iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81260

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

A Survey on Energy Management for Smart Grids Using Deep LakeLambda Architecture

Shwetha AB, Shiva Balaji S Katta, Surabhi Desai, Varshini Patil, Vishalakshi Koti
Dept. of CSE Sapthagiri College of Engineering

Abstract: Smart grids are popping up everywhere now, and that brings in a flood of data from all sorts of sources in energy systems. Some of it is neat and organized, but a lot is just messy unstructuredstuff.Itpushestheneedforwaystostoreeverything without losing track, process it fast, and even analyze right away. Traditional setups for managing energy often struggle with this, liketheycantscaleupeasilyorhandledelayswell,andpredicting things for power networks gets tricky too.

Wearelookingatrecentideasinsmartgridenergymanagement here, focusing on things like Data Lakes for storage and that LambdaArchitectureapproach.Itseemsliketheymixinbigdata tools with machine learning, including deeper stuff like CNNsand LSTM models, plus different forecasting methods. Batch processing works alongside real-time stuff, and that combo helps with load forecasts, better demand responses, and distributing energymoreevenly.Notsureifitcoverseverythingperfectly,but it stands out how these pieces team up.

Current systems still have issues, you know, like not scaling right or missing that quick response at the edge. So moving toward these smarter setups feels important. Pulling together DataLakesforstorage,Lambdaforhandlinglayers,andAI for predictions could make future grids way more efficient and reliable,maybeevengreenerovertime.Somepeoplemightargue its not fully there yet.

Thehybridmodelsgettechnicalfast,andIthinkthatpartis a bit unclear without digging deeper. Overall though, the way integration ties it all seems promising for energy handling, even as it evolves. That catches my attention mostly.

I. INTRODUCTION

Overthepastdecade,theelectricalpowersectorhasevolved significantlyduetorisingenergydemand,increasedintegration of renewable sources, and advances in digital technologies. Unlike traditional centralized systems, modern power grids operate in a dynamic and decentralized environment, where solar energy, wind power, electric vehicles, and smart ap- pliances introduce new operational complexities. Ensuring stability, efficiency, and reliability has therefore become a major challenge.

Smartgridshaveemergedasanadvancedsolution,enabling real-time monitoring, bidirectional communication, and auto- mated control. These systems rely on data-driven decision- makingbutgeneratemassiveanddiversedatasetsfromsources suchassmartmeters,sensors,andmarketplatforms.Traditional data management systems struggle to handle this volume, velocity,andvariety,leadingtoscalabilityandlatencyissues.

To address these challenges, big data technologies such as Data Lakes and distributed computing frameworks are widely adopted. Data Lakes allow storage of raw, heterogeneousdata, while Lambda Architecture combines batch and real-time processing to balance accuracy and low latency, makingit suitable for applications like load forecasting and fault detection.

Additionally,artificialintelligencetechniques—includingDe- cisionTrees,CNNs,andLSTMnetworks—enhancepredictive analytics, anomaly detection, and energy optimization. The integration of big data and AI enables efficient, scalable, and intelligent energy management systems.

This survey focuses on recent advancements in smart grid energymanagementusingDataLakeandLambdaArchitecture, analyzing existing approaches, identifying limitations, and highlighting future research directions for building scalable and low-latency systems.

II. BACKGROUND AND MOTIVATION

Traditionalelectricpowersystemswerebuiltonacentralized modelwithunidirectionalenergyflow,limitedmonitoring,and static decision-making based on predictable demand. While effectiveinthepast,thisapproachlackstheflexibilityrequired for today's dynamic energy landscape.

Modernpowersystemshaveshiftedtowarddecentralization duetotheintegrationofdistributedenergyresourceslikesolar panelsand windturbines,alongwiththeriseofelectricvehicles and dynamic pricing.

These changes introduce variability and uncertainty, requiring real-time, intelligent, and adaptive grid operations. However, traditional infrastructures struggle with bidirectionalenergyflowandlarge-scaledatahandling,leading to inefficiencies and reduced reliability.

Smartgridsaddressthesechallengesbyenablingreal- timemonitoring,automation,anddata-drivendecision-making. However, they generate massive, diverse, and high-speed data from multiple sources, which conventional systems cannot efficiently manage.

To overcome this, advanced solutions such as Data Lakes and Lambda Architecture are used. Data Lakes allow flexible storageofraw,heterogeneousdata,whileLambdaArchitecture combinesbatchandreal-timeprocessingforaccurateandlow-latencyinsights.Additionally,AItechniqueslikeDecision Trees, CNNs, and LSTMs enhance prediction, fault detection, and energy optimization.

TheintegrationofbigdataandAIprovidesascalable and intelligent framework for smart grid energy management, improving efficiency, reliability, and sustainability while sup- porting the development of future smart cities.

III. LITERATURE SURVEY

A. *Energy Load Forecasting Techniques in Smart Grids: A Cross-Country Comparative Analysis*

Author and Year: R. Hachache et al. (2024) Methodol- ogy:Thisstudyprovidesadetailedcomparativeevaluation ofdifferentenergyloadforecastingmethodsusingsmart grid datasets collected from multiple countries. It examines conventional statistical techniques such as ARIMA, alongside machine learning models like Random Forest and Support Vector Machines, and advanced deep learning approaches in- cludingCNN,LSTM,andcombinedCNN–LSTMarchitectures. The analysis focuses on how these models perform under varying geographical conditions and consumption behaviors. The findings indicate that hybrid models deliver superior resultscomparedtostandalonemethods.Inparticular,CNN is effective in capturing spatial patterns in the data, while LSTM excels at modeling temporal relationships, and their combination leads to higher forecasting accuracy and greater reliability across diverse datasets. Limitation: Although the study offers a thorough comparison of forecasting techniques, its primary focus remains on prediction accuracy, with limited attention given to data storage and processing aspects. It does not incorporate big data frameworks such as Data Lakes or Lambda Architecture, which are crucial for managing and analyzing real-time smart grid data. Furthermore, the use of hybrid models increases computational complexity, making them more resource-intensive and potentially challenging to implement in large-scale systems.

B. *Short-Term Electricity Load Forecasting Based on Neural- Prophet and CNN-LSTM*

AuthorandYear:ShuaiLuandTaotaoBao(2024)

Methodology: This study introduces a hybrid forecasting approach that integrates NeuralProphet with a CNN–LSTM framework to enhance short-term electricity load prediction. NeuralProphetisusedtomodelunderlyingtrendsandseasonal variations, while CNN is applied to capture spatial featuresand LSTM to learn temporal dependencies in the data. By combining these components, the model achieves improved accuracy and is better suited for handling complex, nonlinear patterns. The proposed approach is tested on real-world datasets and shows superior performance when compared to conventional methods and individual deep learning models. Limitation: The study mainly concentrates on prediction accuracy and does not consider challenges related to data management or real-time data processing. It also does not incorporate scalable big data frameworks such as Data Lakes orLambdaArchitecture,whichareessentialforhandlinglarge- scalesmartgriddata.Asaresult,itspracticalapplicability in real-world, large-scale smart grid environments remains limited.

C. *Machine Learning Techniques for Load Forecasting in Smart Grids with Renewable Energy Integration*

AuthorandYear:CyrusTinnyPlavila(2025)

Methodology: This paper provides a comprehensive overview of machine learning and deep learning methods applied to load forecasting in smart grid systems, particularly undertheinfluenceofrenewableenergyintegration.Itexamines a range of models, including ARIMA, Random Forest, CNN, LSTM, and various hybrid techniques. The study emphasizes that the inclusion of renewable energy sources increases variabilityanduncertainty,therebymakingaccurateforecasting more complex. In addition, it evaluates the performance of different models and identifies the most suitable approachesfor specific operating conditions and scenarios.

Limitation: The paper is primarily survey-oriented and doesnotincludeanypracticalimplementationorexperimental validation of the discussed methods. It also does not address real-timedataprocessingaspectsandfailstopresentacomplete system architecture tailored for smart grid applications.

D. BigDataReferenceArchitecturefortheEnergySector

AuthorandYear:K.Wehrmeisteretal.(2025)

Methodology: This paper presents a reference architecture tailored for big data applications in the energy domain, with a particular focus on smart grid systems. It outlines a layered framework that includes components for data acquisition, storage, processing, analytics, and application services. The proposed architecture places strong emphasis on secure data sharing, seamless interoperability among diverse systems, and theabilitytoscaleanalyticsasdatavolumegrows.Additionally, it underscores the role of integrating IoT technologies, cloud computingplatforms,anddistributedprocessingframeworksto enable efficient and reliable energy management. Limitation: The paper is mainly conceptual in nature, emphasizing ar- chitecturaldesignratherthanpracticaldeployment.Itdoes not incorporate specific forecasting techniques or provide experimentalvalidationusingreal-worlddatasets.Consequently, itsdirectapplicationtoreal-timesmartgridforecastingremains limited without additional implementation and evaluation.

E. DataLakes:ASurveyofConceptsandArchitectures

Author and Year: S. Azzabi et al. (2024) Methodology: This paper presents a comprehensive overview of Data Lake technologies, covering key aspects such as architecture, data ingestionmethods,storagemechanisms,governancestrategies, and processing approaches. It examines modern frameworks, including Lambda Architecture and Lakehouse models, and explains their benefits in managing large-scale data environ- ments. Additionally, the study addresses important challenges associated with Data Lakes, such as maintaining data quality, effective metadata management, and ensuring data security.

Limitation: The study is primarily theoretical in nature and does not specifically address applications within smart grid systems. It does not incorporate forecasting techniques, nor doesitillustratehowDataLakeframeworkscanbepractically applied to energy management scenarios.

F. Data Integration and Storage Strategies in Heterogeneous Analytical Systems

AuthorandYear:P.Koukaras(2025)

Methodology: This paper reviews different data integration andstoragestrategiesusedinlarge-scaleanalyticalsystems. It covers techniques such as data federation, lakehouse archi- tecture, metadata management, and interoperability between systems. The study emphasizes the importance of scalable and flexibledatamanagementsolutionsforhandlingheterogeneous data sources.

Limitation: The paper does not specifically address smart gridsystemsorenergyforecasting.Italsolacksimplementation of machine learning or deep learning models, making it less relevant for predictive analytics in energy management.

G. Enhancing Short-Term Power Load Forecasting With a TimesNet-Crossformer-LSTM Approach

AuthorandYear:JunHeetal.(2024)

Methodology:Thispaperpresentsanadvancedhybriddeep learningmodelthatcombinesTimesNet,Crossformer,and LSTMforshort-termloadforecasting.TimesNetcaptures temporal patterns, Crossformer models complex dependencies in multivariate time-series data, and LSTM handles sequential learning.Themodelprocessesmultidimensionalsmartgrid dataandachieveshighforecastingaccuracy.Theintegrationof transformer-based architectures improves the model’s abilityto capture long-term dependencies and complex relationships.

Limitation: The model is computationally expensive and complex,requiringhighprocessingpowerandresources.It alsolacksintegrationwithrealtimedataprocessingframeworks,makingitdifficulttodeployinpracticalsmartgridenvironmentswithcontinuousdatastreams.

H. DataLakeLambdaArchitectureforSmartGridsBigData Analytics

AuthorandYear:A.A.MunshiandY.Mohamed(2018)

Methodology: This paper introduces a big data framework forsmartgridsystemsthatintegratesDataLakestorage with Lambda Architecture to efficiently manage large vol- umes of generated data. The Data Lake component allows flexible storage of diverse data types—both structured and unstructured—collected from sources such as sensors, smart meters, and operational platforms. To handle data processing, the framework employs Lambda Architecture, which utilizes both batch processing for historical data analysis and a speed layer for real-time data handling. In addition, the system incorporates analytical tools and visualization dashboards to support informed decision-making, making it well-suited for large-scale smart grid analytics and management.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Limitation: Thestudyprimarilyemphasizesdataarchitectureanddoes not integrate contemporary machine learning or deep learning approaches for forecasting and prediction tasks. As it is based onearlierwork,itdoesnotreflectrecentadvancementsin AI, and it overlooks key challenges such as scaling learning models and incorporating modern neural network techniques.

IV. METHODOLOGY

The proposed methodology for smart grid energy manage- ment leverages a combination of Data Lake and Lambda Architecture to tackle the increasing complexities of modern powersystems,includingrisingelectricitydemand,integration of renewable energy sources, large-scale data generation, and the need for real-time decision-making. Smart grid environ- ments generate vast amounts of data from various operational components, and managing this data effectively requires scalable storage solutions, efficient processing mechanisms, and intelligent predictive models.

To address these requirements, the methodology brings to- getherDataLakeforflexibledatastorage,LambdaArchitecture forhandlingbothhistoricalandreal-timedataprocessing, andmachinelearningtechniquessuchasDecisionTreeand Convolutional Neural Networks (CNN). This integrated approachaimstocreateareliable,scalable,andfuture-oriented framework for efficient smart grid energy management.

A. DataAcquisitionandCollection

The first stage of the methodology involves collecting data from multiple smart grid sources. Modern energy systems generate information from smart meters, electricity usage records, renewable energy production units, market pricing systems, weather services, and utility operational logs. These sources provide valuable insights into energy generation, transmission, consumption, and demand behavior.

Thecollecteddatacanbeofdifferenttypes:

- Structured Data: Numerical tables such as hourly energy usage and pricing data
- Semi-Structured Data: JSON, XML, and logs generated from devices and systems
- Unstructured Data: Text reports, maintenance records, or external documents

Continuous data collection is essential because forecastingand decision-making depend on both historical patterns and real-time changes in the grid environment.

B. CentralizedStorageUsingDataLake

After collection, all raw data is stored in a centralized Data Lake. A Data Lake is preferred over traditional relational databasesbecausesmartgriddatasetsaremassive,diverse,and continuously growing. Traditional databases require predefined schemas and often struggle with scalability, whereas a Data Lake stores data in raw format until processing is required.

KeybenefitsofusingaDataLakeinclude:

- Storage of structured, semi-structured, and unstructured data in one platform
- Supportforpetabyte-scalestorage
- Easyintegrationofmultipledatasources
- Preservationofrawhistoricaldataforfutureanalytics
- Cost-effectiveandscalablestoragesolution

Possible technologies for implementation include HDFS, AWS S3, Azure Data Lake, or Google Cloud Storage.

Thisstageensuresthatthesystemcanefficientlystorelarge datasets without losing important information.

C. DataPreprocessingandQualityImprovement

Raw smart grid data may contain errors, missing values, duplicates,inconsistentformats,andirrelevantattributes.There- fore, preprocessing is performed before applying analytics and machine learning algorithms.

Thepreprocessingprocessincludes:

- MissingValueHandling:Replacingorestimatingunavail- able records
- DuplicateRemoval:Eliminatingrepeatedentries
- NoiseFiltering:Removingabnormalorcorrupteddata
- Normalization:Scalingnumericalvaluestoimprovemodel training
- FeatureSelection:Selectingrelevantvariablessuchas load, weather, time, and price

- Transformation:Convertingdataintoamachine-readable format

For example, date and time values can be transformed into features such as hour, day, month, and season to improve demand forecasting accuracy.

Libraries such as Pandas, NumPy, and Scikit-learn are used for preprocessing tasks.

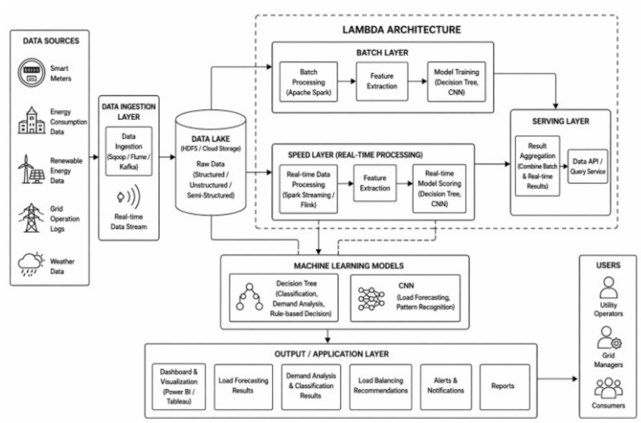

Fig. 1.Proposed System Architecture for Energy Management using DeepLake Lambda Architecture

D. .LambdaArchitectureforDataProcessing

To handle both historical and real-time smart grid data efficiently, the system adopts Lambda Architecture. This architecture combines batch processing and stream processing to achieve high accuracy with low latency.

Thearchitectureconsistsofthreelayers:

1) Batch Layer:The batch layer processes large volumes of historical data stored in the Data Lake. It performs deep analysisoncompletedatasetsandgenerateslong-term insights.

Mainfunctionsofthebatchlayer:

- Historicalloadpatternanalysis
- Seasonaldemandtrendidentification
- Consumerbehavioranalysis
- Long-termforecastingmodeltraining
- Energyconsumptionreporting

For example, the batch layer can identify that elec- tricity demand increases during summer afternoonsor winter nights.

Technologies used:Apache Hadoop,Apache Spark The batch layer ensures accurate and comprehensive analytics by using full historical data.

2) Speed Layer: The speed layer is responsible for pro- cessingreal-timedatastreamswithminimallatency.It playsacrucialroleinenablingimmediatemonitoring andrapiddecision-makinginsmartgridsystems.Un- like batch processing, this layer focuses on handling continuously incoming data and generating instant insights.

Theprimaryfunctionsofthespeedlayerinclude:

- Real-timemonitoringofelectricitydemand
- Detectionofsuddenloadspikes
- Instantidentificationoffaultsandanomalies
- Monitoring fluctuations in renewable energy gen- eration
- Providingimmediaterecommendationsforload balancing

For instance, if there is a sudden surge in electricity demand within a specific region, the speed layer quickly detects the change and generates alerts for system operators, allowing timely corrective actions. Common technologies used for implementing the speed layer include:Apache Kafka,Apache Spark Streaming,Apache Flink.

Overall, the speed layer ensures rapid response capabilities, which are essential for maintaining stability, reliability, and efficiency in modern smart grid operations.

3) ServingLayer:Theservinglayeractsastheinterface between the processing components and end-user applications. It integrates the outputs generated by both the batch layer and the speed layer to deliver final,consistent,andqueryableresults.Thislayer is responsible for making processed data easily accessible for visualization, reporting, and decision- making purposes.

Thekeyfunctionsoftheservinglayerinclude:

- Providingunifiedqueryresultsbycombiningbatch and real-time outputs
- Enabling fast retrieval of processed and analyzed data
- Supporting data visualization for dashboards and monitoring tools
- Offering decision support services for operators and stakeholders

Bymerginghistoricalinsightswithreal-timeupdates, the serving layer ensures that users receive accurate, up-to-date, and comprehensive information. This ca- pabilityisessentialforeffectivemonitoring,analysis, and control in smart grid systems.

Common technologies used in the serving layer include:Apache Cassandra,Elasticsearch

Overall, the serving layer plays a critical role in delivering reliable and efficient access to analytics results,therebysupportinginformeddecision-making in modern energy management systems.

E. MachineLearningModelsinSmartGrid

To enhance the intelligence and adaptability of smartgridsystems,machinelearningmodelsareintegratedinto dataprocessingpipelines.Thesemodelsenablepredictive analysis, pattern recognition, and automated decision- making based on large volumes of energy data.

1) *DecisionTreeAlgorithm:*TheDecisionTreealgorithm is a supervised learning technique widely used for classi- fication and decision-making tasks. It works by splitting dataintobranchesbasedonspecificconditions,ultimately formingatree-likestructurethatrepresentsdecisionrules. Its simplicity and interpretability make it suitable for various smart grid applications.

ApplicationsofDecisionTreeinsmartgridsinclude:

- Classificationofenergydemandlevels(high,medium, low)
- Identificationofpeakloadconditions
- Supportfordemandresponsestrategies
- Detectionofabnormalorirregularconsumption patterns
- Rule-basedoperationaldecision-making

ThekeyadvantagesoftheDecisionTreealgorithmare:

- Easytointerpretandexplain
- Fastexecutionandlowcomputationalcomplexity
- Effectiveperformanceonstructureddatasets
- Minimalrequirementfordatapreprocessing

For example, a Decision Tree model can classify elec- tricity demand as high when conditions such as elevated temperatureandpeakusagehours(e.g.,eveningtime)are met.

2) *ConvolutionalNeuralNetwork(CNN):* Convolutional Neural Networks (CNNs) are a class of deep learning models designed to automatically extract meaningful features from data. Although initially developed for image processing tasks, CNNs have demonstrated strong performance in analyzing time-series data, including energy consumption patterns.

ApplicationsofCNNinsmartgridsinclude:

- Short-termloadforecasting
- Predictionofpeakelectricitydemand
- Recognitionofconsumptionpatterns
- Forecastingofrenewableenergygeneration

Extraction of hidden and complex data features The main advantages of CNN models are:

- Highaccuracyinforecastingtasks
- Automaticfeatureextractionwithoutmanualinter- vention
- Strongcapabilityinidentifyingcomplexpatterns
- Improvedperformancecomparedtomanytraditional techniques
- CommonframeworksusedforimplementingCNNmodels include:

- TensorFlow
- Keras
- PyTorch

For instance, a CNN model can analyze historical energy consumptiondatatolearndaily,weekly,orseasonalusage patterns, thereby improving prediction accuracy in smart grid systems.

F. Forecasting andEnergy Optimization

The outputs generated from machine learning modelssuch as Decision Tree and Convolutional Neural Net- works (CNN) are utilized to enable intelligent energy management.Basedonpredicteddemandpatternsandthe currentstateofthegrid,thesystemrecommendsoptimized control actions to improve overall efficiency.

Keyoptimizationoutcomesinclude:

- Balanced distribution of electrical load across differ- ent regions
- Implementationofpeakshavingtechniquestoreduce demand spikes
- Improvedutilizationofrenewableenergysources
- Reducedrelianceonhigh-costbackuppowergenera- tion
- Minimizationofenergylossesandwastage
- Enhancedcostefficiencyingridoperations

This stage plays a vital role in ensuring that available energy resources are allocated and utilized in the most efficient and sustainable manner.

G. VisualizationandMonitoring

Theprocessedresultsandanalyticaloutputsarepresented through interactive dashboards and reports, enabling operators and decision-makers to effectively monitor and understand system behavior.

Commonlyusedvisualizationtoolsinclude:

- PowerBI
- Tableau
- Plotly
- Matplotlib
- Typicaldashboardfeaturesinclude:
- Real-timevisualizationofenergydemand
- Forecastinggraphsandpredictiveinsights
- Analysisofhistoricalconsumptiontrends
- Alertsforfaultsandabnormalconditions
- –Monitoring of key performance indicators (KPIs) Effectivevisualizationenhancessystemtransparencyand supports informed, data-driven decision-making.

H. PerformanceEvaluation

Theeffectivenessoftheproposedsystemisassessedusing standard evaluation metrics that measure both prediction accuracy and system performance.

Importantevaluationmetricsinclude:

- Accuracy:Measuresthecorrectnessofmodelpredic- tions
- RMSE (Root Mean Square Error): Evaluates predic- tion error magnitude
- MAE(MeanAbsoluteError):Measuresaverage deviation between predicted and actual values
- Latency: Indicates the response time for processing real-time data
- Scalability: Assesses the system's ability to handle increasing data volume
- Reliability: Reflects consistent system performance under varying load conditions
- The proposed approach is expected to deliver improved performance compared to conventional systems, particu- larly in terms of forecasting precision and real-time data processing capabilities.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

I. ExpectedImpact

The integration of Data Lake storage, Lambda Archi- tecture, and machine learning models such as Decision Tree and CNN provides a comprehensive framework for advanced smart grid energy management.

Theanticipatedbenefitsofthisapproachinclude:

- Enhancedaccuracyinenergydemandforecasting
- Fasterandmoreefficientdecision-makingprocesses
- Betterutilizationofrenewableenergyresources
- Reductioninoperationalandenergycosts
- Improvedreliabilityandstabilityofthepowergrid
- Decreasedenergywastage
- Scalableinfrastructurecapableoffutureexpansion
- Strong support for the development of smart cities Overall, the proposed methodology offers a robust, scal- able,andintelligentsolutionfornextgenerationenergymanagementsystems,addressingbothcurrentchallengesandfuturerequirementso fsmartgridenvironments.

V. CONCLUSION

The transition from traditional power systems to smart grids has introduced both opportunities and challenges, includingrisingenergydemand,renewableintegration,dy- namicpricing,andmassivedatageneration.Conventional energy management systems struggle with scalability, speed, and forecasting accuracy.

This study highlights the role of big data and AI in ad- dressingtheseissues.Machinelearninganddeeplearning models—suchasDecisionTrees,CNNs,LSTMs,and hybrid approaches—offer improved forecasting accuracy by capturing complex consumption patterns and trends. Scalabledataarchitecturesareequallyimportant.Data Lakesenableflexiblestorageofdiversedatatypes,while Lambda Architecture combines batch and real-time processing to balance accuracy and low latency, ensuring timely and reliable insights.

The integration of Decision Trees and CNNs enhances decision-making and load forecasting, creating a more adaptive system. Overall, this approach improves effi- ciency,reliability,cost-effectiveness,andrenewableenergy utilization.

Inconclusion,combiningDataLake,LambdaArchitecture, and AI provides a scalable and future-ready solution for smart grid energy management, supporting sustainable energy systems and smart city development.

REFERENCES

[1] R. Hachache, M. Labrahmi, A. Grilo, et al., “Energy LoadForecastingTechniquesinSmartGrids:ACross- Country Comparative Analysis,” Energies, vol. 17, no. 10, pp. 2251, 2024.
[2] S.LuandT.Bao,“Short-TermElectricityLoadFore- casting Based on NeuralProphet and CNN-LSTM,” IEEE Access, 2024.
[3] C.T.Plavila,“MachineLearningTechniquesforLoad Forecasting in Smart Grids with Renewable Energy Integration,” MDPI Energies, 2025.
[4] K. Wehrmeister, et al., “Big Data Reference Archi- tecture for the Energy Sector,” Springer Journal of Big Data, 2025.
[5] S.Azzabi,etal.,“DataLakes:ASurveyofConcepts and Architectures,” Springer, 2024.
[6] P.Koukaras,“DataIntegrationandStorageStrategies in Heterogeneous Analytical Systems,” MDPI, 2025.
[7] J. He, K. Yuan, Z. Zhong, and Y. Sun, “Enhancing Short-Term Power Load Forecasting With a TimesNet- Crossformer-LSTMApproach,”IEEEAccess,vol.12, 2024.
[8] A. A. Munshi and Y. A. I. Mohamed, “Data Lake Lambda Architecture for Smart Grids Big Data Analytics,” IEEE Access, vol. 6, pp. 40463–40474, 2018.