science, Communication and Technology*, 358–363. https://doi.org/10.48175/IJARSCT-22741
2. Barman, S., Borah, A., Das, P., Deka, P., Saikia, P., Borah, J., & Deka, B. (2025). Economics and Adoption Trends of Millet crop in Central Brahmaputra Valley Zone of Assam, India. *Journal of Scientific Research and Reports*, *31*(9), 347–352. https://doi.org/10.9734/jsrr/2025/v31i93498
3. Blbas, H. T. A. (2024). *Descriptive Statistics*. IntechOpen. https://doi.org/10.5772/intechopen.1002179
4. Borbaruah, N., Srivastava, S. K., Kalita, B., & Deka, N. (2025). Assessing the Role of Small Millets in Assam's Agrifood System, India: A Comparative Analysis with Major Crops. *Journal of Experimental Agriculture International*, *47*(7), 953–960. https://doi.org/10.9734/jeai/2025/v47i73638

5. Cheshkova, A. F., Aleynikov, A. F., Grebennikova, I. G., & Stepochkin, P. I. (2017). *Application of the r software environment for analysis of variance in breeding research*. *47*(5), 88–96. https://doi.org/10.26898/0370-8799-2017-5-11
6. Digra, M., Gogoi, M., Gogoi, N., & Jarh, A. (2025). Assessment of Sensory and Nutritional Attributes of Finger Millet-based Traditional Food Products of Assam, India. *Journal of Advances in Biology & Biotechnology*, *28*(1), 364–379. https://doi.org/10.9734/jabb/2025/v28i11891
7. Gowri, M. U. (2020). Millet Scenario in India. *Economic Affairs*, *65*(3), 363–370. https://doi.org/10.46852/0424-2513.3.2020.7
8. Gupta, H., Panigrahi, T., Patle, U., Pandre, E., Singh, T., & Shweta, K. (2024). Revitalizing Millets for Nutritional Security & Fortifying Food and Uplifting Small and Marginal Farmers in India: A Review. *International Journal of Enviornment and Climate Change*. https://doi.org/10.9734/ijecc/2024/v14i13789
9. Kumar, S., Mehta, V., Mourya, K. K., & Kumar, A. (2024). A Comprehensive Study on Trend Analysis of Area, Production and Productivity of Major Millets in India. *Environment and Ecology*, *42*(4C), 2030–2036. https://doi.org/10.60151/envec/QQIP4860
10. Ramavathi, R., & Gandhimathy, B. (2025). A Study on Trends and Instability of Finger Millets in India. *Shanlax International Journal of Economics*, *13*(3), 79–89.
11. Yadav, R., Singh, L. B., Patel, A., Kumar, M., Kumar, M., Pandey, M., Yadav, S., Kumar, S., & Singh, S. (2023). A Review on Yield Gap Analysis of Millets in India. *International Journal of Plant and Soil Science*, *35*(18), 1800–1804. https://doi.org/10.9734/ijpss/2023/v35i183462
12. Varalakshmi, G. (2025). Application of Analysis of Variance (ANOVA) in Biological Science Research. *Journal of Experimental Agriculture International*, *47*(9), 520–527. https://doi.org/10.9734/jeai/2025/v47i93772

COMPREHENSIVE REVIEW OF ORGANIC-INORGANIC HYBRID PERVOKSITE PHOTODETECTORS

Shalu C.*, Sheenu Agarwal, Neelu Trivedi and Shakti Sindhu

Department of Electronics and Communication Engineering,

IFTM University, Moradabad, India,

*Corresponding author E-mail: shalu268@gmail.com, shalu@iftmuniversity.ac.in

Abstract:

The rapid advancement of photonic technologies has created a pressing need for high-performance photodetectors that can efficiently convert light into electrical signals. The development of high-performance photo detectors is crucial for various applications, including space lidars, optical communication, biological system, imaging, and sensing. However, the traditional photo detectors based on inorganic materials often suffer from limitations such as high dark current, low responsivity, and limited spectral range. The recent advancements have led to improved sensitivity, spectral response, and radiation tolerance. This review article presents a study of hybrid organic–inorganic perovskite materials. A brief introduction with the relationship among the perovskite structures, device configurations, and photodetection performances of hybrid organic–inorganic photodetectors is presented in this article. Due to the attractive optoelectronic properties, quantum confinement effect, color tunability, high defect tolerance, high photoluminescence quantum yield, high optical absorption coefficient, high carrier mobility, and long carrier diffusion length, perovskites play a crucial role for the implementation of high-performance photodetectors.

Keywords: Conjugated Polymers, Graphene, Hybrid Photodetectors, Imaging, Organic Photodetectors, Optical Communication, Perovskite Materials.

1. Introduction:

Photodetectors are devices that convert light energy into electrical energy [1]. Traditional photo detectors are based on inorganic materials and have limitations in terms of flexibility, cost, and sensitivity [2]. Organic/inorganic hybrid photodetectors have emerged as a promising alternative, offering improved performance, flexibility, and cost-effectiveness [3]. Photodetectors gather optical signals with wide range of incident irradiance and light intensity that is converted to the electrical signals immediately [4]. They have many important applications including space lidars, optical communication, remote control, chemical/biological system, imaging, and sensing [5]. Currently, GaN (Gallium nitride), Si (Silicon), Ge (Germanium), InGaAs (Indium Gallium Arsenide) and HgCdTe (mercury cadmium telluride) photodetectors are used in commercially available products [1]. HgCdTe are used to create hybrid photodiodes. It is a semiconductor material, commonly used in infrared (IR) detectors [6]. These hybrid photodiodes are used in various applications, including second-generation IR imaging systems [7]

In this review paper, a comprehensive and systematic discussion of the hybrid organic–inorganic perovskite materials their relationship among the perovskite structures, device configurations, and photodetection performances is done. Based on above we seek to develop the understanding for the design and implementation of a hybrid photodetector.

2. Hybrid Photodetectors

Hybrid photo detectors generally combine the dissimilar materials to increase the individual strengths of materials [5,7]. Graphene is combined with quantum dots for enhanced light absorption and carrier transport [8]. Perovskites can be combined with graphene for efficient light harvesting and fast response times [9]. In some cases, semiconductors like Si or InGaAs can be integrated with 2D materials like graphene or MoS_2 to improve sensitivity and performance in specific wavelength ranges [10].

Perovskites and graphene are both materials with diverse applications, but they differ significantly in their composition and properties [10,11]. Perovskites are compounds with a specific crystal structure, often ABX_3, where A and B are cations and X is an anion [12]. They can be categorized into inorganic oxide perovskites, alkaline metal halide perovskites, and organic metal halide perovskites. Graphene, on the other hand, is a single layer of carbon atoms arranged in a hexagonal lattice, a two-dimensional allotrope of carbon. It can also exist in multi-layered forms [13].

For the miniaturization and improved performance of optoelectronic devices, with increasing demands for sustainability and energy efficiency, researchers are trying to find new materials [14]. This new material offers good performance by transforming current technologies and reducing costs and energy usage. Graphene and its derivatives have properties that it can be rendered for the development of new generation of photodetectors [15]. Current advancements in photodetector technology led to improved sensitivity, scalability, spectral response and radiation tolerance. Various developments in the areas of materials and their structure have paved the way for the advancements in the photodetector technology.

Novel materials like perovskites, polymers, graphene, and quantum dots offer unique optoelectronic properties, high tunability that helps to enhance the performance of photodetectors and their physics and structures helps to achieve high figure-of-merit, which in turn improves quantum efficiency and stability. Various types of photo detectors are available for different innovative applications out of which graphene-semiconductor hybrid photodetectors improve quantum efficiency, reduce noise and increase stability and perovskite photodetectors offer high sensitivity, fast response speed, low noise, high gain and narrowband detection.

Perovskites photodetectors gained significant attention due to their excellent optoelectronic properties [16]. Structure of Perovskites photodetectors generally have the ABX3 structure. Different types of perovskites photodetectors are inorganic oxide, alkaline metal halide, and organic metal halide examples are CaTiO3, CH3NH3PbI3 (methylammonium lead iodide [14,17]. Applications of various perovskites photodetectors are fuel cells, memory devices, and photovoltaics [18]. Because of unique properties such as High strength, flexibility, electrical conductivity, high responsivity, high sensitivity, adequate frequency response and high quantum efficiencies of Graphene photodetectors, they have shown exceptional performance in various applications including high-speed data communication, sensing, storage, construction, biomedicine and imaging. Graphene structure is a single layer of carbon atoms in a hexagonal lattice with very few layers graphene (vFLG), few layer graphene (FLG), multi-layer graphene (MLG), and graphene nanoplatelets (GNP) [19].

In essence, perovskites are characterized by their crystal structure and diverse chemical composition, while graphene is defined by its unique two-dimensional carbon structure and exceptional properties [20] Graphene is nearly transparent, absorbing only 2.3 % of visible light, making it an ideal candidate for applications such as touch screens, solar cells, and optoelectronic devices [21].

2.1 Perovskites Photodetectors

Hybrid Perovskite Photodetectors is an organic-inorganic hybrid perovskites (e.g., CH3NH3PbI3) with properties such as high sensitivity, ultrafast response speed, and high gain have applications in the field of visible light detection, X-ray detection, and imaging [21,24]). All-Inorganic Perovskite Photodetectors uses only inorganic perovskites material (e.g., CsPbI3, CsPbBr3) with properties such as improved stability, high sensitivity, and narrowband detection have applications in the field of visible light detection, UV detection, and optoelectronic devices. Photodetectors uses perovskite quantum dots materials (e.g., CsPbX3, X = Cl, Br, I) with properties such as high sensitivity, tunable bandgap, and narrowband detection have applications in the field of Visible light detection, NIR detection, and optoelectronic devices

2D Perovskite Photodetectors uses 2D perovskites material (e.g., (BA)2PbI4) with properties such as high sensitivity, ultrafast response speed, and high anisotropy have applications in the areas of visible light detection, optoelectronic devices, and flexible electronics

Flexible Perovskite Photodetectors uses Perovskites on flexible substrates as the material (e.g., PET, PDMS) with properties such as high sensitivity, flexibility, and finds applications in wearable electronics, biomedical devices, and flexible optoelectronics [19].

These types of perovskite photo detectors offer unique properties and applications, making them promising candidates for various optoelectronic devices [17]. New photodetector structures and physics emphasize devices that achieve high figure-of-merit, such as graphene-semiconductor hybrids, which improve quantum efficiency and stability. Perovskite photo detectors can be classified based on Structures, Material Composition, Dimensionality, Detection Mechanism, Spectral Response, Device Structure or Application [17].

2.2 Diverse Perovskite Photo Detectors

These classifications highlight the diversity of perovskite photo detectors and their potential applications. 1D perovskites refer to perovskite materials with a one-dimensional (1D) structure, typically in the form of nanowires, nanorods, or nanotubes [15]. These materials have unique properties due to their anisotropic shape that is with one dimension much larger than the others and quantum confinement effects, which can enhance optical and electrical properties. 1D perovskites can be synthesized with varying compositions, sizes, and shapes, allowing for tunable properties [23]. 1D perovskites exhibit unique optical properties, such as polarized emission and enhanced absorption. 1D perovskites can exhibit high conductivity and mobility, making them suitable for electronic devices. 1D perovskites can be more stable than their 3D counterparts due to their reduced dimensionality [23]. 1D perovskites are promising materials for optoelectronic devices, such as photodetectors, LEDs, and solar cells [16]. 1D perovskites can be used for sensing applications, such as chemical and biological sensing [20]. 1D perovskites can be used for energy harvesting applications, such as solar energy conversion with controlled size and shape can be challenging. 1D perovskites can be sensitive to environmental factors, such as moisture and heat, which can affect their stability [24]. Scaling up the synthesis and fabrication of 1D perovskites can be challenging [16]. Overall, 1D perovskites offer unique properties and potential applications, but further research is needed to overcome the challenges associated with their synthesis, stability, and scalability.

2D perovskites refer to perovskite materials with a two-dimensional (2D) structure, typically in the form of layered or sheet-like materials [19]. These materials have gained significant attention due to their unique properties and potential applications. 2D perovskites consist of layers of metal halide octahedra separated by organic or inorganic cations [20,23]. The 2D structure leads to quantum confinement effects, which can enhance optical and electrical properties. 2D perovskites can be synthesized with varying compositions, layer thicknesses, and surface chemistries, allowing for tunable properties [24]. 2D perovskites exhibit unique optical properties, such as narrow emission spectra and high photoluminescence quantum yields [24]. 2D perovskites can exhibit high mobility and conductivity, making them suitable for electronic devices. 2D perovskites can be more stable than their 3D counterparts due to their reduced dimensionality. 2D perovskites are promising materials for optoelectronic devices, such as LEDs, solar cells, and photodetectors. 2D perovskites can be used for electronic devices, such as transistors and sensors. 2D perovskites can be used for energy storage applications, such as batteries and supercapacitors. Ruddlesden-Popper (RP) phase is a type of 2D perovskite with a specific crystal structure. Dion-Jacobson (DJ) phase is another type of 2D perovskite with a different crystal structure. Synthesizing high-quality 2D perovskites with controlled layer thickness and composition can be challenging [21]. 2D perovskites can be sensitive to environmental factors, such as moisture and heat, which can affect their stability. Scaling up the synthesis and fabrication of 2D perovskites can be challenging. Overall, 2D perovskites offer unique properties and potential applications, and researchers are actively exploring their synthesis, properties, and applications.

3D perovskites refer to perovskite materials with a three-dimensional (3D) crystal structure [17]. These materials have gained significant attention due to their unique properties and potential applications. 3D perovskites have a crystal structure composed of corner-sharing metal halide octahedral. D perovskites typically have high symmetry, which can lead to unique optical and electrical properties. 3D perovskites can be synthesized with varying compositions, allowing for tunable properties. 3D perovskites exhibit unique optical properties, such as high absorption coefficients and narrow emission spectra. 3D perovskites can exhibit high mobility and conductivity, making them suitable for electronic devices. 3D perovskites can exhibit efficient charge transport, making them suitable for applications such as solar cells. 3D perovskites are promising materials for solar cells due to their high-power conversion efficiency and low-cost fabrication. 3D perovskites can be used for LEDs due to their high luminescence efficiency and narrow emission spectra. 3D perovskites can be used for photodetectors due to their high sensitivity and fast response times. 3D perovskites can be composed of organic and inorganic components or can be composed entirely of inorganic components. 3D perovskites can be sensitive to environmental factors, such as moisture and heat, which can affect their stability. Some 3D perovskites contain toxic elements, such as lead, which can be a concern for environmental and health applications. Scaling up the synthesis and fabrication of 3D perovskites can be challenging. Overall, 3D perovskites offer unique properties and potential applications, and researchers are actively exploring their synthesis, properties, and applications [17,18 and 20].

0D perovskites, also known as perovskite quantum dots or nanoparticles, refer to perovskite materials with a zero-dimensional (0D) structure [20]. These materials have unique properties due to their small size and quantum confinement effects. 0D perovskites have a size range of 1-10 nanometers, which leads to quantum confinement effects. 0D perovskites can be considered as quantum dots, which exhibit unique optical and electrical properties. 0D perovskites can be synthesized with varying compositions, sizes, and surface chemistries, allowing for tunable properties. 0D perovskites exhibit unique optical properties, such as high photoluminescence quantum yields and narrow emission spectra [18,21]. 0D perovskites can exhibit unique electrical properties, such as high charge carrier mobility and conductivity. 0D perovskites can be more stable than their bulk counterparts due to their small size and surface passivation. 0D perovskites are promising materials for optoelectronic devices, such as LEDs, solar cells, and photodetectors. 0D perovskites can be used for biological imaging due to their high photoluminescence efficiency and biocompatibility. 0D perovskites can be used for sensing applications, such as chemical and biological sensing. Synthesizing high-quality 0D perovskites with controlled size and composition can be challenging. 0D perovskites can be sensitive to environmental factors, such as moisture and heat, which can affect their stability. Scaling up the synthesis and fabrication of 0D perovskites can be challenging. 0D perovskites have a high surface area-to-volume ratio, which can enhance their optical and electrical properties. 0D perovskites can be synthesized with tunable properties, which can be beneficial for various applications. Overall, 0D perovskites offer unique properties and potential applications, and researchers are actively exploring their synthesis, properties, and applications. 0D perovskites, also known as perovskite quantum dots or nanoparticles, typically have a structure consisting of a core composed of metal halide octahedra (e.g., PbX6, where X = Cl, Br, I) or a shell surrounding the core, which can be composed of organic ligands or inorganic materials. The surface of the 0D perovskite can be dedicated with ligands or other materials to enhance stability and optical properties. The structure of 0D perovskites can vary depending on the synthesis method, composition, and surface chemistry. Some common structures include: 0D perovskites can form nanocrystals with a specific shape and size. 0D perovskites can exhibit quantum dot behavior, with unique optical properties due to quantum confinement effects. The structure of 0D perovskites plays a crucial role in determining their optical, electrical, and stability properties, making them suitable for various applications.

1D perovskites typically have a structure consisting of Nanowires or nanorods. 1D perovskites often form nanowires or nanorods with a high aspect ratio. The crystal structure of 1D perovskites is anisotropic, meaning it has different properties in different directions. The structure is composed of metal halide octahedra (e.g., PbX6, where X = Cl, Br, I) that are connected in a specific way to form the 1D structure. These are thin, wire-like structures or rod-like structures with a high aspect ratio. These are tube-like structures with a hollow center. The structure of 1D perovskites can be tuned by controlling the synthesis conditions, composition, and surface chemistry, which can affect their optical, electrical, and stability properties.

2D perovskites typically have a structure consisting of layered structure, with metal halide octahedra (e.g., PbX6, where X = Cl, Br, I) forming a 2D network. The layers are separated by organic or inorganic spacers, which can be tuned to control the properties. 2D structure leads to quantum confinement effects, which can enhance optical and electrical properties. Ruddlesden-Popper (RP) phase is a type of 2D perovskite with a specific crystal structure. Dion-Jacobson (DJ) phase is another type of 2D perovskite with a different crystal structure. The structure of 2D perovskites can be tuned by controlling the synthesis conditions, composition, and spacer molecules, which can affect their optical, electrical, and stability properties.

3D perovskites typically have a structure consisting of a crystal structure composed of corner-sharing metal halide octahedra (e.g., PbX6, where X = Cl, Br, I). The octahedra form a three-dimensional network, which can lead to unique optical and electrical properties. The composition of 3D perovskites can be tuned by varying the metal and halide ions, which can affect their properties. The structure of 3D perovskites can be described by the general formula ABX3, where The A-site cation (e.g., Cs+, CH3NH3+) occupies the cube octahedral cavity. B-site cation: The B-site cation (e.g., Pb2+, Sn2+) forms the metal halide octahedra. X-site anion: The X-site anion (e.g., Cl-, Br-, I-) forms the halide ions. The structure of 3D perovskites plays a crucial role in determining their optical, electrical, and stability properties, making them suitable for various applications.

Conclusion:

Perovskite photodetectors have gained significant attention due to their exceptional performance and versatility. Here's comprehensive overview perovskite photodetectors are devices that convert light into electrical signals using perovskite materials. These materials have a unique crystal structure that allows for efficient charge transport and high sensitivity to light. Simple devices that change conductivity in response to light are called perovskite Photoconductors. Perovskite Photodiodes are devices that generate current when exposed to light and Perovskite Phototransistors are devices that amplify current in response to light. Perovskite photodetectors offer efficient charge transport and high absorption coefficients. They are suitable for applications requiring quick detection. Reduced signal fluctuations for accurate detection and are able to detect at specific wavelengths. Have potential applications in imaging, surveillance, and biomedical optics. Device stability and longevity need to be improved. Scaling up production while maintaining performance is key requirement addressing concerns related to lead-based perovskites.

Researchers are actively exploring solutions to these challenges and pushing the boundaries of perovskite photodetector performance [1]. Graphene photo detectors have shown exceptional performance in various applications due to their unique properties. Some key aspects of the performance of graphene photo detectors are they can achieve high responsivity, with values ranging from 0.2 A/W to 114.64 A/W, depending on the specific design and material combination used. The detectivity of graphene photo detectors can reach up to 1.24 × 108 Jones, indicating high sensitivity to light. Graphene photodetectors can operate at high speeds, with some devices demonstrating frequency responses up to 110 GHz. Graphene photo detectors can detect light across a broad wavelength range, including near-infrared, mid-infrared, and even X-ray regions. Some graphene photodetectors have achieved external quantum efficiencies of up to 21,702.23%. Some notable performance metrics for specific graphene photodetector designs include Graphene/silicon/graphene photodetectors devices have shown high performance across a broad wavelength range, from X-ray to near infrared. Graphene/Ge heterojunction photo detectors by modulating the doping level of graphene, these devices can achieve improved responsivity and detectivity. Waveguide-integrated graphene photodetectors have demonstrated high-speed operation, with frequency responses up to 110 GHz and data reception capabilities of up to 100 Gbit/s. Overall, graphene photodetectors offer a promising solution for various applications, including high-speed data communication, sensing, and imaging.

References:

1. Acik, M., & Darling, S. B. (2016). Graphene in perovskite solar cells: Device design, characterization and implementation. *Journal of Materials Chemistry A, 4*(17), 6185–6235.
2. Ahmad, S. O. A., Ashfaq, A., Akbar, M. U., Ikram, M., Khan, K., Wang, F., Ikram, M., & Mahmood, A. (2021). Application of two-dimensional materials in perovskite solar cells: Recent progress, challenges, and prospective solutions. *Journal of Materials Chemistry C, 9*(40), 14065–14092.
3. Ahmadi, M., Wu, T., & Hu, B. (2017a). A review on organic–inorganic halide perovskite photodetectors: Device engineering and fundamental physics. *Advanced Materials, 29*(41), 1605242.
4. Ahmadi, M., Wu, T., & Hu, B. (2017b). A review on organic–inorganic halide perovskite photodetectors: Device engineering and fundamental physics. *Advanced Materials, 29*(41), 1605242.
5. Alasmari, A., Shariq, M., Alhazmi, N. E., Alzahrani, H. S., Bouzgarrou, S. M., Alkhayri, F., Alzahrani, A., Alharbi, S. H., & Azeez, N. A. (2024). Enhancing perovskite solar cells with graphene-based nanocomposites for sustainable energy: A comprehensive review. *Diamond and Related Materials, 148*, 111517.
6. Almutlaq, J., Yin, J., Mohammed, O. F., & Bakr, O. M. (2018a). The benefit and challenges of zero-dimensional perovskites. *The Journal of Physical Chemistry Letters, 9*(14), 4131–4138.
7. Almutlaq, J., Yin, J., Mohammed, O. F., & Bakr, O. M. (2018b). The benefit and challenges of zero-dimensional perovskites. *The Journal of Physical Chemistry Letters, 9*(14), 4131–4138.
8. Ansari, A., Ahmed, S., Mashniwi, M. H. J., Shinde, S. M., Khan, A., Ranjan, P., & Negi, D. S. (2025). Review: Recent advancements in photodetection utilizing inorganic, organic low-dimensional materials and their hybrids. *Journal of Materials Science, 60*(6), 2711–2743.
9. Shalu, C., Mayoorika Shukla, Ananya Tiwari, Anubha Bilgayian, & Singh, V. (2020). Role of solvent used to cast P3HT thin films on the performance of ZnO/P3HT hybrid photodetector. *Physica E, 115*, 113694.
10. Hu, M., Zhang, Y., Meng, N., Wang, W., Lu, Y., Dong, J., Zhao, S., Qiao, B., Song, D., & Xu, Z. (2023). Modulation phase distribution of Ruddlesden–Popper quasi-2D perovskites with a similarly spaced Dion–Jacobson phase. *ACS Applied Materials & Interfaces, 15*(36), 42706–42716.

11. Kim, J. K., Park, M. J., Kim, S. J., Wang, D. H., Cho, S. P., Bae, S., Park, J. H., & Hong, B. H. (2013). Balancing light absorptivity and carrier conductivity of graphene quantum dots for high-efficiency bulk heterojunction solar cells. *ACS Nano, 7*(8), 7207–7212.
12. Shalu, C., Nidhi Yadav, Kshitij Bhargava, Mukesh P. Joshi, & Vipul Singh. (2018). All organic near ultraviolet photodetectors based on bulk heterojunction of P3HT and DH6T. *Semiconductor Science and Technology, 33*, 095021.
13. Lan, C., Zhou, Z., Wei, R., & Ho, J. C. (2019). Two-dimensional perovskite materials: From synthesis to energy-related applications. *Materials Today Energy, 11*, 61–82.
14. Lee, B., Nasir, T., Cho, S., Jung, M., Kim, B. J., Lee, S. H., Jang, H. E., Kang, J., Yu, H. K., & Choi, J.-Y. (2024). Highly sensitive photodetector based on transfer-free 3-D graphene on SiO2. *IEEE Sensors Journal, 24*(5), 6038–6044.
15. Li, G., Wang, Y., Huang, L., & Sun, W. (2022). Research progress of high-sensitivity perovskite photodetectors: A review of photodetectors: Noise, structure, and materials. *ACS Applied Electronic Materials, 4*(4), 1485–1505.
16. Lou, Z., & Shen, G. (2016). Flexible photodetectors based on 1D inorganic nanostructures. *Advanced Science, 3*(6), 1500287.
17. Nabet, B. (2023). *Photodetectors: Materials, devices and applications*. Woodhead Publishing.
18. Kore, P. B., Jamshidi, M., & Gardner, J. M. (2024). The impact of moisture on the stability and degradation of perovskites in solar cells. *Materials Advances, 5*(6), 2200–2217.
19. Razdan, A. K., & Veerabuthiran, S. (2014). LIDAR: A precision long range standoff sensor of atmospheric pollutants, chemical, biological and explosive agents. In *Advances in Laser Physics and Technology* (pp. 308–330). Foundation Books.
20. Koppens, F. H. L., Mueller, T., Avouris, P., Ferrari, A. C., Vitiello, M. S., & Polini, M. (2014). Photodetectors based on graphene, other two-dimensional materials and hybrid systems. *Nature Nanotechnology, 9*(10), 780–793.
21. Korotcenkov, G. (2023). IR detectors array. In *Handbook of II-VI semiconductor-based sensors and radiation detectors: Volume 2, Photodetectors* (pp. 79–106). Springer International Publishing.
22. Shalu, C., Kshitij Bhargava, Nidhi Yadav, Mukesh P. Joshi, & Vipul Singh. (2019). Effect of concentration of DH6T on the performance of photoconductor fabricated using blends of P3HT and DH6T. *Optical Materials, 89*, 214–223.
23. Rogalski, A., Antoszewski, J., & Faraone, L. (2009). Third-generation infrared photodetector arrays. *Journal of Applied Physics, 105*(9), 091101.
24. Sahudin, M. R., Zourob, M., Noor Akashah, M. H., Abdul Rani, R., & Makhsin, S. R. (2023). Physical properties of graphene. *Journal of Mechanical Engineering (JMechE), 12*(1), Article 1.

REMOTE SENSING AND UAV APPLICATIONS FOR SMART FARMING SYSTEMS

K. Prabhu*[1] and A. Jegatheesan[2]

[1]Department of Physics, AMET University, Kanathur, Tamilnadu - 603112.

[2]Department of Physics, Paavai Engineering College (Autonomous), Namakkal, Tamilnadu- 637018.

*Corresponding author E-mail: k.prabhuresearch@gmail.com

Abstract:

Remote sensing and unmanned aerial vehicle (UAV) technologies are rapidly reshaping smart farming systems by enabling timely, high-resolution monitoring of crops and soils across diverse agricultural landscapes. Satellite imagery provides repeated, large-area observations for assessing vegetation dynamics, drought stress, and seasonal crop development, while UAV platforms offer fine-scale measurements that reveal within-field variability in plant vigor, nutrient deficiencies, irrigation performance, and pest damage. When integrated with geographic information systems, Internet-of-Things sensor networks, and farm management software, these data streams support predictive and site-specific decision making rather than uniform field treatments. This chapter reviews recent advances in sensor technologies, image-processing workflows, and analytics frameworks that underpin satellite- and drone-based smart farming applications. Multispectral, hyper-spectral, thermal, and LiDAR sensors are examined for their capacity to characterize canopy structure, biomass distribution, and crop water status. The growing use of machine learning and computer vision to convert imagery into operational recommendations, such as variable-rate input maps and early-warning alerts is also discussed, along with practical challenges including calibration, regulatory compliance, data fusion, and large-volume image management. Emerging trends such as autonomous UAV operations, edge computing, and near-real-time analytics are highlighted as key enablers of next-generation digital agriculture.

Keywords: Remote Sensing, UAV operations, Smart Farming, Precision Farming, Crop Monitoring, Sustainable Agriculture.

Introduction:

Agriculture is increasingly confronted with complex and interrelated challenges that include climatic variability, declining water availability, soil degradation, labor shortages, and the necessity to produce more food from shrinking arable land. These pressures are particularly acute in rapidly developing agrarian economies, where smallholder-dominated systems must adapt to unpredictable monsoon patterns, rising temperatures, and intensified pest outbreaks. In India, where agriculture sustains a large proportion of rural livelihoods and contributes substantially to national food security, the adoption of digital and spatial technologies is gaining strategic importance (Chakraborty *et al.,* 2016; Sharma *et al.,* 2019). Smart farming systems integrating sensing platforms, analytics, and automated field operations are increasingly viewed as pathways toward resilient and resource-efficient agricultural production.

Remote sensing and unmanned aerial vehicles (UAVs) have become central components of these systems because they enable non-invasive, repeatable observation of crops and soils across wide spatial extents. Satellite missions provide continuous temporal records that are essential for monitoring crop growth, drought progression, and land-use transitions, while UAVs allow targeted, high-resolution surveys during critical growth stages or immediately after extreme weather events. In the Indian context, satellite-derived vegetation indices and evapotranspiration products have been widely applied for crop acreage estimation, yield forecasting, and irrigation planning at regional and national scales (Roy *et al.,* 2014; Panigrahy *et al.,* 2019). Drone-based sensing is similarly gaining traction for plot-level nutrient diagnosis, pest scouting, and crop insurance assessment, supported by evolving regulatory frameworks and government-led digital agriculture initiatives.

The growing availability of geospatial data alone, however, does not automatically translate into improved farm decisions. Agricultural landscapes are characterized by strong spatial heterogeneity in soil properties, topography, and management practices, as well as temporal variability driven by weather and crop phenology. Extracting meaningful signals from these complex datasets requires advanced analytical methods capable of handling nonlinearity, noise, and multiscale interactions. Machine-learning techniques, coupled with geographic information systems and cloud-computing infrastructures, have therefore become indispensable for converting remote-sensing observations into actionable agronomic recommendations (Kamilaris & Prenafeta-Boldú, 2018; Benos *et al.,* 2021). Indian researchers have similarly emphasized the role of data-driven approaches in linking satellite imagery with field surveys and agro-meteorological data to support site-specific nutrient management and early-warning systems for crop stress (Chakraborty *et al.,* 2016; Sharma *et al.,* 2019).

This chapter aims to clarify current capabilities, identify operational constraints, and outline research pathways that can enhance the contribution of remote sensing technologies to sustainable intensification and climate-resilient food production.

Satellite Remote Sensing

Satellite remote sensing has long been the backbone of agricultural monitoring at regional, national, and global scales. Indian agricultural research and planning have benefitted substantially from satellite programs such as the Indian Remote Sensing (IRS) series, as well as global missions like Landsat and the European Space Agency's Sentinel-2. These platforms collect data across multiple spectral bands ranging from visible light to near-infrared and shortwave infrared enabling the derivation of key vegetation indices such as the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Soil Adjusted Vegetation Index (SAVI) that correlate with photosynthetic activity and canopy conditions (Rao *et al.,* 2011; Panigrahy *et al.,* 2019). In India, these indices have been widely used for crop acreage estimation, monitoring crop phenology, and assessing drought impacts over large agricultural belts such as the Indo-Gangetic Plains (Rao *et al.,* 2011; Roy *et al.,* 2014).

Beyond vegetation indices, satellite data support estimation of biophysical parameters such as leaf area index, fractional vegetation cover, chlorophyll content, and surface temperature. Thermal infrared bands from satellites like Landsat 8 and MODIS allow calculation of land surface temperature and evapotranspiration, providing insights into crop water stress and irrigation management requirements (Mulla, 2013). Recent work has also employed hyperspectral satellite data to improve discrimination between crop types and to detect early signs of disease or nutrient deficiencies, although sensitivity and spatial resolution remain technological challenges at larger scales.

Unmanned Aerial Vehicles (UAVs)

While satellites offer frequent and broad coverage, their moderate spatial resolution can limit the detection of field-level variability. Unmanned Aerial Vehicles (UAVs), commonly referred to as drones, bridge this gap by capturing ultra-high resolution imagery that can identify small patches of crop stress, weed infestations, or irrigation issues that satellites cannot resolve (Sharma *et al.,* 2019). Indian research institutions and private agritech enterprises have increasingly adopted UAVs for precision crop monitoring, particularly in cash crops such as sugarcane, paddy, and cotton; where plot-level decisions can significantly affect yield and profitability (Reddy *et al.,* 2020; Singh & Gupta, 2021).

Equipped with a range of sensors including RGB cameras, multispectral systems, thermal imagers, and light detection and ranging (LiDAR) instruments UAVs can estimate canopy height, biomass distribution, and soil moisture variability with centimeter-level accuracy. For example, multispectral UAV imagery has been applied in India to map crop nitrogen status during critical growth stages, enabling targeted fertilizer application that reduces costs and environmental impact (Reddy *et al.,* 2020). Thermal drones are also used to identify irrigation inefficiencies or water stress before visible wilting occurs, supporting more precise application of water in water-limited environments.

Synergies and Challenges

Remote sensing systems provide complementary strengths: satellites deliver multispectral and multi-temporal data over large areas, while UAVs offer high spatial detail and flexibility in data acquisition timing. Integrating these data within geographic information systems (GIS) enables layered analyses that support decision making from farm to landscape scales.

However, several challenges persist. Satellite observations can be hindered by cloud cover, especially during monsoon seasons in India, reducing usable data frequency. UAV operations face regulatory constraints, require skilled pilots, and generate large volumes of data that demand robust processing infrastructure. Radiometric calibration, data fusion, and ground-truth validation remain critical areas of ongoing research to ensure that remote sensing products reliably inform agronomic interpretations (Panigrahy *et al.,* 2019; Singh & Gupta, 2021).

By advancing sensor technologies, data analytics, and integration frameworks, remote sensing platforms are poised to become even more central to future smart farming systems, enabling scalable, precise, and sustainable agricultural management.

Sensor Technologies and Information Extraction

Smart farming systems rely on a diverse set of airborne and space borne sensors to capture biophysical and biochemical properties of crops and soils. Multispectral sensors, typically operating in visible, near-infrared, and short-wave infrared regions, are widely used for mapping vegetation vigor, crop density, and nutrient stress. Hyperspectral sensors, which record reflectance in hundreds of narrow spectral bands, provide richer diagnostic capabilities for detecting subtle biochemical changes associated with disease onset, chlorophyll concentration, and water content (Mahlein *et al.,* 2013).

Thermal sensors record surface temperature variations that are closely linked to transpiration and stomatal conductance, making them valuable for irrigation scheduling and drought assessment (Jones *et al.,* 2004). LiDAR instruments mounted on UAVs or aircraft measure three-dimensional canopy structure, enabling estimation of plant height, biomass, lodging, and surface roughness (Villa-Henriksen *et al.,* 2020).

Indian studies increasingly emphasize multisensor integration for crop assessment. Panigrahy *et al.* (2019) demonstrated that combining multispectral satellite data with agro meteorological observations improved seasonal crop forecasting at regional scales, while Reddy *et al.* (2020) showed that UAV-derived spectral metrics enhanced nitrogen stress detection in irrigated rice systems.

Data Processing, GIS Integration, and Analytical Frameworks

Raw satellite or UAV imagery must undergo several preprocessing steps before it can be operationally applied. Radiometric correction compensates for atmospheric effects and sensor noise, while geometric correction aligns images to real-world coordinates. UAV datasets are commonly transformed into orthomosaics and digital surface models using photogrammetric techniques.

Geographic information systems (GIS) serve as integrative platforms where remote-sensing products are combined with soil maps, yield records, irrigation layouts, and management zones. Time-series analysis within GIS environments supports monitoring of phenological development and detection of anomalous growth patterns (Roy *et al.,* 2014).

Machine-learning algorithms play a central role in extracting agronomic meaning from high-dimensional imagery. Random Forest, support-vector machines, and deep convolutional neural networks have been applied for crop classification, disease detection, and biomass estimation (Kamilaris & Prenafeta-Boldú, 2018; Sharma *et al.,* 2019). Cloud-based infrastructures, such as large geospatial processing engines, increasingly facilitate national-scale monitoring efforts by enabling rapid analysis of multi-year satellite archives (Gorelick *et al.,* 2017).

Operational Applications in Smart Farming

1. Crop Health and Nutrient Diagnosis

Remote sensing has become indispensable for early identification of crop stress caused by nutrient deficiencies, pests, or diseases. Hyperspectral UAV imagery combined with classification algorithms can discriminate between healthy and infected plants at early growth stages, enabling timely intervention (Mohanty *et al.,* 2016). In Indian production systems, UAV-based spectral indicators have been tested for diagnosing nitrogen stress in cereals and cotton, supporting site-specific fertilizer recommendations (Reddy *et al.,* 2020).

2. Irrigation Management and Drought Monitoring

Thermal imagery from satellites and drones allows estimation of crop water stress and evapotranspiration, providing spatially explicit guidance for irrigation scheduling (Jones *et al.,* 2004). In semi-arid regions of India, satellite-derived drought indices and land-surface temperature products are increasingly used by agricultural agencies to guide contingency planning and crop insurance assessments (Roy *et al.,* 2014; Panigrahy *et al.,* 2019).

3. Weed Detection and Targeted Crop Protection

High-resolution UAV imagery enables discrimination between crops and weeds based on spectral and textural features, facilitating selective spraying and mechanical control (Dyrmann *et al.,* 2016). Such approaches reduce chemical usage and environmental contamination, aligning with sustainable intensification objectives promoted in Indian precision-agriculture programs (Singh & Gupta, 2021).

4. Yield Forecasting and Risk Assessment

Time-series satellite data integrated with weather observations and crop models have improved yield forecasting at district and national levels (Crane-Droesch, 2018). Indian crop-monitoring initiatives have adopted similar approaches to support food-security planning and market regulation, particularly for staple cereals (Panigrahy *et al.,* 2019).

Integration with IoT and Farm Management Systems

Smart farming increasingly relies on networks of in-field sensors that measure soil moisture, salinity, microclimate, and nutrient fluxes. These Internet-of-Things (IoT) devices complement aerial observations by providing continuous point-based measurements that calibrate and validate remote-sensing models (Ojha *et al.,* 2015).

Farm-management information systems integrate UAV and satellite products with machinery telemetry and agronomic records to generate decision-support dashboards for growers and extension services (Wolfert *et al.,* 2017). Indian digital agriculture initiatives have similarly emphasized interoperable platforms that link geospatial data with advisory services delivered through mobile applications (Chakraborty *et al.,* 2016).

Technical and Institutional Challenges

Despite their promise, several barriers constrain widespread adoption of remote-sensing-based smart farming. Persistent cloud cover during monsoon seasons limits optical satellite usability in South Asia. UAV deployment requires trained operators and compliance with evolving aviation regulations. Data volumes from hyperspectral or LiDAR surveys impose heavy computational demands, while calibration across sensors and seasons remains a methodological challenge (Tsouros *et al.,* 2019).

Socio-economic factors are equally important. Smallholder farmers may lack access to capital, digital infrastructure, or advisory support necessary to translate geospatial insights into management actions (Benos *et al.,* 2021). Addressing these constraints requires participatory technology design, policy support, and capacity-building programs tailored to regional farming systems.

Conclusion:

Remote sensing and UAV technologies now form the observational backbone of smart farming systems. By providing multi-scale, repeatable measurements of crop and soil conditions and by integrating these data with GIS, IoT networks, and machine-learning analytics, they enable more precise input management, early stress detection, and climate-resilient production strategies. Continued advances in sensor fusion, automation, and inclusive deployment models are expected to further embed these technologies within sustainable agricultural intensification efforts, particularly in data-rich yet resource-constrained regions such as India.

References:

1. Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sensing*, 5(2), 949–981.
2. Bah, M. D., Hafiane, A., & Canals, R. (2018). Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. *Remote Sensing*, 10(11), 1690.

3. Basso, B., & Antle, J. (2020). Digital agriculture to design sustainable agricultural systems. *Nature Sustainability*, 3, 254–256.
4. Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors*, 21(11), 3758.
5. Chakraborty, S., Newton, A. C., & Sharma, M. (2016). Digital agriculture and climate adaptation in developing countries: Perspectives from India. *Agricultural Systems*, 148, 72–82.
6. Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment. *Environmental Research Letters*, 13(11), 114003.
7. Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant species classification using deep convolutional neural networks. *Biosystems Engineering*, 151, 72–80.
8. García, L., Parra, L., Jimenez, J. M., Lloret, J., & Lorenz, P. (2020). IoT-based smart irrigation systems: An overview on the recent trends. *Computers and Electronics in Agriculture*, 165, 104812.
9. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27.
10. Jones, H. G., Serraj, R., Loveys, B. R., Xiong, L., Wheaton, A., & Price, A. H. (2004). Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress. *Functional Plant Biology*, 31(7), 689–703.
11. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90.
12. Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in Plant Science*, 10, 621.
13. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.
14. Mahlein, A.-K., Oerke, E.-C., Steiner, U., & Dehne, H.-W. (2013). Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology*, 133, 197–209.
15. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419.
16. Mulla, D. J. (2013). Twenty-five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114(4), 358–371.
17. Ojha, T., Misra, S., & Raghuwanshi, N. S. (2015). Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges. *Computers and Electronics in Agriculture*, 118, 66–84.

18. Panigrahy, S., Ray, S. S., Parihar, J. S., & Singh, A. K. (2019). Satellite-based crop monitoring systems in India: Applications and future prospects. *Current Science*, 116(12), 2011–2022.
19. Reddy, G. P. O., Kumar, N., & Rao, P. V. N. (2020). UAV-based multispectral data for nitrogen stress detection in rice ecosystems. *Journal of the Indian Society of Remote Sensing*, 48(5), 689–703.
20. Roy, P. S., Roy, A., Joshi, P. K., Kale, M. P., Srivastava, V. K., Srivastava, S. K., … Ramachandran, R. M. (2014). Development of decadal land use and land cover database for India. *Current Science*, 107(3), 401–410.
21. Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K.-R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3), 247–278.
22. Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2019). Machine learning applications for precision agriculture: A review. *Computers and Electronics in Agriculture*, 162, 678–692.
23. Singh, R., & Gupta, R. (2021). Drone technology for precision agriculture in India: Opportunities and constraints. *Indian Journal of Agricultural Sciences*, 91(10), 1421–1429.
24. Tsouros, D. C., Bibi, S., & Sarigiannidis, P. G. (2019). A review on UAV-based applications for precision agriculture. *Information*, 10(11), 349.
25. Villa-Henriksen, A., Edwards, G. T. C., Pesonen, L. A., Green, O., & Sørensen, C. A. G. (2020). Internet of Things in arable farming: Implementation, applications, challenges and potential. *Biosystems Engineering*, 191, 60–84.
26. Viscarra Rossel, R. A., Adamchuk, V. I., Sudduth, K. A., McKenzie, N. J., & Lobsey, C. (2016). Proximal soil sensing: An effective approach for soil measurements in space and time. *Advances in Agronomy*, 135, 243–291.
27. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming – A review. *Agricultural Systems*, 153, 69–80.
28. Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: A review. *Precision Agriculture*, 13(6), 693–712.

RENEWABLE ENERGY DEVELOPMENT POTENTIAL AND GOVERNMENT INITIATIVES IN INDIA

Ritu Raj Kaur*, Gursharan Kaur and Sakshi Sahni
GRD School of Planning, Guru Nanak Dev Univerity, Amritsar
*Corresponding author E-mail: rituraj.plan@gndu.ac.in

Abstract:
Rapid growth in energy demand, whether renewable or non-renewable, has driven India to increasingly adopt renewable energy as a key component of its national energy strategy. Due to the limited availability of non-renewable energy sources and the ever-increasing demand for energy, worldwide emphasis is being placed on the development and use of renewable energy sources. The Government of India has implemented a range of fiscal and non-fiscal incentives to promote renewable energy technologies. This paper highlights the role of government incentives in promoting renewable energy development and the potential of renewable energy development in India.

Keywords: Renewable Energy, Solar Energy, Small Hydro Power, Wind Energy, Biomass Energy, Geothermal Energy, Ocean Energy.

1. Introduction:

Since the beginning of human civilisation, renewable energy has been the primary form of energy on which humans relied after the discovery of oil in the US in 1859. Before this, only 5% of global energy was supplied by fossil fuels, but after this, the fossil fuel dependency increased, and by 1950, almost 93% of energy supplied was fossil fuel-based. After the 1973 energy crisis, the global energy crisis, the search for alternative energy started, and the focus was given to nuclear energy, hydroelectric power and renewables

India also began installing experimental solar PV systems and windmills in the late 1970's and early 1980s. The Department of Science and Technology was a pioneer department that initially managed solar and wind programmes. Since 2006, the Ministry of New and Renewable Energy Sources has been the leading ministry responsible for renewable energy sources, with initiatives at the central level. Ever since, the renewable energy capacity has grown rapidly. Thus, highlighting renewable energy as a viable and sustainable alternative.

Renewable energy refers to any form of energy that is not depleted by use over time and can be replenished within a very short period, for instance, hours, days, weeks, or months. Different renewable energy sources include solar, wind, hydro, and biomass. For example, solar energy can be utilised within hours, while other renewable sources may take days or weeks. This paper provides details on various forms of energy, their origins, and the time required to use them. It also includes the potential for electricity from renewable energy at the national level, encompassing various sources. It further highlights the government initiatives to encourage renewable energy generation.

2. Renewable Energy: Definition and Origin

Energy forms the backbone of modern society as almost every human activity depends upon its availability. Over time, energy has emerged as an important commodity as any disruption or uncertainty in its supply can destabilise the entire economy. Consequently, patterns of energy production and consumption play a decisive role in achieving rapid and sustainable economic growth, since almost all developmental needs are energy-intensive. Within the broader energy sector, the power sector occupies a central position, supporting growth across domestic, transportation, industrial, commercial, and agricultural sectors in both urban and rural regions. Electricity generation relies on a mix of renewable and non-renewable energy resources.

Renewable energy has been defined in various ways by international organisations. The National Renewable Energy Laboratory (NREL) describes '*renewable energy as energy derived from natural sources such as sunlight, wind, water, geothermal heat, and biomass, which are continuously replenished by natural processes. Technologies based on these resources convert them into usable energy forms, primarily electricity, but also heat, mechanical power, and fuels'* (NREL, U.S. DoE, 2001). Similarly, the World Energy Council defines *'renewable energy as energy obtained from sources that do not diminish with use and can be restored within a relatively short period'* (WEC, 2016). Thus, renewable energy sources are primary energy resources that regenerate naturally over short timeframes, ranging from hours and days to weeks and months. Thus, making its availability continuous and sustainable.

The Sun is the fundamental source from which most renewable energy sources originate. Solar energy can be harnessed directly and almost instantaneously. Wind energy arises from the uneven heating of the Earth's surface by solar radiation. Similarly, solar energy drives the hydrological cycle through evaporation and transpiration. It leads to river flows that enable the generation of hydroelectric power. Similarly, plants capture solar energy through photosynthesis and indirect solar energy is stored as biomass. Similarly, when organic matter remains buried under specific geological and climatic conditions for millions of years, it is eventually transformed into fossil fuels such as oil and natural gas, which are again used as sources of energy.

It has been projected that renewable energy will play a significant role in meeting global energy demand, especially in developing countries, by 2060. As energy demand increases every year, a growing share of this demand is expected to be met by renewable sources. To assess the renewable energy potential, the following sections include the potential and capacities of different sources.

3. Renewable Energy Potential at the National Level

The sharp rise in energy demand and increasing gaps between energy supply and consumption have prompted governments to focus on expanding energy generation. Ongoing reforms in the power sector, increasing electricity requirements, and growing environmental concerns have collectively accelerated the adoption of renewable energy technologies. In this light prefrence is given to renewable sources. These efforts aim not only to enhance energy availability but also to minimise the environmental damage associated with conventional fossil-fuel-based power generation.

Renewable Energy: Sources, Assessed Potential, and Utilised Capacity

India possesses a vast estimated renewable energy potential of approximately 895 GW. at present, only a small fraction, around 4 per cent, has been harnessed. Advances in engineering and energy technologies are expected to significantly improve the utilisation of renewable resources. Among renewable energy sources, solar energy has the highest potential. It is projected that with continued technological improvements, nearly 84 per cent of the total renewable energy potential could be generated from solar power alone (PwC, 2015).

At present, wind energy accounts for nearly two-thirds of total renewable power generation, making it the dominant source of installed renewable capacity. This has positioned India as the fifth-largest wind energy producer globally. Other renewable sources such as small hydropower, biomass, and waste-to-energy have also made notable contributions, though their exploited capacity remains significantly below their assessed potential.

A range of renewable energy resources can be further developed for national-scale electricity generation. These include solar and wind energy, biomass and bioenergy, small hydropower, geothermal energy, ocean-based energy sources (such as tidal, wave, and ocean thermal energy), and hydrogen energy and fuel cell technologies. Strategic investment and policy support in these areas can substantially strengthen the country's energy security while promoting sustainable development.

Solar Energy

The sun is the primary source of energy that sustains life on Earth. It is a clean and inexhaustible form of energy. It is a widely available source. India receives an estimated over 5,000 trillion kWh of solar energy annually (Khan, 2017). The theoretical solar energy potential available for capture is approximately 5 kWh per square meter per day.

Despite its abundance, solar energy availability varies seasonally. Large-scale deployment faces challenges of high initial capital costs and land requirements. However, compared to wind energy, solar power is relatively more predictable. Solar energy can be harnessed primarily through two technologies, including solar thermal systems and solar photovoltaic (PV) systems, discussed as following.

Solar Thermal Systems:

Solar thermal technology captures solar radiation as heat and utilises it for various applications. These include domestic water heating, industrial process heating, drying of agricultural and industrial products, and space heating. The use of solar thermal systems reduces reliance on conventional heating fuels. In addition, solar thermal energy can be used for electricity generation with technologies similar to those of conventional thermal power plants.

In India, the government has actively promoted solar thermal applications, particularly for water heating. This is done by incorporating supportive provisions in municipal building byelaws. Cities such as Bengaluru and Thane have been designated as solar thermal cities. MNRE has also introduced Solar Master Plans for around 60 cities and towns under the Solar Cities Programme. These plans involve assessing city-level solar potential by conducting baseline studies. The primary objective of this initiative is to encourage decentralised renewable energy generation. The aim is to reduce conventional energy demand in urban areas by approximately 20 per cent. This approach draws inspiration from international models implemented in cities such as New York, Tokyo, and London.

India has also planned the development of 25 solar parks dedicated to concentrated solar power (CSP) generation. One notable example is the Charanka Solar Park in the Patan district of Gujarat, developed as a public–private partnership. The park spans about 2,024 hectares, with land contributions from both the Government of Gujarat and private entities, and has an approximate installed capacity of 500 MW.

Solar Photovoltaic (PV) Systems

Solar photovoltaic technology using PV cells converts solar radiation directly into electricity. These solar cells were first developed in 1954 for powering space satellites. Their terrestrial applications expanded significantly following the 1973 global oil crisis, which highlighted the need for alternative energy sources. Today, solar PV systems are increasingly deployed across a wide range of applications. It is used in residential areas and street lighting. It is also being used in transportation, irrigation, and even for the electrification of remote and off-grid areas. As technology continues to evolve, solar photovoltaic energy is expected to play an increasingly vital role in meeting sustainable energy goals.

Wind Energy:

Wind energy is the kinetic energy of moving air. In this mechanical energy of wind can be converted into electricity using wind turbines. Effective power generation requires moderate to high wind speeds, as very low or stagnant winds are unsuitable for electricity production. Typically, wind turbines operate efficiently within a wind speed range of 5 to 25 meters per second. The global potential for large-scale, grid-connected wind power generation has been estimated at around 100,000 MW (Khan, 2017).

The utilisation of wind energy dates back to 2000–1700 BC, when windmills were employed in regions such as Babylon and China for tasks like water pumping and grain grinding. Wind energy was also historically used for maritime navigation through sailing ships. The concept of generating electricity from wind was first introduced in Denmark in 1890. However, significant development occurred after the 1973 oil crisis, which renewed global interest in alternative energy sources. Technological advancements since the 1980s have enabled the deployment of large-scale wind power systems worldwide, with installations ranging from a few kilowatts to several megawatts.

Improvements in wind turbine technology have substantially reduced the cost of wind-generated electricity from approximately ₹17 per kWh in 1980 to about ₹2.50 per kWh in recent years. Wind energy projects are developed in both onshore and offshore locations. Capacities range from small-scale microturbines (around 225 kW) used at the local or building level to large wind farms exceeding 1,500 MW at the regional scale. In India, notable wind power installations include the Muppandal wind farm in Tamil Nadu (1,500 MW), the Greenshore wind farm in Gujarat (600 MW), and several other wind farms across Tamil Nadu with capacities of around 500 MW.

Biomass Energy:

Biomass energy is produced using organic materials, primarily agricultural residues and waste, as feedstock. Broadly, biomass includes all living or recently living matter, such as plants and organic waste. It has been estimated that only one-eighth of the biomass produced annually could meet the current global energy demand (Khan, 2017). Like most renewable resources, biomass energy ultimately originates from the sun, as plants store solar energy through photosynthesis. Residues from plant growth are then utilised for energy production.

Biomass can be converted into multiple forms of energy and by-products. These include electricity, biofuels (ethanol and biodiesel), charcoal, and biogas. The use of agricultural residues and organic waste from different sources makes biomass a renewable and sustainable energy option.

Biomass is one of the oldest energy sources. Traditionally, it was used for cooking and heating. In modern times, technologies such as waste-to-energy (WTE) plants and cogeneration systems have enabled efficient electricity generation from biomass. Urban solid waste also presents significant potential for WTE projects.

Several successful biomass initiatives exist in India. For instance, the Chennai Corporation has implemented a community-level biomethanation plant capable of processing 2 tonnes of waste per day, generating gas equivalent to four cylinders or about 200 units of electricity, along with 180 kg of organic manure daily. In Pune, an industrial WTE facility processes 72 tonnes per day of hazardous waste using plasma technology, producing approximately 1.6 MW of electricity through gasification. Additionally, rice mills and sugar industries across states such as Punjab and Maharashtra utilise agricultural residues for captive power generation. A sugar mill in Maharashtra, for example, generates 30 MW of electricity through a bagasse-based cogeneration project.

Geothermal Energy

Geothermal energy is the thermal energy stored beneath the Earth's surface. It is accessible in regions with volcanic activity, geysers, or hot water springs. Although geothermal energy has significant potential, its exploitation is geographically limited to specific locations. This constraint, along with the requirement for advanced, expensive technologies, has restricted the widespread development of geothermal power.

Geothermal resources with temperatures exceeding 90°C are suitable for electricity generation. The total global installed capacity is approximately 9,031 MW, with an annual growth rate of about 3 per cent (Khan, 2017). The United States is the world's leading producer of geothermal energy. A major advantage of geothermal energy is its continuous availability, as it operates independently of weather conditions.

Ocean Energy:

Ocean energy involves converting the mechanical energy of tides and waves. It also includes the thermal energy from temperature differences in ocean water, into electricity. Oceans cover nearly 71 per cent of the Earth's surface, and continuous wave and tidal movements generated by ocean winds represent a vast energy resource. Tidal energy is the most developed form of ocean energy. While wave energy and ocean thermal energy conversion (OTEC) remain largely underdeveloped due to challenges related to transmission, maintenance, and utilisation in marine environments. Coastal regions offer significant potential sites for tidal energy projects.

Small Hydro Resources:

Small hydropower utilises the mechanical energy of flowing or falling water. The energy drives turbines and generates electricity. It is one of the most mature and cost-effective renewable energy technologies. While large hydropower projects are well established, small hydropower projects (SHPs) have gained attention for their lower environmental impact and suitability for decentralised power generation.

Definitions of SHPs vary internationally, with capacity limits ranging from 5 MW to 50 MW. In India, the capacities of small hydropower projects are up to 25 MW. The MNRE has identified 6,474 potential SHP sites along rivers and canals, with a cumulative estimated capacity of approximately 19,749 MW.

Hydrogen Energy:

Hydrogen is considered a clean and renewable energy carrier. It can be produced through processes such as the electrolysis or thermolysis of water. In hydrogen-based energy systems, the electrochemical energy of hydrogen is converted directly into electricity using fuel cells. The only by-product is water. This makes hydrogen energy a non-polluting and environmentally friendly option.

However, the widespread adoption of hydrogen energy faces several challenges. These include limited production infrastructure, complex and costly storage requirements, and transportation and distribution challenges. These factors currently make hydrogen energy an expensive option, although continued research and technological advancements may improve its feasibility in the future.

4. Government Incentives and Challenges in Renewable Energy Development

The Government of India has introduced a wide range of fiscal and non-fiscal policy measures to promote the growth of renewable energy. These measures aim to improve the financial viability of renewable energy projects. Different renewable energy sources are supported through technology-specific incentive frameworks. Common incentives include higher feed-in tariffs, tax concessions, and direct or indirect subsidies. However, the expansion of renewable energy remains constrained by challenges such as high initial capital costs, inadequate infrastructure, intermittent resource availability, and lengthy approval processes. A summary of key incentives and challenges associated with major renewable energy sources is presented below.

Solar Energy

Solar energy development is strongly supported under the National Solar Mission, implemented by the Ministry of New and Renewable Energy (MNRE) in coordination with State Renewable Energy Agencies. Incentives include attractive feed-in tariffs for grid-connected projects, typically ranging from ₹10–14 per kWh for concentrated solar power (CSP) plants and ₹12–18 per kWh for photovoltaic (PV) systems. Despite these incentives, solar power projects face significant barriers, including high upfront costs, limited availability of urban land for large installations, evolving, capital-intensive technologies, and the intermittent nature of solar radiation.

Wind Energy:

Feed-in tariffs generally range from ₹3.39 to above ₹5 per kWh. Generation-based incentives and accelerated depreciation benefits are also available for wind energy. These measures have contributed to rapid capacity addition in wind-rich regions. However, challenges persist, including inadequate transmission and distribution infrastructure, variability of wind resources, and the need for wind-zone-specific incentive mechanisms due to significant spatial variations in wind potential.

Biomass and Waste-to-Energy Power:

Biomass and waste-to-energy projects receive multiple fiscal incentives, including exemptions from customs duty, excise duty, and central sales tax, as well as 100 per cent accelerated depreciation. In addition, project developers are eligible for a 10-year income tax holiday, which can be claimed over a 15-year period. Nevertheless, these projects face operational challenges, particularly ensuring a continuous and affordable supply of feedstock, managing fluctuating biomass prices, competing with alternative uses for feedstock, and dealing with intermittent power generation.

Small Hydropower:

Small hydropower projects are supported through capital subsidies and various tax and duty exemptions. Various challenges these projects face include long approval timelines, procedural delays, and limited access to transmission and distribution networks, especially in remote and hilly regions.

Geothermal Energy:

In India, geothermal energy remains largely at an exploratory stage. While the government and a few private entities have initiated preliminary investigations. Still, the sector lacks a clearly defined policy framework and incentive structure. High capital requirements, limited technical expertise, and the absence of proven commercial projects have restricted progress in this area.

Wave Energy:

Wave energy development in India is minimal, with only initial efforts undertaken in selected coastal states such as Gujarat. Key challenges include uncertainty regarding resource potential, lack of mature technologies, high project costs, and the absence of a comprehensive policy and incentive framework.

Tidal Energy:

Tidal energy has attracted limited attention, with some exploratory initiatives supported by the government. However, the sector faces substantial obstacles, including high capital investment requirements, insufficient transmission infrastructure, intermittent power output, and a lack of large-scale demonstration projects.

Conclusion:

Renewable energy is a critical pillar of India's strategy to ensure energy security and long-term economic growth. Besides their unmatched potential, the renewable energy sector continues to face several challenges. Various emerging renewable technologies, such as geothermal, wave, and tidal energy, remain underdeveloped due to limited policy and incentive mechanisms. To sustain and accelerate renewable energy growth, it is essential to address the challenges. It is important to focus on localised power generation, tailored regional incentives, and the combination of different renewable energy sources to improve overall renewable energy development in the country.

References:

1. Energy Alternatives India (EAI). (2011). *The India renewable energy advisor*. Energy Alternatives India. http://www.eai.in/ref/services/wp/india_renewable_energy_advisor_aug_2011.pdf
2. Khan, M. J., *et al.* (2017). Techno-economic feasibility analysis of different combinations of PV-wind-diesel-battery hybrid system for telecommunication applications in different cities of Punjab, India. *Renewable and Sustainable Energy Reviews, 76,* 577–607. https://doi.org/10.1016/j.rser.2017.03.056
3. Ministry of New and Renewable Energy (MNRE). (n.d.). *Renewable energy installed capacity and programs*. Government of India. http://mnre.gov.in/
4. Monga, P. (1997). Institutional barriers to renewable energy development in India. In V. Ramana (Ed.), *Rural and renewable energy: Perspectives from developing countries* (pp. xx–xx). Tata Energy Research Institute.
5. U.S. Department of Energy (DOE), National Renewable Energy Laboratory (NREL). (2001). *Renewable energy: An overview*. U.S. Department of Energy. http://www.nrel.gov/docs/fy01osti/27955.pdf
6. PwC India. (2015). *Renewable energy's transformation of the Indian electricity landscape*. PwC India. https://www.pwc.in/assets/pdfs/publications/2015/renewable-energys-transformation.pdf
7. Raven, P. H., & Berg, L. R. (2008). *Environment*. Wiley.
8. Stockholm International Peace Research Institute (SIPRI). (1974). *Oil and security*. Humanities Press.
9. Sobel, L. A. (Ed.). (1974). *Energy crisis, volume 1: 1969–73*. Facts on File.
10. World Energy Council (WEC). (2016). *World energy scenarios 2016* (in collaboration with Accenture Strategy and Paul Scherrer Institute). World Energy Council. https://www.worldenergy.org/wp-content/uploads/2016/10/World-Energy-Scenarios-2016_Full-Report.pdf

Contemporary Innovations in Science, Engineering and Technology

(ISBN: 978-93-47587-99-3)

About Editors

Dr. Sukanta Das, a native of Tripura, India, born on 25 April 1985, completed graduation in Veterinary Sciences in 2008 from Central Agricultural University, Imphal. He earned his Master's degree in Veterinary Anatomy and Histology in 2010 from GBPUA&T, Pantnagar, Uttarakhand, and obtained a Ph.D. in Veterinary Anatomy in 2018 from Assam Agricultural University, Assam. Presently, he serves as Associate Professor and Head, Department of Veterinary Anatomy, College of Veterinary Sciences and Animal Husbandry, R.K. Nagar, Tripura West. With fifteen years of teaching and research experience, he has published over twenty five research articles, designed farmer training manuals, specializes in histology, histochemistry, and electron microscopy, undertook institutional responsibilities, attended national and international seminars, workshops, brainstorming sessions, and received research presentation awards and a research fellowship.

Dr. Shweta Tanwar, MDS, MPH, is currently serving as a Medical Scientist and Technical Editor at the Indian Journal of Medical Research, ICMR Headquarters. Her subject expertise and research interests include oral health, public health, health systems research, research ethics, peer review, and publication practices. Her professional role involves managing editorial workflows, coordinating peer review, overseeing manuscript quality, content accuracy, and ensuring timely publication of high quality biomedical research. She is actively involved in research practices with interests encompassing oral health, bioethics, and population based health research. She has presented her work at several national and international scientific conferences and serves as an expert reviewer for national and international journals. With a strong background in public health and biomedical research.

Mujahid M. Hussain is an Assistant Professor in the Department of Physics at Maharashtra College of Arts, Science and Commerce, Nagpada, Mumbai 400612. He holds an M.Sc. in Physics with specialization in Nuclear Physics and has qualified NET, SET, JEST, and GATE examinations. He possesses nine years of teaching experience at degree college level, primarily teaching undergraduate students. His area of academic and research interest is Non Linear Dynamics and Chaos Theory. Earlier, he worked with Target Publications, contributing to solving and compiling previous year entrance examination questions for their competitive examination books, thereby supporting student preparation and conceptual understanding in physics through clear explanations and structured problem solving approaches.

Dr. R. Karthick Manoj is an Assistant Professor in the Department of Electrical and Electronics Engineering at AMET Deemed to be University, Chennai. He earned his Ph.D. in Machine Learning from AMET Deemed to be University in 2024, with research focused on artificial intelligence–based disease detection frameworks. His research interests span machine learning, computer vision, smart agriculture, Internet of Things, and medical image analysis. He has published over 20 research articles in SCI, Scopus, and Web of Science indexed journals and more than 25 papers in national and international conferences. He has authored three books and contributed over twelve book chapters with leading international publishers. Dr. Manoj holds four granted patents and two published patents in AI, IoT, and smart systems, has served as Co-Principal Investigator for externally funded projects, and is an active member of professional bodies.

Bhumi Publishing

ISBN 978-93-475-8799-3

9 789347 587993