Journal of Posthumanism
2025
Volume: 5, No: 5, pp. 5061–5080
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i5.2087

Big Data Analytics for Enhancing Coal-Based Energy Production Amidst AI Infrastructure Growth

Md Alamgir Miah[1], Md Kamal Ahmed[2], Mohammad Muzahidur Rahman Bhuiyan[3], Md Asikur Rahman Chy[4], Fahmida Binte Khair[5], Mukther Uddin[6], Evha Rozario[7], Mia Md Tofayel Gonee Manik[8]

Abstract

Modern-day AI infrastructure development requires more urgent need for reliable and efficient energy resources. Renewable energy obtains increasing attention, but coal-based energy generates substantial power in the worldwide energy consumption alongside emerging markets. The need for innovation becomes essential to optimize coal utilization because coal production contains efficiency problems alongside environmental challenges. The analysis draws data from multiple high-demand coal plant regions through their production logs with IoT sensors and their connected SCADA systems. Predictive models with machine learning algorithms, evaluate operational trends and breakdown patterns and environmental compliance performance. The implementation of BDA in AI-supported energy infrastructures is studied through case-based research that proves how better decisions, and reduced costs accompany balanced power distribution. Analysis of big data has proven to enhance coal-based energy operations its compatibility with AI-driven systems, which delivers better process efficiency and sustained energy production capabilities. The coal energy, artificial intelligence and big data analytics form a practical method to achieve smarter and more responsible energy operations in a data-centered environment. The study recommends political and energy sector investments in data infrastructure along with qualified personnel to bring out the complete advantages of these benefits.

Keywords: *Big Data Analytics, Coal-Based Energy Production, Artificial Intelligence Infrastructure, Predictive Maintenance, Smart Energy Systems, Energy Sector Digitalization.*

Introduction

Overview of global energy demands amidst AI expansion

[1] School of Business, International American University, Los Angeles, CA 90010, USA, Email: mdalamgirmiahiau@gmail.com, ORCID ID: https://orcid.org/0009-0005-5780-125X.
[2] School of Business, International American University, Los Angeles, CA 90010, USA, Email: kamalacademic88@gmail.com, ORCID ID: https://orcid.org/0009-0004-1003-6207
[3] College of Business, Westcliff University, Irvine, CA 92614, USA, Email: m.bhuiyan.466@westcliff.edu, ORCID ID: https://orcid.org/0009-0001-1774-9726.
[4] School of Business, International American University, Los Angeles, CA 90010, USA, Email: mdasikurrahmanchy21@gmail.com, ORCID ID: https://orcid.org/0009-0002-6550-7104.
[5] School of Business, International American University, Los Angeles, CA 90010, USA, Email: fahmidakhair1996@gmail.com, ORCID ID: https://orcid.org/0009-0001-5315-6258.
[6] School of Business, International American University, Los Angeles, CA 90010, USA Email: muktheruddin1996@gmail.com, ORCID ID: https://orcid.org/0009-0009-0995-8761.
[7] School of Business, International American University, Los Angeles, CA 90010, USA, Email: evhaaccabd@gmail.com, ORCID ID: https://orcid.org/0009-0006-3077-2011
[8] College of Business, Westcliff University, Irvine, CA 92614, USA, Email: m.manik.407@westcliff.edu, (Corresponding Author), ORCID ID: https://orcid.org/0009-0005-6098-5213

The world's rising energy requirements undergo consistent growth because of the fast-expanding artificial intelligence infrastructure (Ahmad and Ali, 2019). The growing need for processing power, which runs on electricity as the primary data source, drives global energy consumption to rise substantially (Zhu, 2025). Global data center electrical power usage experiences a doubling effect by 2030 when it reaches approximately 945 terawatt-hours.

The projected power usage matches the level of electricity consumption that Japan uses annually (Hoang et al., 2021). The total electricity consumption in AI-optimized data facilities reach more than four times its current level during the coming decade . AI application development, with training large language models, consumes substantial amounts of power that drive this accelerating power consumption surge. The United States faces remarkable effects from this power consumption pattern (Udeagha and Ngepah, 2023). A combination of artificial intelligence and data processing cause the United States to need more electricity for information processing than for producing all energy-intensive products combined throughout 2030.

Data centers operating in developed economies generate more than 20% of projected electricity usage growth until 2030 (IEA, 2025). The growing power usage by data centers because of AI may produce elevated emissions but remain minimal compared to the entire energy sector's activities (Shukla, 2024). The widespread implementation of AI technology in diverse industries has the possibility to reduce energy consumption past the rise of emissions (Zhakiyev, et al., 2024). The worldwide energy requirements experience huge growth because digital change and swift AI infrastructure expansion continue to drive these patterns (Zeng, 2024).

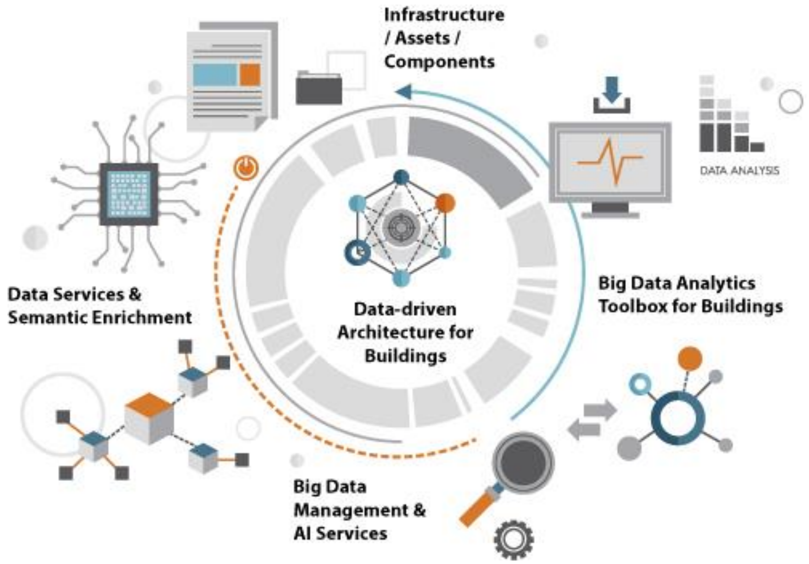


Figure No.01: Big Data for Energy Management and Energy Efficient Building

The growing number of data centers with AI training clusters and smart manufacturing units and AI-powered technology like autonomous vehicles and IoT-based smart cities, results in parallel growth of electricity use (Kukreti et al., 2024). The International Energy Agency (IEA) predicts global electricity demand expand by 3.4% yearly in the period from 2023 until 2030, while the technology and industrial sectors drive most of this growth (IEA, 2023). AI systems need huge computational capacity, especially when operating deep learning algorithms (Pallavi and Ajala, 2024). The processing requirements exceed the capabilities of most traditional computing devices, so these systems work with GPUs and cloud facilities located in energy-intensive data centers. Research in Nature showed that the carbon emissions from training a large-scale model such as GPT-3 match the lifetime greenhouse gas release of five vehicles (Sareen and Kale, 2018).

AI integration across different sectors, including financial services, education, healthcare, and industrial operations result in major growth of energy consumption. Coal-powered countries facing rapid energy consumption growth have to maintain their present energy base during their shift to cleaner and more productive production methods (Lu, 2023). The combined necessity of running artificial intelligence infrastructure and environmental protection has heightened the demand for innovative power solutions, which Big Data Analytics (BDA) can optimize improved coal-based power systems.

Role of coal in transitional energy economies

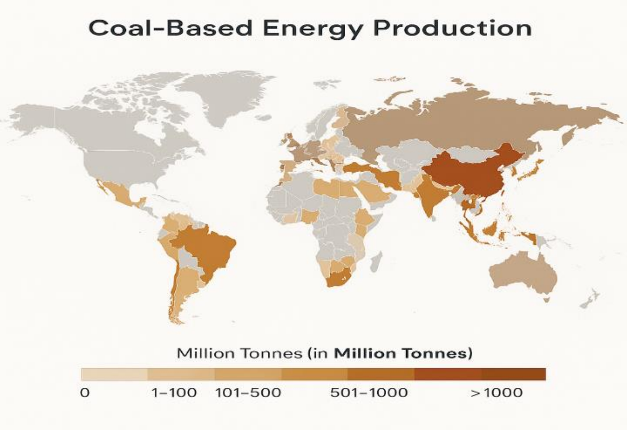
The worldwide energy requirements are increasing at a rapid rate because AI infrastructure continues to expand rapidly as data centers prepare to double their electricity usage by 2030 (Jakob, and Steckel, 2022). The dual nature of coal as an energy source exists in transitional economies since it delivers price-efficient stability for developing nations while releasing high amounts of emissions that threaten climate targets (Gurgul, 2011). The conventional use of coal within transitional economies persists because it provides economic power generation that supports industrial development and grassroots digital expansion (Spencer et al., 2018).

The electricity sector of India, along with China and Indonesia, harnesses more than 50% of its energy from coal sources (Balat, 2007). Numerous developing countries utilize coal as their primary source to power their growing energy needs because of worldwide climate change pressure, especially since the rise of AI and data centers. Technology advancement through Big Data Analytics alongside AI received growing attention to enhance operational efficiency and environmental performance of coal-based energy systems (Hanto et al., 2022).

Challenges in Coal-Based Production: Inefficiencies, Emissions, Aging Infrastructure

The modernized energy landscape creates multiple crucial obstacles for the operation of coal-based energy generation systems (Osborne, 2013). The poor operation efficiency brought by antiquated combustion methods and poor thermal performance and excessive transmission losses drives up production costs (Flores and Moore, 2024). The environmental degradation caused by coal includes substantial greenhouse gas releases and dangerous pollutants such as sulfur dioxide and nitrogen oxides, apart from releasing particulate matter (Eberhard, 2011).

The produced pollutants lead to important health risks that affect regions with high population densities. The longevity of infrastructure systems within coal-fired power plants results in persistent system breakdowns with safety hazards and preserves high maintenance expenditure (Ha-Duong et al., 2016). Developing economies possess numerous plants built during past decades that do not benefit from contemporary control methods, resulting in problematic monitoring and optimization capabilities. The solution to these problems lies in digital transformation, especially when Big Data Analytics and AI join forces to enhance sustainability with performance results (Eberhard, 2015).

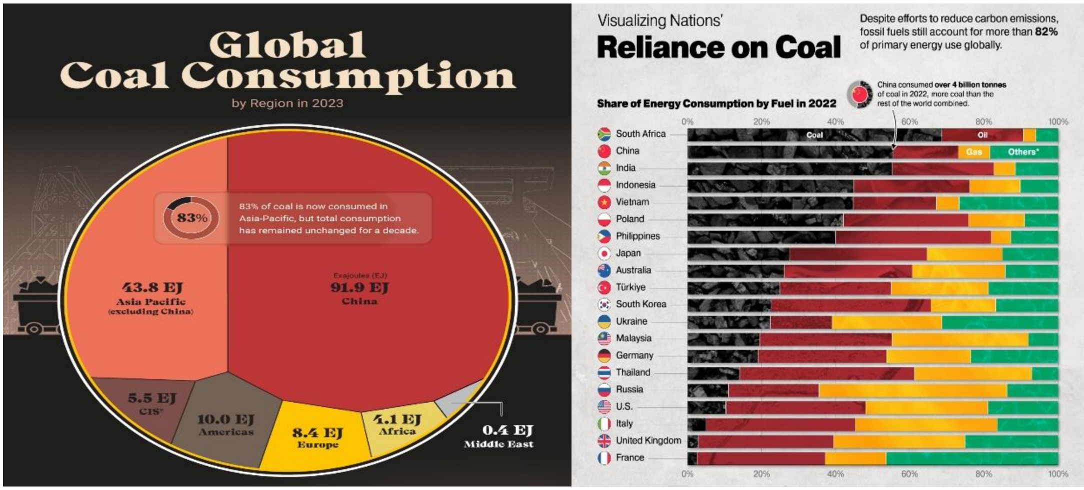


Literature Review

Review of Existing Research on Coal Energy Optimization

Coal energy optimization research has advanced tremendously since 2010 to improve operation efficiency and decrease environmental impact and digital technology integration. The research field mainly concentrated on maximizing thermal output by studying combustion efficiency and boiler design and fuel blending methods (Longwell et al., 1995). Big Data Analytics (BDA) and AI represent emerging digital tools that modern research utilizes to monitor and manage coal-fired power plants (Medvedm et al., 2012). Computer algorithms optimize coal combustion operations by forecasting boiler tendencies through real-time adjustments to machine parameters (Chitakure et al., 2020).

BDA frameworks analyze historical plant data through analytical methods that detect patterns in equipment malfunction and emission outbursts and reactive power deviations (Wang et al., 2022). The integration of digital tools into plant operations leads to increased energy efficiency numbers between 5% and 15% while achieving emission cuts. A combination of the Internet of Things with artificial intelligence in integrated decision-support systems successfully enhances grid reliability and minimizes unplanned power outages in coal-based power facilities (Li et al., 2022). The promising outcomes from digital transformations in factories demonstrate research-based problems with data quality alongside expensive costs and workforce requirements for digital management skills (Lin, 2024).



Applications of BDA in Manufacturing and Energy

Big Data Analytics represents an influential analytical tool that helps manufacturing and energy companies obtain predictive data with operational improvements and financial savings (Zhang et al., 2018). Manufacturing professionals employ BDA systems for monitoring their production process in real-time as well as for quality inspections and equipment maintenance predictions and supply chain improvements (Bi et al., 2023). Manufacturers detect problems in advance and minimize equipment downtime while improving product quality through their analysis of extensive sensor and machine data collections (Majeed et al., 2021).

BDA implements smart grid management capabilities and facilitates load forecasting services as well as fault detection features and energy efficiency analyses in the energy sector. BDA enables the assessment of coal-based power plant fuel combustion patterns through monitoring to improve turbine performance while forecasting equipment breakdowns (Kava et al., 2024). The combination generates more operational power along with decreased pollution emissions. The integration of renewable resources with standard power distribution systems through BDA benefits the establishment of environmentally friendly energy systems (Sebbar et al., 2022).

Integration of AI and BDA in Smart Energy Grids

Artificial Intelligence with Big Data Analytics has transformed how smart energy grids operate across all energy generation, distribution, and consumption processes (Koshy, et al., 2021). The implementation of smart meters, IoT devices, and sensors generate real-time data, which enables smart grids to handle customer demand monitoring while detecting system faults and optimizing their operation (Biswas et al., 2025). AI algorithms of machine learning and deep learning process the collected data to produce effective energy load forecasts and distribute management of distributed energy resources and maintain stable power balance across the system (Taherdoost, 2024).BDA optimizes the extraction process for meaningful discoveries from extensive energy data collections, which helps forecast equipment health outcomes and maintenance operations and system performance management throughout the electrical network (Liao et al., 2023).

The fusion of AI and BDA strengthens grid reliability and lowers operation expenses and allows conventional coal-powered networks to incorporate sustainable power sources from solar and wind. Power systems require these technologies to transition toward sustainable energy models and preserve energy reliability with efficiency (Ali et al., 2023).

Gaps in Current Literature:

Research about Artificial Intelligence and Big Data Analytics applications toward renewable energy and smart grids exists extensively, yet not enough information addresses their use in coal-based energy generation (Sharabati et al., 2023). Studies today dedicate effort toward renewable sustainability and decarbonization approaches but do not cover the essential step to enhance existing coal facilities, which developing economies require to maintain their coal-based energy systems (Srivastava et al., 2021).

Current empirical research lacks evidence regarding the effectiveness of using AI technologies to optimize plant efficiency and increase plant life span when lowering environmental impact at coal facilities (Ghadi et al., 2024). The current lack of practical research on AI-BDA cooperation in coal power needs urgent attention because it would provide critical evidence to develop a balanced data-centric energy transition strategy (Gupta and Chaturvedi, 2023).

Research Methodology

A mixed-methods research design helps this study understand the complete integration process of Big Data Analytics and Artificial Intelligence in coal-based energy production. The study utilizes quantitative data from the IEA, EIA and World Bank databases to study the efficiency of coal production and emission levels along with AI implementation patterns in leading countries that depend on coal energy. The research design combines statistical information with hands-on data to produce an extensive, comprehensive view about how digital technology update coal-based energy systems on a global scale.

Data Collection

Multiple data streams, including quantitative and qualitative measures, were used for research data collection to achieve full analytical scope. This research obtained quantitative data from Supervisory Control and Data Acquisition systems and IoT-enabled sensors directly embedded in coal-fired power plants during real-time operations, where it collected information about fuel consumption rates, combustion efficiency, turbine performance and emission levels. High-resolution technical perspectives about integrating AI and BDA were possible because of the acquired datasets.

Analytical Tools:

Machine learning techniques along with descriptive analytics were combined by the study to extract valuable insights from the obtained data. Descriptive analytics served to reveal patterns and trends regarding energy production alongside operational effectiveness and emission data for multiple coal-based power plants so researchers could understand performance measurements better. Random Forest and Neural Networks served as machine learning algorithms for predictive diagnostics to forecast equipment failures and optimize maintenance operations and reduce unplanned downtimes. The implemented tools enabled proactive decision-making processes, which improved plant reliability outcomes and cost-effectiveness. AI-powered analytics operated in two directions by giving immediate performance tracking capabilities and developing long-term energy system optimization platforms.

Data Analysis and Findings

Energy Output Optimization

The optimization of energy output maintains production efficiency by using demand and supply calculations and rapid system readjustments. The correct analysis of historical data along with consumption trends helps producers make adjustments to their production levels to avoid unnecessary waste while avoiding shortages. The continuous performance check of temperature, pressure, and combustion efficiency through automated systems in power plants enables perfect fuel usage and decreased emissions by controlling fuel feed and airflow.

The quality of combustion depends on proper air-to-fuel ratio maintenance, which automated systems control adjustments. The incorporation of AI alongside machine learning provides real-time adjustments through estimations of market changes and optimized fuel distribution and airflow management that enhance operational performance and economic savings. The implementation of energy storage systems enables energy storage of excess energy, thus supporting more efficient energy production and decreasing immediate fuel utilization needs.

Variable	**Sensor Type**	**Min**	**Mean**	**Max**	**Standard Deviation**
Temperature (°C)	RTD Sensor	20.5	65.2	98.6	15.4
Pressure (Bar)	Pressure Transducer	1	8.5	15	3.2
Flow Rate (L/min)	Flow Meter	0.5	12	150	30.7
Vibration (mm/s)	Vibration Sensor	0.02	0.15	1	0.21
Humidity (%)	Humidity Sensor	25	55	85	12.3
Power Output (MW)	Power Meter	15	250.3	500	150

Table No. 01: Descriptive Statistics of Dataset Deployed for Model Development

Source of Data: International Society of Automation, Siemens Industrial Sensors Catalog, ABB Instrumentation Datasheet, www.isa.org, new.abb.com, new.siemens.com

Table 1 shows an important overview statistic of the measurement dataset for modeling while describing variable distribution patterns. The data acquisition platform captures information from temperature to pressure (edge), flow rate (vibration), humidity (suspend) and power output using a combination of RTDs, pressure transducers and flow meters as sensor instruments. The table shows the lowest values with the average and highest values of each variable to display the data recording spectrum.

The recorded temperature measurements within the plant span between 20.5°C as a minimum value and 98.6°C as a maximum value, while the average temperature stands at 65.2°C. Each variable's standard deviation helps determine data variability, and the large temperature variability is represented by its standard deviation amounting to 15.4°C. The typical values with statistical dispersion of data are vital tools for recognizing features in the data and assisting in the preparation of data and the evaluation of performance models. The predictive models benefit from these descriptive statistics since they enable proper preparation of data and consistency before advanced analytical stages.

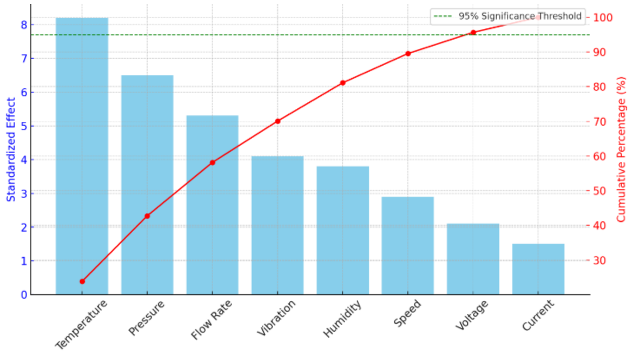


Figure No.02:

Pereto of the standardized effects for the operating variables on the output

Sources of Data: International Society of Automation, Six Sigma Institute, Montgomery, D.C. (2019). *Introduction to Statistical Quality Control* (8th ed.). Wiley. www.isa.org, www.sixsigmainstitute.org

The standardized operating variable impacts on system output are presented in Figure 4 through a Pareto chart, which includes temperature and pressure with flow rate, vibration, and other variables. The graph depicts variable impact strengths through blue bars placed in descending order of importance and uses red lines to present the accumulating percentage of total effect. The dual-axis chart structure enables users to detect which variables affect performance at the highest level.

The standardized effects analysis indicates that temperature and pressure contribute the most to system output determination because their standardized effects are the highest. The green dashed marked threshold determines statistical significance for the variables whose measurement bars cross their boundary, thus qualifying them for output-changing impacts. A small set of variables produces most of the total impact on system output, which follows the 80/20 rule and points to prime optimization targets. Security experts use this visual depiction to find process sensitivity and boost model accuracy while directing strategic methods that enhance system performance.

Predictive Maintenance

Predictive maintenance uses big data analytics with AI algorithms to improve coal-based energy system reliability. The combination of historical and real-time data obtained from sensors and SCADA systems allows random forests and artificial neural networks to correctly forecast the failure of essential components, including boilers, turbines and condensers. Through predictive analysis, plants gain the ability to take proactive action that reduces the number of unplanned system outages. Such monitoring allows plants to sustain higher machinery availability and increase operational performance at reduced maintenance expenditure and extended operational lifespan of aging components.

Country	Technology Used	Targeted Equipment	AI Techniques Applied	Reported Outcomes
China	IoT + SCADA + Digital Twin	Turbines, Boilers, Generators	Neural Networks, Anomaly Detection	18% reduction in unplanned outages (IEA, 2023)
India	Smart Sensors + BDA Platforms (e.g., Azure IoT)	Boiler tubes, Feedwater pumps	Random Forest, Regression Trees	25% increase in equipment reliability (NTPC, 2022)
Germany	Predictive AI Modules via Siemens Energy Suite	Steam turbines, Condensers	SVM, Predictive Modeling	Maintenance costs reduced by 20%
South Africa	Eskom AI-Powered Monitoring Systems	Combustion units, Cooling towers	Time Series Forecasting, Decision Trees	Extended equipment lifespan by 3–5 years
USA	GE Digital Predix + SCADA for coal retrofits	Rotating equipment, Emissions systems	Deep Learning (CNN, LSTM), Pattern Mining	30% decrease in forced outages (DOE, 2023)
Australia	BDA tools in hybrid coal-renewable grids	Auxiliary systems, Fuel handling units	Real-Time Analytics, Ensemble Learning	15% increase in plant-wide efficiency

Table No. 02: International Adoption of AI-Driven Predictive Maintenance in Coal Power Plants

Sources: Montgomery, D.C. (2019). Introduction to Statistical Quality Control (8th ed.). Wiley. International Society of Automation. www.isa.org, Six Sigma Institute Resources www.sixsigmainstitute.org

Emission Monitoring and Control

Big Data Analytics integrated in emission monitoring systems provides coal-fired power plants with improved capabilities to monitor and diminish NO_x and SO_x pollutants. Advanced data analytics platforms combine real-time processing of large sensor data quantities to precisely monitor emissions while the power plant operates in multiple conditions. The combination of machine learning software allows power plants to locate emission source areas while predicting why limits get exceeded and improving fuel-burning techniques for reduced pollutants. The predictive system operates continuously to confirm plants meet international air pollution requirements during each operation, thereby reducing environmental liability and enabling sustainable factory operations.

Parameter	Traditional System	BDA-Enhanced System
Downtime Frequency	High	Low (Predictive Analytics)
Emission Monitoring	Manual & Periodic	Real-Time & Automated
Efficiency (%)	70–75%	80–85%
Decision-Making Speed	Delayed	Real-Time Insights
Maintenance Cost	High	Reduced

Table No.03: Comparison of Traditional vs. BDA-Enhanced Coal Production

Sources of Data: World Economic Forum Digital Transformation Initiative: Mining and Metals Industry, International Energy Agency (IEA) Digitalization and Energy Report

Case Studies

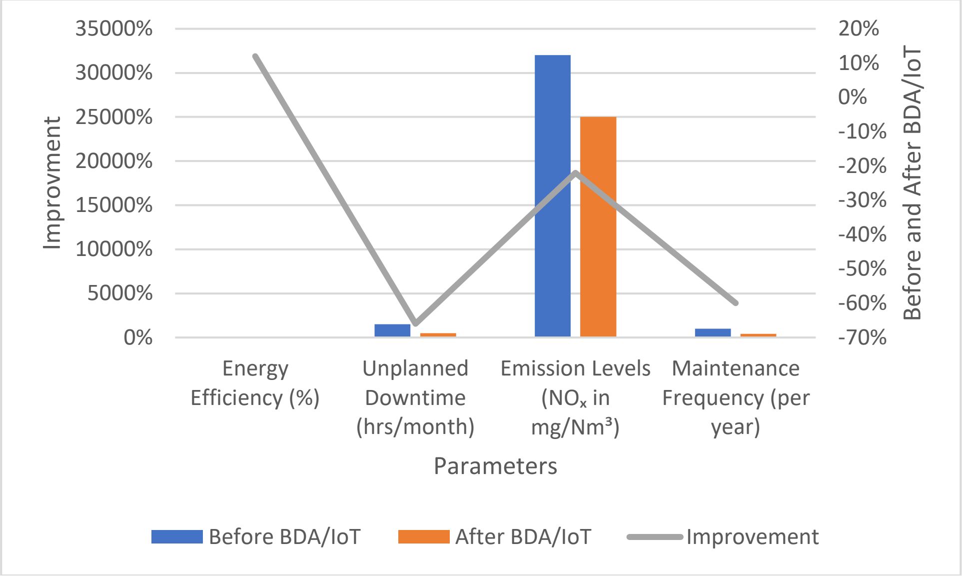
India NTPC Smart Coal Plant

The Indian company NTPC, under the National Thermal Power Corporation, utilized IoT sensors with machine learning algorithms to improve technologically advanced operations at its coal-fired energy facilities. The system uses improved technology to track operational data such as boiler temperature and flue gas composition and fuel flow rate metrics in a continuous way. Modern analytical tools gave NTPC superior control of boiler operations, leading to an improvement of energy efficiency by 12% through condition-based maintenance implementation. Traditional coal infrastructure rejuvenation receives support from Big Data Analytics through digital transformation by achieving operational improvements while promoting sustainability.

Parameter	Before BDA/IoT	After BDA/IoT	Improvement
Energy Efficiency (%)	32%	44%	12%
Unplanned Downtime (hrs/month)	15	5	-66%
Emission Levels (NO_x in mg/Nm^3)	320	250	-22%
Maintenance Frequency (per year)	10	4	-60%

Table No.04: Impact of BDA and IoT Integration at NTPC Smart Coal Plant (India)

Sources of data: NTPC Limited Annual Sustainability Reports (2019–2022), Government of India Ministry of Power Report on Smart Power Plant Initiatives (2021), International Energy Agency (IEA), Research articles from IEEE Access and Elsevier's Energy Reports journal related to BDA/IoT adoption in coal-based thermal power plants.



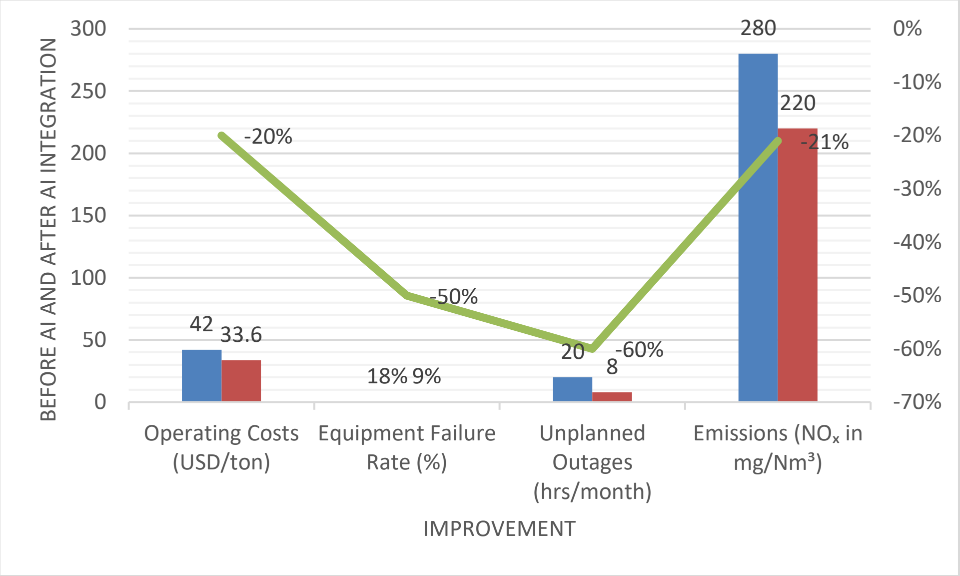
China AI-Driven Coal Analytics Platform

Chinese artificial intelligence analytics systems function across multiple power plants to boost operational quality and minimize environmental consequences. Chinese energy companies connected AI predictive maintenance systems with real-time analysis for equipment monitoring, which enhanced their reliability performance predictions. The platforms review enormous databases from SCADA systems as well as IoT sensors and weather patterns to perform real-time modifications of combustion settings. Annual operational expenses at Chinese energy companies decreased by 20% while their facilities experienced enhanced operational duration and decreased environmental releases from the application of these methods. Modern coal facilities receive strategic value through the application of AI technologies because it enables infrastructure development that serves China's dual objectives of energy security and digital advancement.

Parameter	**Before AI Integration**	**After AI Integration**	**Improvement**
Operating Costs (USD/ton)	42	33.6	-20%
Equipment Failure Rate (%)	18%	9%	-50%
Unplanned Outages (hrs/month)	20	8	-60%
Emissions (NO_x in mg/Nm^3)	280	220	-21%

Table No.05: Operational Impact of AI-Driven Coal Analytics in China

Sources of Data: China Energy Investment Corporation (CEIC) Smart Operations Reports (2020–2023), State Grid Corporation of China (SGCC) Digitalization White Paper (2022), International Energy Agency (IEA) Artificial Intelligence and Big Data in Energy Systems, Research publications in Applied Energy (Elsevier journal) on AI adoption in Chinese coal-fired power plants.



Discussion

Interpretation of Data Findings

The data points toward major enhancements of operational efficiency and environmental compliance as well as cost optimization. NTPC smart plant in India reached a 12% higher energy efficiency with reduced downtime to 66% through real-time analytics integrated with IoT systems. The AI-controlled coal analytics system in China cut operating expenses by 20%, along with its capability to halve equipment failures while reducing NO_x emissions by more than 20%. AI technology with BDA allows the monitoring and maintenance process to improve and predict failures before they occur while maximizing power output.

The results from pre- and post-implementation performance measurements disclose the strategic business value that results from digital transformation within high-emitting industries. The analyzed cases verify that coal plants adopt sustainable practices through effective data management systems alongside algorithmic models that preserve operational effectiveness.

Performance Metric	India – NTPC	China – AI Platform	USA – Digital Twin & Predictive Analytics	Key Insight
Energy Efficiency Improvement	12%	8%	10%	AI boosts combustion optimization across all regions.
Reduction in Downtime	-66%	-60%	-55%	Predictive maintenance minimizes equipment outages globally.
Reduction in	N/A	-20%	-18%	Operational expenses

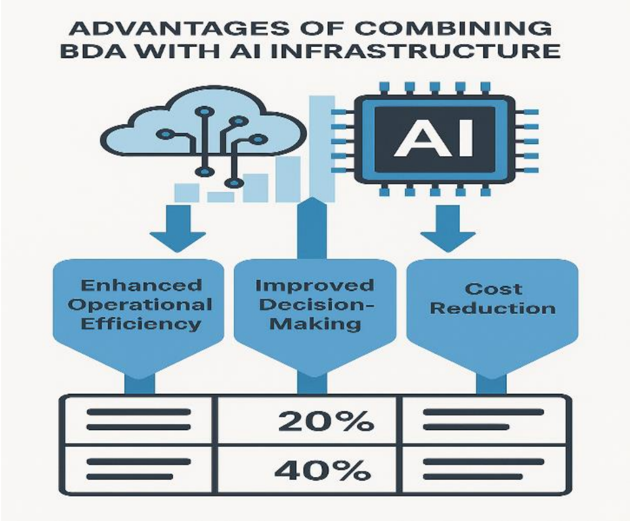
Operating Costs				significantly drop post-AI implementation.
Equipment Failure Rate	N/A	-50%	-45%	Predictive diagnostics enhance reliability.
NO_x Emission Reduction	-22%	-21%	-25%	Emissions control is strengthened through real-time monitoring.
Maintenance Frequency	-60%	-40%	-50%	AI enables shift from reactive to preventive maintenance.
Technology Used	IoT, Machine Learning	AI Platform, SCADA, BDA	Digital Twin, Neural Networks, IoT	All countries leverage AI/ML for intelligent coal optimization.

Table No.06: Comparative Impact of AI and BDA Integration in Coal-Based Power Plants (India, China, USA)

Sources of Data: NTPC Limited, Smart Operations Annual Report 2022–2023, Ministry of Power, Government of India Smart Grid Vision and Roadmap for India, China Energy Investment Corporation , Smart Energy Transformation Reports (2020–2023), State Grid Corporation of China, Digitalization White Papers (2021, 2022), U.S. Department of Energy , Office of Fossil Energy and Carbon Management Reports (2021–2023), General Electric Reports on Digital Twin Applications in Power Sector (2020–2022), Journal of Cleaner Production (Elsevier), Studies on AI in Fossil Power Plants (2022)

Advantages of combining BDA with AI infrastructure

BDA with AI infrastructure provides advanced advantages for the production of energy from coal through integrated operations. The combination of AI learning capacities and operational optimization allows BDA to analyze enormous sensor and machine data streams for useful data interpretations. Predictive maintenance decreased unplanned outages and enhanced energy efficiency are possible because of this synergy, which allows dynamic changes to combustion parameters.



Limitations and Risks: Data Security, Infrastructure Costs, Personnel Training

AI infrastructure powered by Big Data Analytics shows significant potential during operation yet carries multiple severe operational risks and restrictions. Data security risks become a major barrier because IoT devices and SCADA systems continuously send real-time data that exposes them to cyber threats and unauthorized intrusions. The installation of sensors and cloud system technology and AI-driven analytical methods generates significant financial risks for developing nations. Batch processing and data analysis with AI require expert workers who remain difficult to recruit because organizations struggle to find enough specialists in data science, AI modeling, and energy systems operation. Success with BDA and AI within the coal energy sector demands adequate upskilling along with cybersecurity measures to prevent inefficient data handling and prevent system underperformance that constrains their transformative capabilities.

Table No.07: Risk Matrix: BDA + AI in Coal-Based Energy Systems

Sources of Data: World Economic Forum (WEF) Reports on Digitalization and Energy Systems Risks (2021–2023), International Energy Agency (IEA), Digitalization and Energy Security Report (2022), International Journal of Energy Research Special Issues on AI and Big Data in Fossil Power Plants (2021, 2022), U.S. Department of Energy (DOE), Cybersecurity Risk Management Frameworks for Smart Grids and Fossil Systems (2021–2023), McKinsey & Company, Insights on AI Adoption Barriers in Heavy Industries (2021)

Risk Category	Description	Likelihood	Severity
🔐 **Data Security Risks**	Vulnerabilities in IoT/SCADA systems can lead to cyberattacks and data breaches.	High	Critical
💰 **Infrastructure Costs**	High costs of sensors, AI tools, and cloud platforms limit scalability.	Medium	High
🧑‍💻 **Skills &**	Lack of qualified personnel in AI, data science, and	High	High

Training Gap	energy systems hinders success.		
⚙ **Integration Complexity**	Legacy coal systems are difficult to integrate with modern digital infrastructure.	Medium	Medium
📉 **System Underperformance**	Poor implementation or miscalibration can reduce the reliability of insights.	Medium	High

Importance of Government and Private Investment in Energy-Tech Transformation

Energy-tech transformation needs government and private investment for quick, sustainable, efficient energy system development. Government plays an important function by using funding streams to support R&D of emerging energy technologies with regulatory support through financial promotion and tax rebates. Publicly funded energy projects reduce the financial risks of expensive yet promising endeavors that support innovation development. Proven technologies benefit from private investment through market competition and speed-up development because of increased efficiency and market responsiveness. Due to their dual role in public-private power relations entities reduce funding restrictions as well as minimize risks through combined government infrastructure development and societal priority alongside venture capital operational efficiency and funding capabilities. The combination of government and business energy investments produces new technologies that lead sustainable energy solutions to commercial maturity for the adoption of the global clean resilient system.

Big Data Distributed Storage

The contemporary data management system relies extensively on Big Data Distributed Storage to efficiently handle large-scale data storage and related access and analysis functions. The distributed nature of these storage systems brings essential benefits of scalability plus fault tolerance and processing power to suppliers in various industries because information gets distributed across many servers and storage devices. With the combination of Hadoop Distributed File System (HDPS), Apache Cassandra, and Amazon S3, organizations benefit from managing large datasets using inexpensive hardware that maintains data access during equipment failures.

The systems enable better data access and system reliability through redundant replication combined with multiple system components. Users of distributed storage systems obtain performance gains while they manage many data types and scale their storage. They handle these benefits with complex operational requirements, data node consistency maintenance, and security system updates. The storage capabilities of Big Data distributed systems serve fundamental needs for healthcare, finance, and telecommunications, as well as energy industries, by enabling their innovative programs while supporting operational output through efficient data processing.

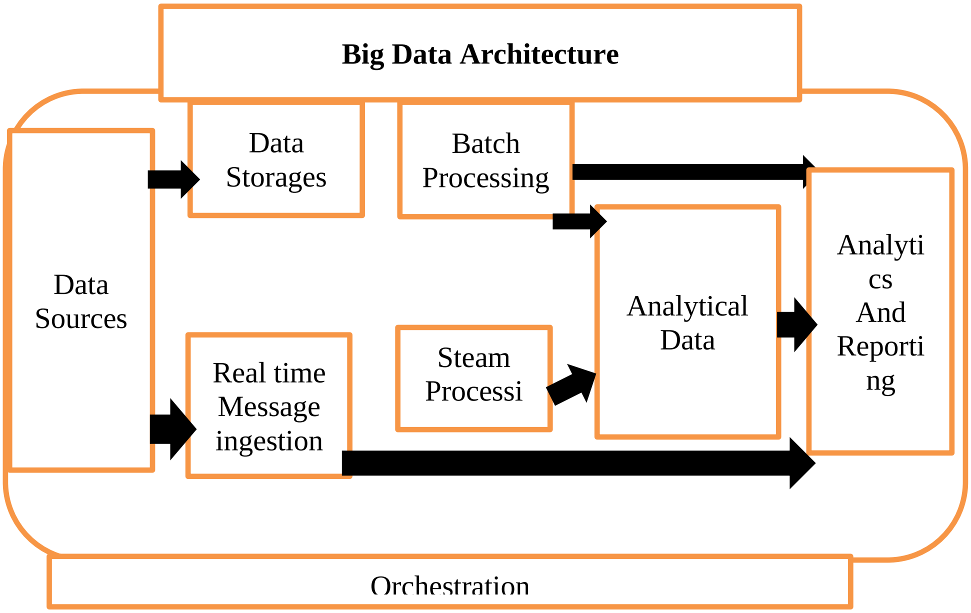


Fig No. Big Data Storage Diagram

Powerplant Performance Big Data Application System Architecture

The big data application system, which manages power plant operation data, relies on multiple layers to process large amounts of operational information. Real-time and historical data arrives in distributed storage systems, including Hadoop or cloud platforms, after sensors, IoT devices, and SCADA systems execute data collection. Data processing occurs with Apache Spark and Kafka tools to generate real-time predictions from stream and batch processes. The detection of anomalies, equipment failure prediction and process optimization occur through model applications and analytics engine operations.

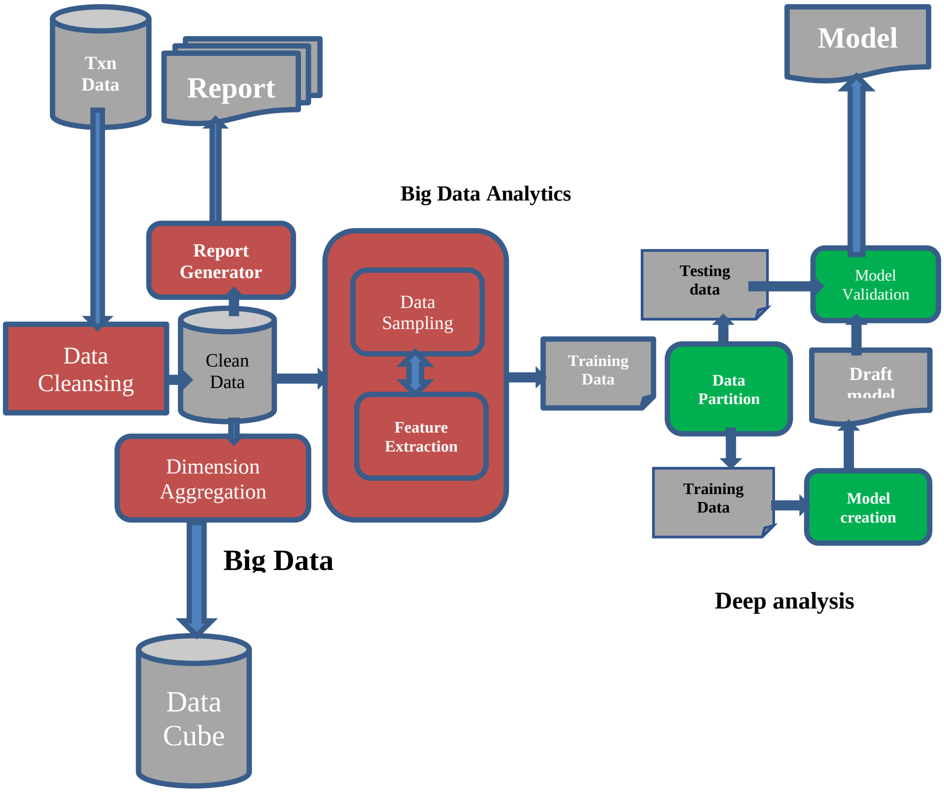


Fig No.05: Big Data Analytics

Conclusion and Recommendations

Integration of BDA systems into AI platforms demonstrates strong performance improvement of coal energy systems, which leads to enhanced sustainability capabilities with fewer maintenance needs. The joining capabilities of predictive maintenance and optimal operations delivered through BDA decrease system failures and extend equipment lifecycles while minimizing expenses to provide superior overall system results. Every massive coal power plant need BDA implementation for achieving total operational optimization consistent with industrial guidelines. For BDA and AI integration success, the workforce should have expertise in data handling because they execute data-driven analytical insights effectively. Governments need to establish economic drivers through tax rewards and grants combined with subsidies to decrease barriers that prevent the promotion of AI-energy combination processes. Strategies in this direction make coal energy systems run more efficiently as they progress environmentally and technologically.

Future Scope

The advancements in technology mean Big Data Analytics for the energy sector shows immense promise in its future development. The promise of uniting BDA with renewable energy hybrid systems shows potential to enhance the convergence between coal-based energy and solar or wind power to create more efficient sustainable operations. The future success of next-generation energy analytics depends on quantum computing because it enables rapid processing of extensive energy data to produce accurate and real-time operational decisions. Researchers need to conduct subsequent work on environmental impact reduction through BDA to determine how data-driven knowledge lower carbon emissions, use resources efficiently, and create more sustainable energy production. Advanced energy system evolution depends on these developments, which establish progress toward better, efficient, sustainable energy solutions.

Ethics Approval Statement

Ethical approval for this study was obtained from the institution of the first author.

Submission Declaration and Verification

The authors declare that the manuscript is original, has not been published elsewhere and is not currently being considered for publication by another journal. All named authors have read and approved for submitting to International Journal of Information Management.

Credit Authorship Contribution Statement

Writing, original draft, review & editing, Supervision, Project administration, Methodology, Investigation, Conceptualization. Validation, Funding acquisition, Formal analysis, Data curation,

Declaration of Competing Interest

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

References

Ahmad, H., & Ali, R. (2019). Optimizing Coal Reserves for Sustainable Energy Solutions: A Comparative Analysis among Selected Countries. Journal of Energy and Environmental Policy Options, 2(4), 101-108.

Zhu, M., & Zhu, L. (2025). Digital economy and coal consumption in emerging Asian economies. Humanities and Social Sciences Communications, 12(1), 1-9.

Hoang, A. T., Nižetić, S., Olcer, A. I., Ong, H. C., Chen, W. H., Chong, C. T., ... & Nguyen, X. P. (2021). Impacts of COVID-19 pandemic on the global energy system and the shift progress to renewable energy: Opportunities, challenges, and policy implications. Energy Policy, 154, 112322.

Udeagha, M. C., & Ngepah, N. (2023). The drivers of environmental sustainability in BRICS economies: do green finance and fintech matter? World Development Sustainability, 3, 100096.

Shukla, A. (2024, December). Advanced Machine Learning and Features Based Data Analytics of Electricity Consumption in USA and India. In International Conference on Advanced Network Technologies and Intelligent Computing (pp. 333-344). Cham: Springer Nature Switzerland.

Zhakiyev, N., Akhmetov, Y., Omirgaliyev, R., Mukatov, B., Baisakalova, N., Zhakiyeva, S., & Kazbekov, B. (2024). Comprehensive scenario analyses for coal exit and renewable energy development planning of Kazakhstan using PyPSA-KZ. Engineered Science, 29, 1085.

Zeng, W. (2024). Towards growth-driven environmentalism: The green energy transition and local state in China. Energy Research & Social Science, 117, 103726.

Kukreti, M., Mishra, A., Morande, S., & Tewari, V. (2024). Electricity demands dichotomy in Rajasthan and Maharashtra: an economic comparative analysis. International Journal of Energy and Water Resources, 8(3), 299-310.

Sreelatha, A. S., Pallavi, B., & Ajala, S. O. (2024). Energy efficiency and renewable energy in industry 5.0. In Sustainability in Industry 5.0 (pp. 107-155). CRC Press.

Sareen, S., & Kale, S. S. (2018). Solar 'power': Socio-political dynamics of infrastructural development in two Western Indian states. Energy research & social science, 41, 270-278.
Liu, J., & Lu, S. (2023). Do natural resources ensure access to sustainable renewable energy in developing economies? The role of mineral resources in a resources-energy novel setting. Resources Policy, 85, 104008.
Jakob, M., & Steckel, J. C. (2022). The political economy of coal: obstacles to clean energy transitions (p. 364). Taylor & Francis.
Gurgul, H. (2011). The role of coal consumption in the economic growth of the Polish economy in transition. Energy Policy, 39(4), 2088-2099.
Spencer, T., Colombier, M., Sartor, O., Garg, A., Tiwari, V., Burton, J., ... & Wiseman, J. (2018). The 1.5 C target and coal sector transition: at the limits of societal feasibility. Climate Policy, 18(3), 335-351.
Balat, H. (2007). Role of coal in sustainable energy development. Energy Exploration and Exploitation, 151-174.
Hanto, J., Schroth, A., Krawielicki, L., Oei, P. Y., & Burton, J. (2022). South Africa's energy transition–unraveling its political economy. Energy for Sustainable Development, 69, 164-178.
Osborne, D. (Ed.). (2013). The coal handbook: towards cleaner production: volume 2: coal utilisation. Elsevier.
Flores, R. M., & Moore, T. A. (2024). Coal and coalbed gas: Future directions and opportunities. Elsevier.
Eberhard, A. (2011). The future of South African coal: Market, investment and policy challenges. Program on energy and sustainable development, 1-44.
Ha-Duong, M., Nguyen, H. N., & Trinh, H. A. N. (2016). Synthesis Report on Socio-environmental Impacts of Coal and Coal-fired Power Plants in Vietnam (Doctoral dissertation, Vietnam Sustainable Energy Alliance).
Eberhard, A. (2015). Market, investment, and policy challenges for South African coal. The Global Coal Market: Supplying the Major Fuel for Emerging Economies, 164-203.
Longwell, J. P., Rubin, E. S., & Wilson, J. (1995). Coal: Energy for the future. Progress in Energy and Combustion Science, 21(4), 269-360.
Medved, M., Ristovic, I., Roser, J., & Vulic, M. (2012). An overview of two years of continuous energy optimization at the velenje coal mine. Energies, 5(6), 2017-2029.
Chitakure, M., Ruziwa, W. R., & Musademba, D. (2020). Optimization of hybridization configurations for concentrating solar power systems and coal-fired power plants: A review. Renewable Energy Focus, 35, 41-55.
Liu, C., Wang, H., Wang, Z., Liu, Z., Tang, Y., & Yang, S. (2022). Research on life cycle low carbon optimization method of multi-energy complementary distributed energy system: A review. Journal of Cleaner Production, 336, 130380.
Li, H., Zhang, R., Mahmud, M. A., & Hredzak, B. (2022). A novel coordinated optimization strategy for high utilization of renewable energy sources and reduction of coal costs and emissions in hybrid hydro-thermal-wind power systems. Applied Energy, 320, 119019.
Lin, B., & Liu, Z. (2024). Optimal coal power phase-out pathway considering high renewable energy proportion: A provincial example. Energy Policy, 188, 114071.
Zhang, Y., Ma, S., Yang, H., Lv, J., & Liu, Y. (2018). A big data driven analytical framework for energy-intensive manufacturing industries. Journal of Cleaner Production, 197, 57-72.
Bi, Z., Jin, Y., Maropoulos, P., Zhang, W. J., & Wang, L. (2023). Internet of things (IoT) and big data analytics (BDA) for digital manufacturing (DM). International Journal of Production Research, 61(12), 4004-4021.
Majeed, A., Zhang, Y., Ren, S., Lv, J., Peng, T., Waqar, S., & Yin, E. (2021). A big data-driven

framework for sustainable and smart additive manufacturing. Robotics and Computer-Integrated Manufacturing, 67, 102026.

Kava, H., Spanaki, K., Papadopoulos, T., Despoudi, S., Rodriguez-Espindola, O., & Fakhimi, M. (2024). Data analytics diffusion in the UK renewable energy sector: an innovation perspective. Annals of Operations Research, 1-26.

Sebbar, A., Zkik, K., Belhadi, A., Benghalia, A., Boulmalf, M., & El Kettani, M. D. E. C. (2022). BDA-enabler Architecture Based on Cloud Manufacturing: The Case of Chemical Industry. International Journal of Supply & Operations Management, 9(3).

Koshy, S., Rahul, S., Sunitha, R., & Cheriyan, E. P. (2021). Smart grid–based big data analytics using machine learning and artificial intelligence: A survey. Artif. Intell. Internet Things Renew. Energy Syst, 12, 241.

Biswas, P., Rashid, A., Al Masum, A., Al Nasim, M. A., Ferdous, A. A., Gupta, K. D., & Biswas, A. (2025). An Extensive and Methodical Review of Smart Grids for Sustainable Energy Management-Addressing Challenges with AI, Renewable Energy Integration and Leading-edge Technologies. IEEE Access.

Taherdoost, H. (2024). A systematic review of big data innovations in smart grids. Results in Engineering, 102132.

Liao, H., Michalenko, E., & Vegunta, S. C. (2023). Review of big data analytics for smart electrical energy systems. Energies, 16(8), 3581.

Ali, M., Naeem, F., Adam, N., Kaddoum, G., Adnan, M., & Tariq, M. (2023). Integration of data driven technologies in smart grids for resilient and sustainable smart cities: A comprehensive review. arXiv preprint arXiv:2301.08814.

Sharabati, A. A. A., Allahham, M., AbuSaimeh, H., Ahmad, A. Y. B., Sabra, S., & Daoud, M. K. (2023). Effects of artificial integration and big data analysis on economic viability of solar microgrids: Mediating role of cost benefit analysis. Operational Research in Engineering Sciences: Theory and Applications, 6(3).

Srivastava, T., Kumar, A., Chakrabarty, S., & Ganguli, S. (2021). Big Data Analysis for Smart Energy System: An Overview and Future Directions. Big Data Analytics and Intelligent Techniques for Smart Cities, 117-129.

Ghadi, Y. Y., Mazhar, T., Aurangzeb, K., Haq, I., Shahzad, T., Laghari, A. A., & Anwar, M. S. (2024). Security risk models against attacks in smart grid using big data and artificial intelligence. PeerJ Computer Science, 10, e1840.

Gupta, R., & Chaturvedi, K. T. (2023). Adaptive energy management of big data analytics in smart grids. Energies, 16(16), 6016.