International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s,2025
https://theaspd.com/index.php

AI Meets Energy: Forecasting The Future Of Country-Level Energy Consumption

Chandan Reddy Chittimalle[1], Naveen Sandra[2], Macha Babitha[3]
[1]Sr. Electric Engineer, Viking Yachts, New Gretna, chandanreddyc@gmail.com
[2]Sr. Business Analyst, Lotus Data LLP, Princeton, NJ, USA, naveensandra2@gmail.com
[3]Head of Research & Development, Aroha Daily Essentials, Tirupati, India, machababitha@gmail.com

Abstract
Accurate forecasting of national energy consumption is critical for enabling effective energy planning, sustainable development, and informed policy formulation. This study investigates and compares a diverse set of forecasting models—including deep learning architectures (LSTM, CNN-BiLSTM, Transformer), machine learning ensemble methods (XGBoost, LightGBM, CatBoost), and classical time series approaches (SARIMA)—to predict annual oil, gas, and renewable energy consumption across three major economies: the United States, China, and India. Leveraging a harmonized multi-decadal dataset, extensive preprocessing techniques were employed to ensure temporal consistency, normalize consumption metrics, and enhance feature representation. Each model was trained using consistent time-series cross-validation and evaluated using a standard suite of performance metrics. The research aims to assess the strengths and limitations of each modeling paradigm in the context of national-level energy forecasting, and to provide a foundation for data-driven model selection strategies in energy systems analytics.
Keywords: *Energy Forecasting; Machine Learning; Deep Learning; Time Series Analysis; XGBoost; LSTM; SARIMA; Renewable Energy; National Energy Consumption; Comparative Modeling; Sustainability Analytics; Cross-Validation*

1. INTRODUCTION

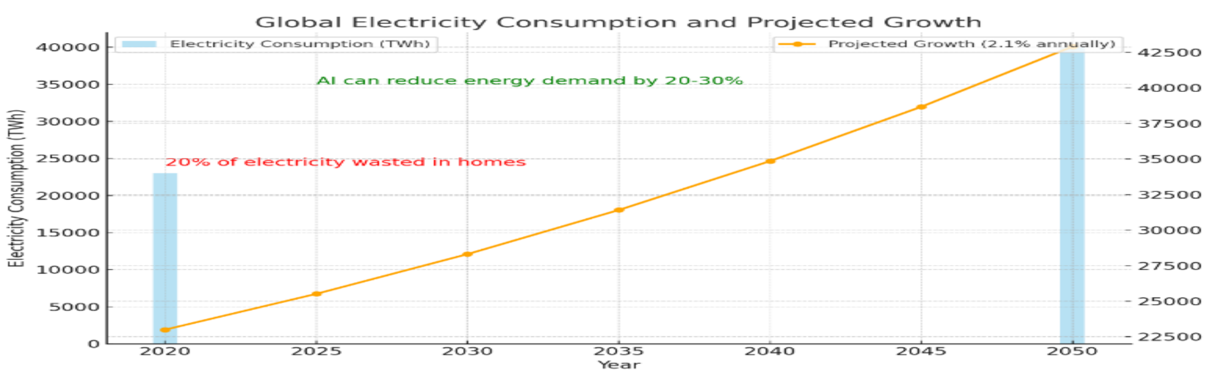
The global energy demand has witnessed an exponential rise due to factors such as population growth, industrialization, and urbanization. According to the International Energy Agency (IEA), global energy consumption is projected to increase by 25% between 2020 and 2040, primarily driven by rising energy demands from developing nations. This trend places considerable pressure on existing energy infrastructure, and the need for efficient energy management has never been more urgent. Traditional energy meters provide a basic measure of energy consumption, but they fall short of delivering insights into consumption patterns, identifying inefficiencies, or providing actionable solutions for optimization.

AI-based energy meters, however, integrate advanced machine learning (ML) and deep learning (DL) algorithms to bridge this gap. By utilizing large datasets, AI-driven systems can predict energy consumption, identify unusual patterns, and provide recommendations for reducing energy waste. These systems can be particularly beneficial for residential, industrial, and commercial applications, improving energy efficiency and promoting sustainable practices. Key statistics that highlight the importance of AI-based energy management is shown in Table 1 and Graph 1.

Table 1. Key Statistics and Visualization Insights

Statistic	Value/Insight	Source	Visualization Insights
Global electricity consumption in 2020	23,000 TWh (terawatt-hours)	IEA	Bar chart shows global consumption in 2020.
Projected global electricity consumption by 2050	Over 40,000 TWh	IEA	Bar chart shows projected consumption in 2050.
Annual growth rate of electricity consumption [1]	2.1%	IEA	Line graph shows projected growth over time.

Energy efficiency improvements by 2040	20-30% reduction in energy demand across sectors	McKinsey & Company	AI-driven improvements in energy efficiency by 2040.
AI's role in energy efficiency improvements	AI-based technologies to drive 20-30% reduction in energy demand by 2040	McKinsey & Company	Annotations highlight AI's impact on reducing energy demand.
Electricity waste in U.S. homes	20% of electricity consumed is wasted	U.S. Department of Energy	Annotations highlight the 20% waste in U.S. homes due to inefficiency.
Key cause of electricity waste in U.S. homes	Inefficient appliances, faulty devices, and lack of awareness of usage patterns	U.S. Department of Energy	AI-powered energy meters help address waste issues by providing data-driven recommendations.
AI-powered energy meters' contribution [2]	Provides data-driven recommendations and automates optimization to reduce electricity waste	U.S. Department of Energy	Line graph and bar chart annotations show how AI can reduce waste and optimize energy use.



Graph 1. Global Electricity Consumption Trends and AI's Role in Energy Efficiency

The growing concern over energy inefficiency and the associated environmental impact necessitates the development of intelligent systems that not only measure consumption but also optimize it in real-time. AI-based energy meters are capable of analyzing patterns in real-time, detecting anomalies in consumption, and offering tailored solutions to reduce energy use, making them essential components in the transition towards more sustainable and efficient energy systems[16][17].

2. LITERATURE SURVEY

Each of the following studies contributes uniquely to the growing body of knowledge in energy consumption forecasting using ML and DL, highlights unresolved challenges, and presents the outcomes of implemented models, and summary of reviewed studies is specified in Table 2.

[3] Focusing on transfer learning with transformers, the authors compared PatchTST, Informer, and vanilla Transformer models for energy prediction across smart buildings. PatchTST emerged as the most accurate and flexible when applied to feature-sparse domains. Transfer learning enabled the reuse of pre-trained models across buildings, but performance varied based on dataset alignment [18].

[4] This empirical evaluation of foundation models in energy analytics used the ComStock dataset to test RNNs and large-scale pretrained models. The findings suggested that heterogeneity in building data impacts performance more than model complexity or parameter count. The study advocated for better data harmonization rather than deeper models alone [19].

[5] Tailoring LSTM and GRU architectures for smart meter data, this study developed lightweight models that could be deployed on household-level devices. LSTM slightly outperformed GRU and ARIMA in RMSE and MAE metrics. The results were promising, especially for short-term forecasting, though challenges remained in ensuring adaptability to different home environments [20].
[6] This study employed CNN-GRU architectures with appliance clustering to improve residential short-term forecasting. The model was trained to recognize patterns in appliance-level behavior using convolutional layers, followed by GRUs for temporal sequence modeling. Its accuracy was notably higher than baseline LSTM models due to the incorporation of behavioral associations, though explainability remained limited due to the deep architecture [21].
[7] The authors examined the application of deep learning models to predict instantaneous energy usage in a manufacturing context. Among CNN-LSTM, LSTM, and TCN, the TCN model achieved the best performance due to its ability to manage long input sequences efficiently. Despite the promising results, the authors noted the model's sensitivity to hyperparameter tuning and input sequence length [22].
[8] Combining Stationary Wavelet Transform (SWT) with deep learning models such as CNNs, LSTMs, and Transformers, this study tackled multistep household energy forecasting. SWT improved signal quality by reducing noise, while the deep models captured sequential dependencies. The proposed cascaded hybrid system was effective for long-term forecasting, although error accumulation over extended horizons was a concern [23]. [9] Hoshino's work introduced a visual economic analysis tool using the Screening Curve Method for evaluating solar PV and battery investment decisions at the household level. It allowed users to simulate various configurations and estimate cost-efficiency and payback periods. While the method provided quick insights, its simplification of load patterns was cited as a limitation for precision applications [14]. [10] Investigating monthly consumption forecasts across large residential datasets, the study tested fully connected networks (FCN), CNNs, and LSTMs. With over 9 million records used for training, LSTM outperformed other architectures due to its memory retention capacity. This work emphasized the scalability of deep models for long-term aggregated consumption prediction, while noting computational costs[15]. [11] The research explored weather-based probabilistic forecasting for microgrids, focusing on load, renewable generation, and energy prices. Models incorporated uncertainty via quantile forecasts and probabilistic density estimations. The system showed value in day-ahead scheduling and minimizing reliance on backup sources, though accuracy was highly dependent on meteorological inputs.
[12] The authors integrated rough set theory and Deep Belief Networks (DBNs) to build a predictive model for public building energy consumption. Rough set theory was used for attribute reduction, improving model efficiency and interpretability. The DBN structure helped capture non-linear interactions between features, outperforming classical ANN and fuzzy logic approaches, especially in terms of generalization on unseen data.
[13] Focused on mining behavioral patterns from household appliance-level energy data, this study developed Bayesian network models capable of learning appliance usage dependencies. Their work revealed correlations between device usage and household routines, which were then used to improve day-ahead forecasting accuracy. The approach worked well in behavior-rich datasets but faced challenges scaling new, unseen households without retraining. [14] A comprehensive analysis of over 150 studies focused on data-driven energy models. The review outlined various ML methods including SVM, ANN, decision trees, and k-NN applied to both residential and commercial buildings. The authors noted the inconsistency of evaluation metrics and the lack of standard datasets as key limitations in the literature. The review helped categorize existing approaches based on input features, learning techniques, and building types, providing a strong base for future model development.

Table 2. Summary of Reviewed Studies

Ref. No.	Methodology	Application Area	Key Findings
[3]	PatchTST, Informer, Transformer + TL	Smart Buildings	PatchTST effective in TL with sparse data

[4]	RNNs, Foundation Models	Commercial Buildings	Heterogeneity impacts more than model size
[5]	Customized LSTM/GRU, ARIMA	Household	LSTM superior for short-term load prediction
[6]	CNN-GRU with Appliance Clustering	Household	Behavior-based associations improved accuracy
[7]	CNN-LSTM, TCN	Manufacturing	TCN best for high-frequency data
[8]	SWT + CNN/LSTM/Transformer	Household	Improved long-term forecasts using hybrid model
[9]	Screening Curve Method	Solar Households	Economic tool for PV and battery investment
[10]	FCN, CNN, LSTM	Residential (monthly)	LSTM best with large datasets
[11]	Probabilistic ML + Weather Forecasting	Microgrids	Supports uncertainty-aware scheduling
[12]	Rough Set + Deep Belief Networks	Public Buildings	Improved generalization, reduced complexity
[13]	Bayesian Networks	Household	Behavioral dependencies boost prediction
[24]	ML Survey (SVM, ANN, Decision Trees)	General	Review of 150+ papers; identified gaps in standardization

3. METHODOLOGY

This section outlines the end-to-end methodology adopted for forecasting global energy consumption trends. It encompasses data acquisition, preprocessing, feature engineering, model development, and evaluation.

3.1 Data Acquisition

This study utilizes the "World Energy Consumption" dataset, sourced from an open-access repository on Kaggle. The dataset provides annual records of global energy consumption alongside key demographic and economic indicators, enabling robust long-term forecasting across multiple countries. It spans from 1900 to 2022, with most energy-related metrics reliably populated from 1965 onward, and includes data from over 300 countries and regional aggregates.

The dataset integrates information from highly credible sources, including the BP Statistical Review of World Energy, Ember Climate, and Our World in Data (OWID). It has been harmonized for unit consistency (e.g., TWh), standardized country names, and temporal alignment. This ensures both accuracy and comparability across time and regions.

For the purpose of this study, the dataset was filtered to retain only key attributes relevant to energy demand modeling, including in table 3.

Table 3. Key Attributes considered for modeling

Attribute	Description	Type
Country	ISO-aligned country or region name	Categorical
Year	Year of observation	Integer
Population	Total population (mid-year estimate)	Numeric
Gdp	Gross Domestic Product (in current USD)	Numeric
oil_consumption	Annual oil consumption (TWh)	Numeric
gas_consumption	Annual gas consumption (TWh)	Numeric
renewables_consumption	Aggregate of solar, wind, biofuel, and hydro (TWh)	Numeric

Other variables, such as coal, nuclear, and individual renewable sources, were available but excluded from the primary modeling pipeline. The dataset was downloaded from [25] Kaggle website and stored locally in CSV format for further preprocessing and model development.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s,2025
https://theaspd.com/index.php

3.2 Data Preprocessing and feature engineering

Prior to model development, a structured and standardized preprocessing pipeline was implemented to prepare a clean, temporally aligned, and feature-rich dataset suitable for a diverse set of forecasting models. These include deep learning architectures (LSTM, CNN-BiLSTM, Transformer), gradient boosting ensemble models (XGBoost, LightGBM, CatBoost), and statistical methods (SARIMA). The preprocessing steps were designed to enhance data integrity, ensure chronological consistency, reduce noise, and generate predictive features that could be uniformly applied across all model types. The following stages summarize the transformation of the raw energy consumption dataset into a model-ready format.

3.2.1 Feature Selection

To reduce noise and retain only relevant macroeconomic and energy indicators, the dataset was filtered to include the following columns:

- country, year, population, gdp
- oil_consumption, gas_consumption, renewables_consumption (computed as the sum of individual renewables):

renewables_consumption=solar_consumption+wind_consumption+biofuel_consumption

This ensured that both demographic and economic drivers of energy use were retained while discarding unrelated or redundant attributes.

3.2.2. Temporal Alignment and Missing Value Imputation

Time-series integrity was preserved through chronological sorting and imputation:

- **Sorting**:

Data was sorted by country and year to maintain the sequential order of observations.

- **Forward-Fill Imputation**:

Missing values were imputed using forward-fill (ffill) within each country group, preserving time-order consistency without backward leakage.

- **Final Cleanup:**

Any remaining rows with missing values were dropped.

This produced a consistent, temporally aligned dataset with no missing values.

3.2.3 Per Capita Feature Engineering

To normalize energy demand across countries with varying population sizes, per capita consumption metrics were computed:

oil_per_capita = oil_consumption / population

gas_per_capita = gas_consumption / population

renewables_per_capita = renewables_consumption / population

This step enabled better cross-country comparisons and scale-invariant modeling.

3.2.4 Lag and Rolling Features

To incorporate temporal context into the forecasting models, lag-based and rolling statistical features were generated for each country's energy consumption data. These features capture both short-term and long-term dependencies essential for time series forecasting. Specifically, for each of the three energy types—oil, gas, and renewables—a 1-year lag (lag1) and a 12-year lag (lag12) were computed to reflect recent history and potential seasonal or cyclical patterns, respectively. In addition, a 3-year rolling mean (roll3) was calculated to smooth short-term fluctuations and highlight broader consumption trends. This enriched feature set enables models like LSTM, CNN-BiLSTM, and Transformer to learn from past temporal dynamics, while also providing valuable predictors for tree-based models such as XGBoost, LightGBM, and CatBoost, as well as for traditional statistical models like SARIMA. After the lag and rolling features were created, any rows containing NaN values—introduced due to lagging at the start of each series—were removed to maintain dataset consistency.

3.2.5 Logarithmic Transformation

To address skewness in the distribution of energy consumption variables and promote statistical stability, logarithmic transformations were applied to the raw consumption values. This transformation compresses

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s,2025
https://theaspd.com/index.php

the scale of large values while preserving the relative differences between data points, which is particularly beneficial when dealing with countries exhibiting vast differences in consumption levels.

Specifically, a natural log transformation with a shift was used, computed as log(1 + x), where x represents the original value. This formulation avoids undefined values for zero entries and is widely used in time series modeling to stabilize variance and reduce the impact of outliers. The transformation was applied to the following variables:

- oil_consumption
- gas_consumption
- renewables_consumption

The resulting features,
oil_consumption_log, gas_consumption_log and renewables_consumption_log—were added as new columns to the dataset. These transformed variables enhance model performance by making patterns more linear and easier for both statistical and machine learning models to interpret.

3.2.6 Output Export and Country Selection

The final preprocessed dataset, enriched with per capita metrics, lag features, rolling means, and log-transformed variables, was saved as a CSV file in the designated project folder on Google Drive.

After export, the dataset was filtered to include only United States, China, and India, selected for their significant global energy consumption and reliable historical data coverage. This refined dataset was used uniformly across all forecasting models: LSTM, CNN-BiLSTM, Transformer, XGBoost, LightGBM, CatBoost, and SARIMA.

3.3 Modeling and Training :

This research implements a comprehensive suite of forecasting models to estimate annual energy consumption across three leading global economies: the United States, China, and India. The target variables encompass three energy types: oil, natural gas, and renewables. The models deployed span both deep learning and statistical domains, enabling rigorous benchmarking under a unified evaluation pipeline.

3.3.1 Long Short-Term Memory (LSTM)

The LSTM model [Figure 1] is designed to capture long-range temporal dependencies in annual energy consumption data. Each input consists of six years of historical features; the model predicts the seventh year's consumption for oil, gas, or renewables.

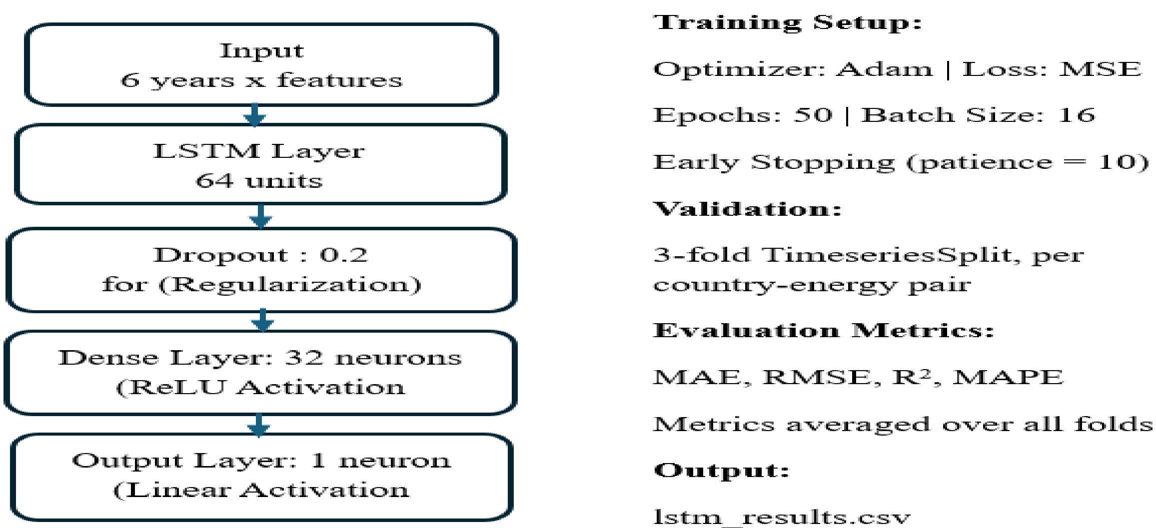


Figure 1. LSTM Model Structure for Energy Forecasting

3.3.2 CNN–BiLSTM Hybrid

The CNN–BiLSTM model [Figure 2] combines the strengths of convolutional and recurrent neural networks to capture both short-term and long-range dependencies in annual energy consumption data. A onedimensional convolution layer first detects local patterns, followed by a bidirectional LSTM that learns temporal dynamics in both forward and backward directions.

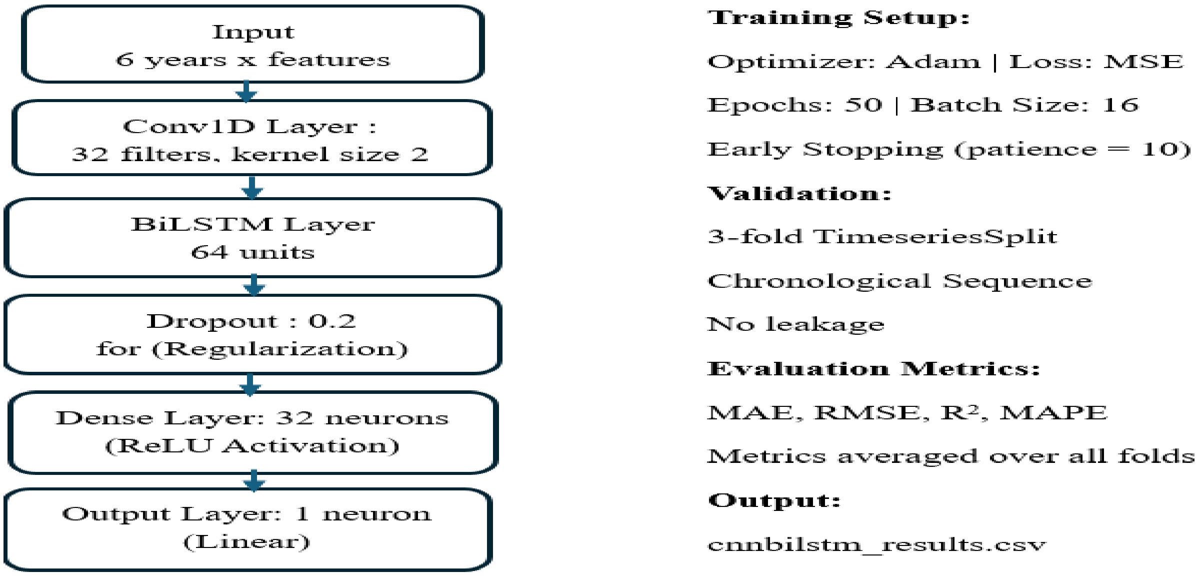


Figure 2. CNN-BiLSTM Model Structure for Energy Forecasting

3.3.3 Transformer

The Transformer model [Figure 3] introduces self-attention mechanisms to learn temporal dependencies without recurrence. It is particularly effective for capturing both short- and long-range relationships in structured energy consumption data.

Input: 6 years x features

Projection: Dense Layer to project input into 64 Dimensions

Self-Attention: Multi-Head Attention (4 heads)

Normalization: Add & LayerNorm

FeedForward Block: Dense (ReLU), Add & LayerNorm

Pooling & Output: GlobalAveragePooling -> Dense(64) -> Dense(1)

Training Setup:

Optimizer: Adam | Loss: MSE | Epochs: 50 | Batch Size: 16 | Early Stopping (patience = 10) | Validation split: 10%

Validation:

3-fold TimeseriesSplit | NaN filtering

Evaluation Metrics:

MAE, RMSE, R^2, MAPE (Metrics averaged over all folds)

Output:

transformer_results.csv

Figure 3. Transformer Model Structure for Energy Forecasting

3.3.4 XGBoost

XGBoost [Figure 4] is a scalable, tree-based ensemble algorithm optimized for speed and accuracy. It operates on tabular input, making it well-suited for structured macroeconomic and demographic energy forecasting data.

Model Configuration:

Estimators: 100

Learning Rate: 0.1

Max Depth: 4

Random Speed: 42

Training Setup:

No sequence modeling required | Predicts Consumption year-by-year | Trained using static feature vectors

Validation:

3-fold TimeseriesSplit | Temporal Separation across folds

Evaluation Metrics:

MAE, RMSE, R^2, MAPE (Metrics averaged over all folds)

Output:

xgboost_results.csv

Figure 4. XGBoost Model Structure for Energy Forecasting

3.3.5 LightGBM

LightGBM [Figure 5] is a high-performance gradient boosting framework that employs a leaf-wise tree growth strategy. It efficiently models non-linear dependencies from structured data without requiring sequence modeling.

Model Configuration:

Estimators: 100

Learning Rate: 0.1

Max Depth: 4

Random Speed: 42

Training Setup:

Operates on Static Tabular features | Performs year-by-year regression | Fast convergence with low memory footprint

Validation:

3-fold TimeseriesSplit | Temporal Consistency ensured

Evaluation Metrics:

MAE, RMSE, R^2, MAPE (Metrics averaged over all folds)

Output:

lightgbm_results.csv

Figure 5. LightGBM Model Structure for Energy Forecasting

3.3.6 CatBoost

CatBoost [Figure 6] is a robust gradient boosting framework designed for efficient handling of categorical and numerical features. In this study, it is applied on purely numerical inputs, eliminating the need for normalization or encoding.

Model Configuration:

Estimators: 100

Learning Rate: 0.1

Depth: 4

Random Speed: 42

Verbose: OFF

Training Setup:

Operates on tabular data | Automatically manages missing values | Minimal preprocessing required

Validation:

3-fold TimeseriesSplit | Fold-wise evaluation with preserved temporal order

Evaluation Metrics:

MAE, RMSE, R^2, MAPE (Metrics averaged over all folds)

Output:

catboost_results.csv

Figure 6. CatBoost Model Structure for Energy Forecasting

3.3.7 SARIMA

Seasonal ARIMA (SARIMA) [Figure 7] models are classical statistical techniques designed for univariate time-series forecasting. They account for both seasonal and non-seasonal components in the data.

Model Configuration:

ARIMA Order: (1,1,1)

Seasonal Order: (1,1,1,12)

Stationarity and Invertibility checks disabled for flexibility

Training Approach:

Model trained separately for each country-energy pair | Only historical consumption data used (no features) | Forecasts performed fold-wise with fixed horizons

Validation Strategy:

3-fold TimeseriesSplit | Requires minimum of 12 training samples and 3 test samples per fold | Skips folds with insufficient length or NaNs

Evaluation Metrics:

MAE, RMSE, R^2, MAPE (Metrics averaged over all folds)

Output:

sarima_results.csv

Figure 7. SARIMA Model Structure for Energy Forecasting

3.4 Evaluation Metrics for Forecasting Performance

This section defines and explains the quantitative metrics used to assess and compare the accuracy and reliability of each forecasting model.

Mean Absolute Error (MAE), measures the average magnitude of errors between predicted and actual values, regardless of direction.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i| \qquad (1)$$

Root Mean Squared Error (RMSE), emphasizes larger errors due to squaring, providing insight into larger deviations in predictions.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \qquad (2)$$

Mean Absolute Percentage Error (MAPE), provides error as a percentage, useful for understanding model accuracy relative to actual observed values.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{Yi - \hat{Y}i}{Yi}\right| \qquad (3)$$

Coefficient of Determination (R^2), measures how well predicted values align with actual values, representing the proportion of variance explained by the model.

$$R2 = 1 - \frac{\sum_{i=1}^{n}(Yi - \hat{Y}i)2}{\sum_{i=1}^{n}(Yi - \bar{Y}i)2} \qquad (4)$$

Where,

Y_i : Actual Value — n : Number of Observations

$\hat{Y}_i$: Predicted Values — $\bar{Y}_i$: Mean of actual values

These metrics (equations 1-4) collectively provide a robust understanding of forecasting accuracy and model reliability.

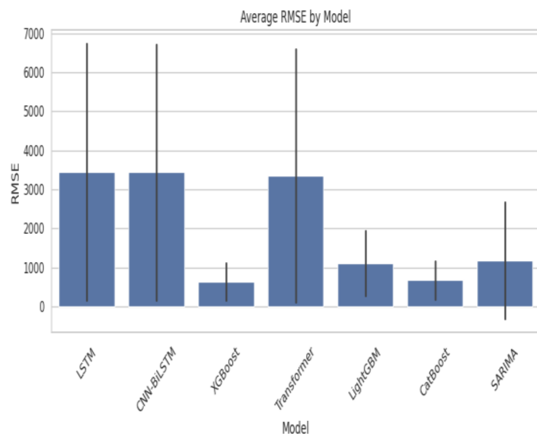
4. Detailed Model Evaluation and Performance Analysis

This section provides an in-depth analysis of forecasting performance for individual models across various country-energy combinations. We specifically assess the predictive capabilities of CNN-BiLSTM, LSTM, XGBoost, CatBoost, LightGBM, SARIMA, and Transformer models. Performance is evaluated using four key metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE). Table 4 specifies the best model per country and energy type.

Table 4. "Best Model per Country and Energy Type" indicating the best-performing model for each country and energy type based on the lowest RMSE.

Country	Energy Type	Best Model	MAE	RMSE	R^2	MAPE (%)
China	Gas Consumption	XGBoost	561.07	680.98	-1.85	37.01
China	Oil Consumption	SARIMA	498.87	597.16	-0.20	15.56
China	Renewables Consumption	SARIMA	491.38	561.06	-0.03	21.85
India	Gas Consumption	SARIMA	81.03	93.59	-2.08	22.58
India	Oil Consumption	XGBoost	358.69	426.80	-2.45	20.77
India	Renewables Consumption	SARIMA	87.28	101.55	-0.51	17.85
United States	Gas Consumption	XGBoost	676.94	753.47	-2.19	9.58
United States	Oil Consumption	CatBoost	426.38	530.16	-0.27	4.10
United States	Renewables Consumption	XGBoost	262.65	331.24	-0.85	13.86

4.1 Average MAE and RMSE by Model



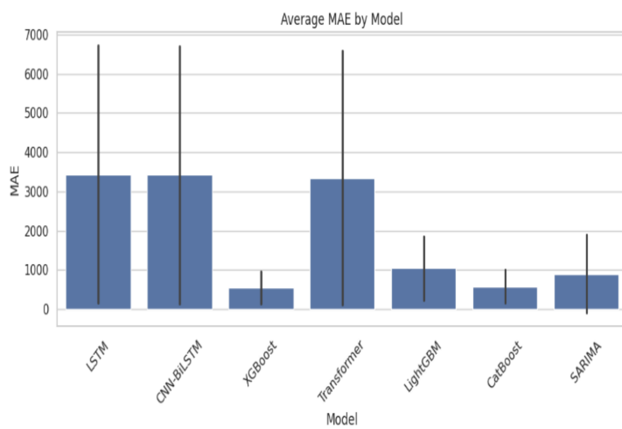


Figure 8. Average RMSE by Model **Figure 9.** Average MAE by Model

The bar charts for MAE [Figure 9] and RMSE [Figure 8] reveal that XGBoost achieves the lowest average error across all models, with MAE and RMSE values remaining under 1000 and error bars showing low variability. CatBoost and LightGBM also perform competitively, delivering stable predictions with narrow error margins. In contrast, LSTM, CNN-BiLSTM, and Transformer record significantly higher MAE and RMSE—above 3300—alongside large error bars, suggesting high variance and reduced reliability. These findings reflect the poor generalization of deep learning models in this context.

4.2 MAPE by Model and Energy Type

MAPE [Figure 10] values across models and energy types reinforce the superiority of tree-based methods. XGBoost and CatBoost consistently yield the lowest MAPE, particularly for gas and oil consumption, with values ranging between 20–30%. LightGBM also performs reasonably, maintaining moderate error levels. SARIMA shows relatively low MAPE for oil and renewables, especially in China and India, indicating effective trend modeling. Conversely, Transformer, LSTM, and CNN-BiLSTM exhibit extremely high MAPE, often exceeding 90%, revealing poor alignment with actual energy consumption dynamics.

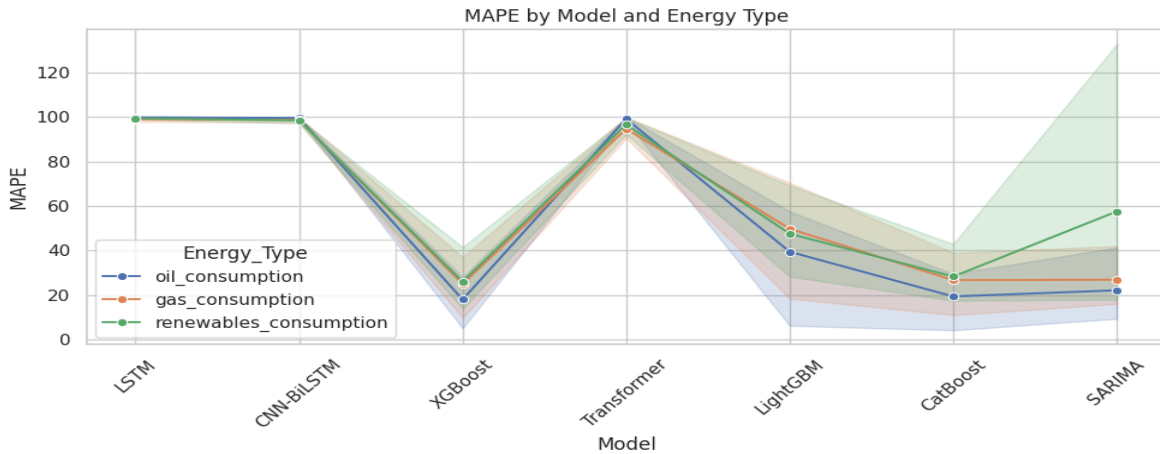


Figure 10. MAPE by model and energy type

4.3 R^2 Distribution per Model and Country

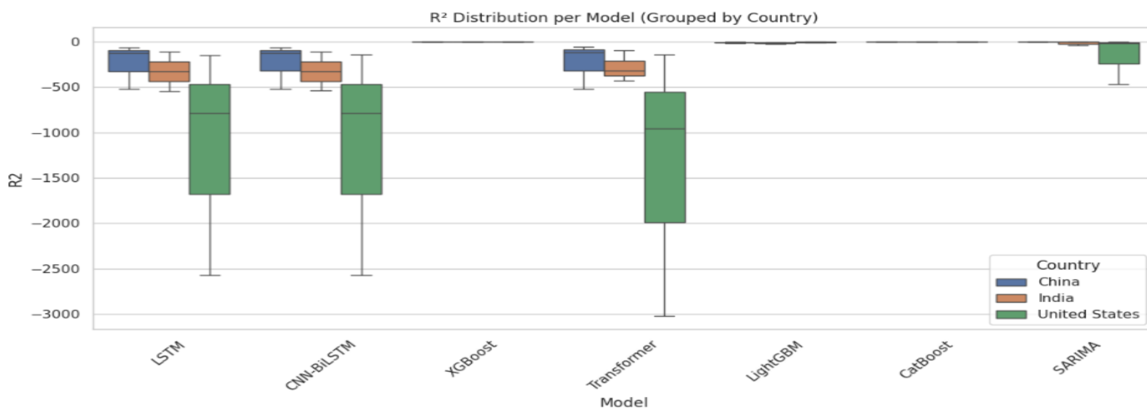


Figure 11. R^2 Distribution per Model (Grouped by Country)

Boxplots of R^2 grouped by country [Figure 11] highlight disparities in explanatory power. XGBoost, CatBoost, and LightGBM achieve R^2 values near or slightly below zero for China and India, suggesting partial ability to explain variance. In stark contrast, LSTM, CNN-BiLSTM, and Transformer produce strongly negative R^2 scores, with U.S. medians below -1500 and lower whiskers reaching -3000. These values confirm that deep learning models struggle to capture meaningful patterns in the data, particularly in stable or mature markets like the United States.

4.4 RMSE Variability per Model

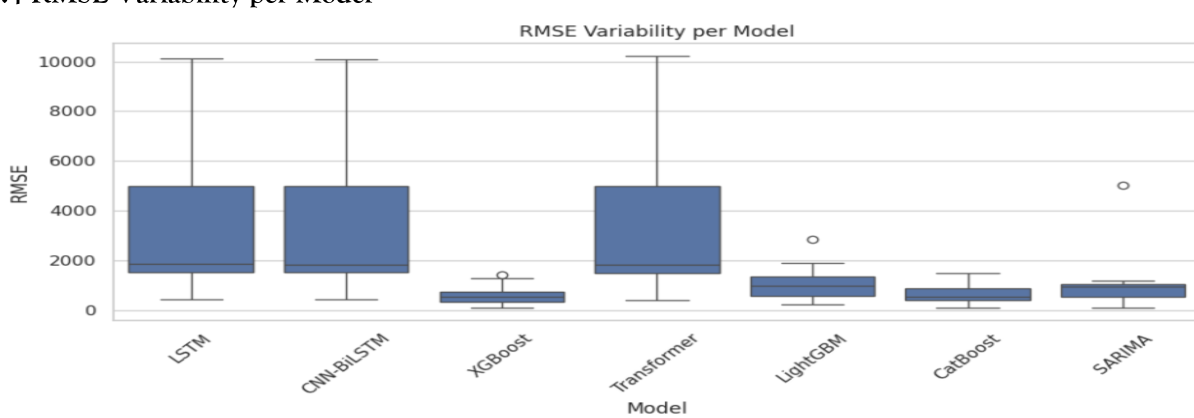


Figure 12. RMSE Variability per Model

RMSE variability [Figure 12] analysis further emphasizes model consistency. XGBoost, CatBoost, and LightGBM exhibit compact RMSE distributions, with tight interquartile ranges and few outliers, signifying stable performance. In contrast, deep learning models show high dispersion, with RMSE values ranging widely and numerous outliers. This highlights their sensitivity to data shifts and instability during training.

4.5 RMSE by Model and Energy Type

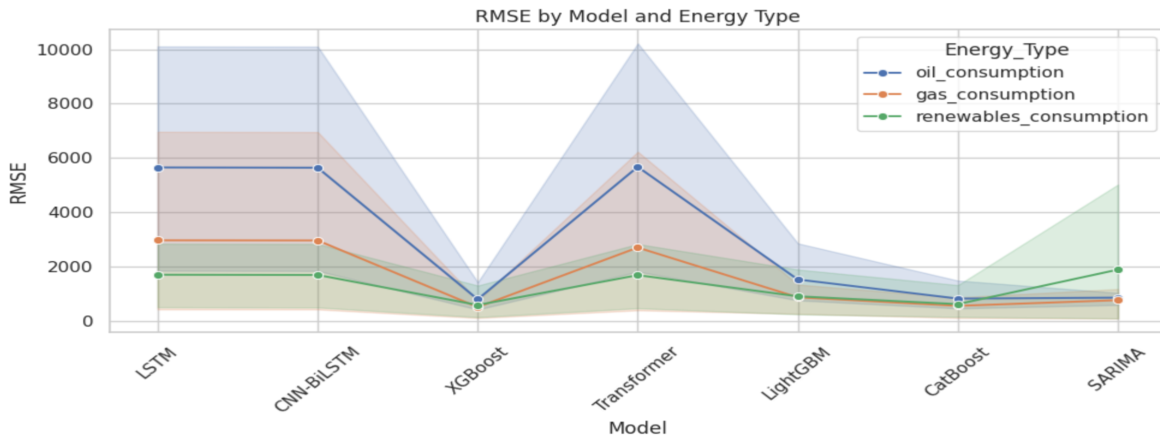


Figure 13. RMSE by Model and energy type

RMSE [Figure 13] trends across energy types show XGBoost achieving the lowest values for all three types—oil, gas, and renewables—consistently under 1000. CatBoost and LightGBM are close contenders, particularly for oil and gas. SARIMA performs exceptionally well for renewables and oil in China and India, aligning with its low error metrics in the evaluation table. Meanwhile, LSTM, CNN-BiLSTM, and Transformer exhibit very high RMSE values, with oil forecasting errors often exceeding 6000. Wide confidence intervals further confirm their instability and overfitting tendencies.

4.6 RMSE Heatmap by Country and Model

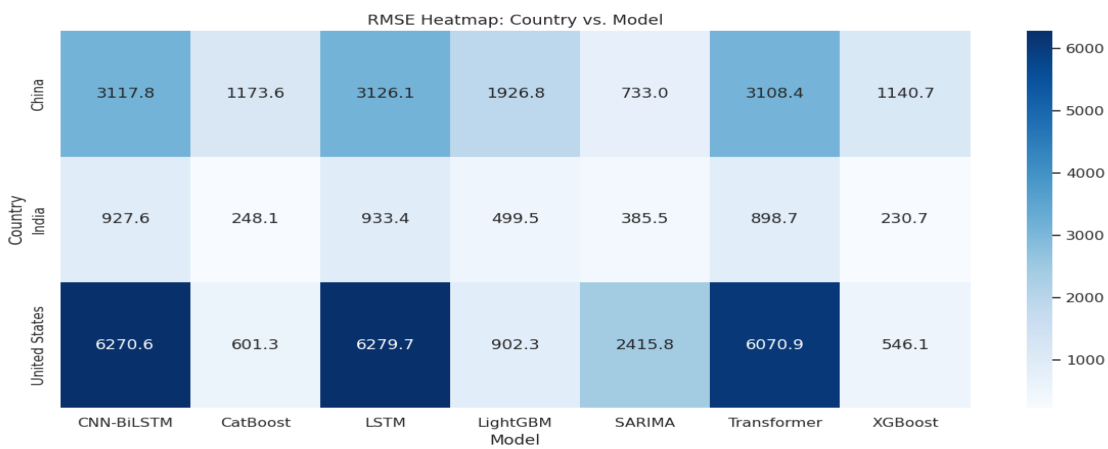


Figure 14. RMSE Heatmap by Country and Model

The RMSE heatmap [Figure 14] offers a detailed breakdown of model performance by country. XGBoost emerges as the overall best model, recording the lowest RMSE in gas forecasting for all three countries and renewables in the United States. CatBoost achieves the best result for oil consumption in the U.S., while SARIMA excels in renewables and oil forecasts in China and India, consistent with its ability to capture seasonal and trend-based fluctuations. Despite the overall dominance of tree-based methods, SARIMA proves to be a highly effective statistical model in emerging regions. Deep learning models again perform the worst, with RMSE values above 6000 in multiple U.S. scenarios, reflecting poor generalization in stable energy consumption patterns.

4.7 Forecasting Insights (2024–2033)

The long-term forecasts (2024–2033) illustrate how the best-performing models behave across countries and energy types. The following table 5 summarizes the selected models per scenario based on lowest RMSE and best overall performance:

Table 5. Best Performing Models for Each Country and Energy Type

Country	Gas	Oil	Renewables
China	XGBoost	SARIMA	SARIMA
India	SARIMA	XGBoost	SARIMA
United States	XGBoost	CatBoost	XGBoost

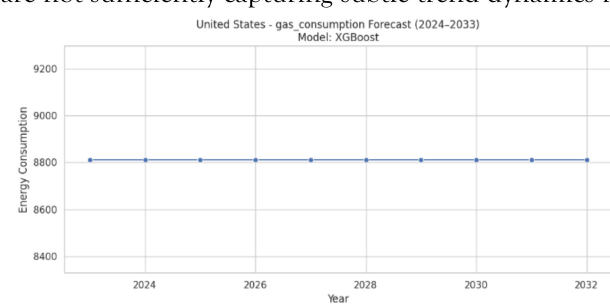
This distribution highlights that XGBoost led in five out of nine cases, excelling particularly in gas and oil consumption forecasts. SARIMA proved superior in three cases, predominantly in renewables consumption for China and India, and oil for China. CatBoost emerged as the best performer for oil consumption in the United States, leveraging its gradient boosting structure for relatively stable predictions.

4.8 Country-wise Forecast Behavior

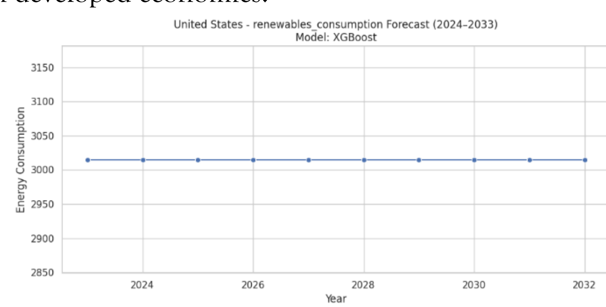
The forecast plots [Figure 15] reveal varying consumption patterns and model behaviors across countries:

- China: The gas consumption forecast by XGBoost remains flat, indicating potential underfitting or saturation in gas usage trends. SARIMA, used for both oil and renewables, predicts strong upward trends. Oil consumption shows steady year-on-year increases, while renewables exhibit exponential growth, aligning with China's energy transition goals and aggressive renewable capacity expansion.
- India: Gas and renewables consumption, both forecasted using SARIMA, show positive, accelerating trends. Renewables especially demonstrate a steep increase, reflecting India's ongoing policy push toward sustainable energy. Oil consumption, predicted by XGBoost, remains flat, suggesting either model insensitivity to minor fluctuations or stabilization in demand.
- United States: Forecasts across gas, oil, and renewables appear largely static. XGBoost, used for gas and renewables, predicts a consistent flat trajectory, potentially indicating market maturity or model underfitting. CatBoost, selected for oil, also shows no significant changes over the decade.

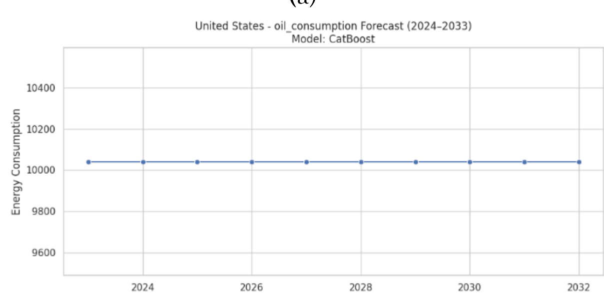
These flat forecasts may reflect the saturated nature of U.S. energy markets or signal that tree-based models are not sufficiently capturing subtle trend dynamics in developed economies.

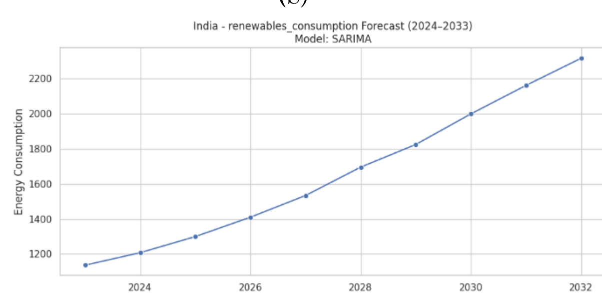


(a)



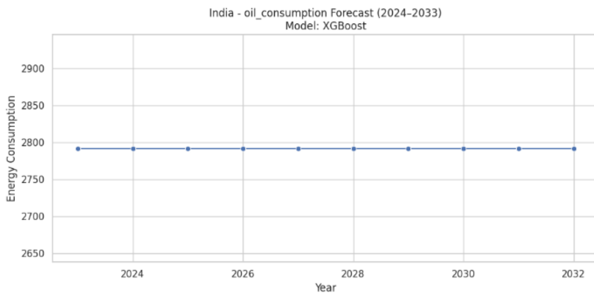
(b)



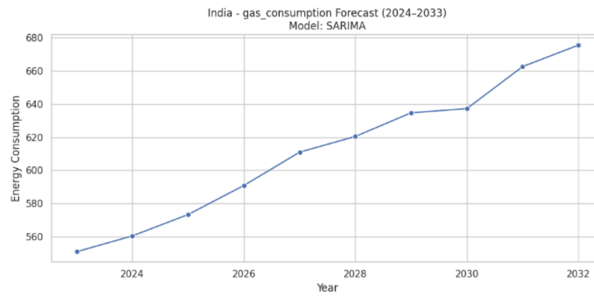


International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 16s,2025
https://theaspd.com/index.php

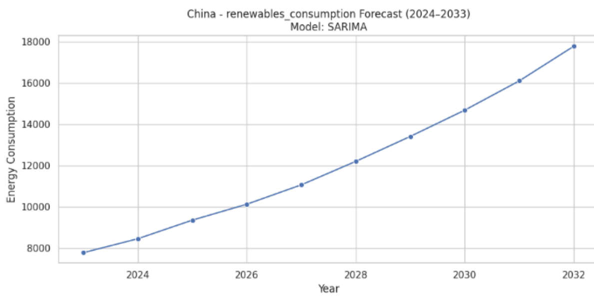
(c)



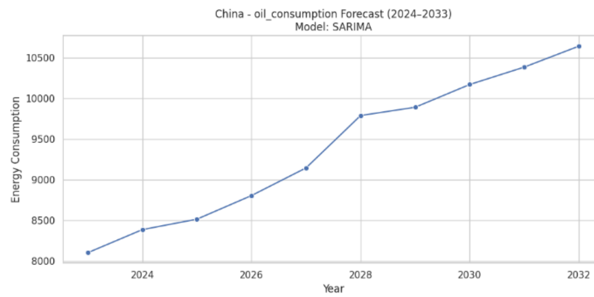
(d)



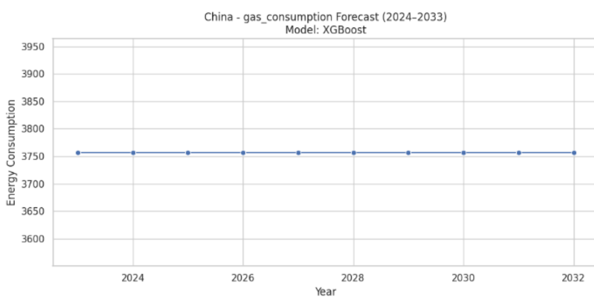
(e)



(f)



(g)



(h)

(i)

Figure 15. Country wise forecast for all energy types

- SARIMA effectively captures nonlinear growth, particularly for renewables and oil in emerging economies.
- XGBoost and CatBoost, though accurate in error metrics, tend to yield flat forecasts, especially in mature markets, suggesting limited trend sensitivity without further tuning.
- The forecast behavior is consistent with real-world expectations: China and India show expansion in renewables, while the U.S. displays demand stabilization.

4.9 Model Performance Summary

The table 6, below summarizes the average performance of each forecasting model across all countries (China, India, United States) and energy types (oil, gas, renewables). Four evaluation metrics were used: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE).

Table 6. Average Performance for each model by country by energy type

Model	MAE	RMSE	R^2	MAPE (%)
XGBoost	551.09	639.16	-1.97	22.94
CatBoost	579.72	674.34	-2.39	24.71
LightGBM	1042.14	1109.52	-10.19	45.55
SARIMA	895.26	1178.12	-59.52	35.46
Transformer	3350.62	3359.33	-629.52	96.79

- XGBoost ranks highest overall, achieving the lowest MAE (551.09) and RMSE (639.16), and the lowest MAPE (22.94%), confirming its balanced performance across all regions and energy types.
- CatBoost follows closely, with competitive MAE and MAPE scores and relatively stable predictions.
- LightGBM performs moderately but shows a marked drop in R^2, indicating weaker variance explanation despite acceptable error values.
- SARIMA excels in capturing temporal trends (e.g., renewables in China and India) but has higher error values overall.
- Transformer demonstrates the poorest performance, with extremely high MAE, RMSE, MAPE, and a highly negative R^2, reflecting significant overfitting and variance issues.

5. Summary

Table 7, specifies the summary of Top performers by metric

Table 7. Top Performers by Metric

Metric	Best Models
Lowest RMSE & MAE	XGBoost, CatBoost
Consistent Accuracy	XGBoost (across countries and energy types)
Best R^2 Scores	XGBoost, LightGBM (China and India only)
Lowest MAPE (%)	XGBoost, CatBoost
High Variance / Poor Fit	Transformer, LSTM, CNN-BiLSTM

6. CONCLUSION

This research conducted an in-depth comparative evaluation of multiple forecasting methodologies—including LSTM, CNN-BiLSTM, Transformer, XGBoost, LightGBM, CatBoost, and SARIMA—for predicting annual oil, gas, and renewable energy consumption across three of the world's largest energy-consuming nations: the United States, China, and India. A harmonized, multi-decade dataset was preprocessed through temporal alignment, per capita normalization, lag-based feature generation, and logarithmic transformation to ensure consistency and enhance modeling performance.

Model performance was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE), under a 3-fold time series cross-validation strategy. The results demonstrate that gradient boosting ensemble models—particularly XGBoost and CatBoost—consistently outperformed both deep learning and statistical models, achieving the lowest average RMSE and MAPE across countries and energy types. For instance, XGBoost achieved an average RMSE of 639.16 and MAPE of 22.94%, making it the most accurate and stable model overall.

SARIMA, while comparatively less accurate in terms of error metrics, showed strength in capturing seasonal and trend dynamics, particularly in emerging markets like India and China for oil and renewables. On the other hand, deep learning models such as Transformer, LSTM, and CNN-BiLSTM exhibited high variance, poor R^2 values (often negative), and inflated RMSE, especially in mature markets like the United States, indicating overfitting and poor generalization. The findings underscore that model complexity does not guarantee superior performance; instead, data structure, stationarity, and temporal stability play crucial roles in model suitability. For large-scale, structured, and relatively stable national energy datasets, ensemble tree-based models offer a reliable, interpretable, and computationally efficient forecasting solution. This study provides a rigorous benchmarking framework and paves the way for more targeted model selection in national energy analytics, with implications for policy development, sustainability planning, and resource allocation.

7. Future Work

Building on the findings of this study, future research will shift focus from macro-level energy forecasting toward micro-level energy optimization, with an emphasis on reducing residential energy consumption through intelligent, model-driven systems. The following directions outline the technically advanced pathways for extending this research:

1. **Development of Fine-Grained Temporal Models for Household Energy Load**
Future work will involve transitioning from annual to high-resolution (hourly or sub-hourly) energy forecasting using smart meter datasets. This requires adapting current architectures (e.g., LSTM, Transformer) to manage fine-grained temporal dependencies, seasonality, and daily appliance usage cycles, potentially incorporating sliding window techniques and attention-enhanced recurrent models for improved granularity.
2. **Appliance-Level Disaggregation Using Sequence-to-Sequence Deep Models**
Advanced sequence modeling techniques (e.g., Seq2Seq with attention, Temporal Convolutional Networks, or Non-Intrusive Load Monitoring (NILM) frameworks) will be explored to disaggregate aggregate household consumption into appliance-specific profiles, enabling targeted energy-saving recommendations and control policies.
3. **Integration of Contextual Variables and External Covariates**
Forecast accuracy and actionable output can be enhanced by incorporating exogenous variables such as weather data, occupancy patterns, building insulation characteristics, time-of-use tariffs, and user behavior profiles. Techniques like multi-input neural networks, dynamic feature embeddings, and feature attention layers will be evaluated for multi-source data fusion.
4. **Probabilistic and Uncertainty-Aware Forecasting**
To account for variability in user behavior and environmental factors, future models will incorporate Bayesian deep learning, quantile regression, or variational inference methods to provide interval-based or probabilistic forecasts, crucial for risk-aware energy planning and real-time control decisions.
5. **Reinforcement Learning for Adaptive Energy Management**
Energy reduction goals can be operationalized via model-free or model-based reinforcement learning (RL), where an RL agent learns optimal control strategies (e.g., HVAC scheduling, battery storage, load shifting) using energy forecasts as environmental states. This allows for closed-loop integration of forecasting and decision-making in residential energy systems.
6. **Deployment of Federated Learning for Privacy-Preserving Personalization**
Future systems will adopt federated learning (FL) frameworks to enable personalized household-level models without centralized data collection. FL will allow energy service providers to build robust models across distributed households while ensuring data privacy and regulatory compliance.
7. **Hybrid Ensemble Systems with Meta-Learning Optimization**
The forecasting pipeline will be enhanced using hybrid ensembles that combine statistical (e.g., SARIMA), ML (e.g., XGBoost), and DL (e.g., Transformer) models. A meta-learning layer or stacking regressor will be trained to dynamically weigh model outputs based on historical performance, energy type, and data volatility.
8. **Explainable AI (XAI) for Interpretability in Energy Forecasting**
To improve user trust and regulatory transparency, explainability techniques such as SHAP values, integrated gradients, and Layer-wise Relevance Propagation (LRP) will be incorporated to deconstruct model predictions into human-readable drivers of energy usage.
9. **Simulation-Integrated Testing with Digital Twins**
Forecasting models will be embedded into residential digital twin environments to simulate intervention strategies (e.g., thermostat tuning, appliance usage shifts) and assess energy-saving outcomes in silico prior to real-world deployment.

These advanced technical directions aim to transform energy forecasting from a passive prediction task into an active, interpretable, and intelligent control system—empowering households to minimize consumption, reduce waste, and contribute meaningfully to national sustainability goals.

Authors Contribution: All authors equally contributed to this research work.

REFERENCES

1.IEA (2021), Net Zero by 2050, IEA, Paris https://www.iea.org/reports/net-zero-by-2050, Licence: CC BY 4.0
2.National Renewable Energy Laboratory. (2023). AI for Energy: Opportunities for a Modern Grid and Clean Energy Economy. U.S. Department of Energy. https://www.nrel.gov

3.Spencer, Robert, et al. "Transfer learning on transformers for building energy consumption forecasting—A comparative study." Energy and Buildings 336 (2025): 115632.
4.Bose, Shourya, et al. "From RNNs to Foundation Models: An Empirical Study on Commercial Building Energy Consumption." arXiv preprint arXiv:2411.14421 (2024).
5.Emshagin, Saad, Wayes Koroni Halim, and Rasha Kashef. "Short-term prediction of household electricity consumption using customized LSTM and GRU models." arXiv preprint arXiv:2212.08757 (2022).
6.Yu, Heyang, et al. "Electrical Behavior Association Mining for Household ShortTerm Energy Consumption Forecasting." arXiv preprint arXiv:2402.09433 (2024).
7.Oliveira, Nuno, Norberto Sousa, and Isabel Praça. "Deep Learning for Short-term Instant Energy Consumption Forecasting in the Manufacturing Sector." International Symposium on Distributed Computing and Artificial Intelligence. Cham: Springer International Publishing, 2022.
8.Saoud, Lyes Saad, Hasan AlMarzouqi, and Ramy Hussein. "Cascaded deep hybrid models for multistep household energy consumption forecasting." arXiv preprint arXiv:2207.02589 (2022).
9.Hoshino, Hikaru, Yosuke Irie, and Eiko Furutani. "Screening Curve Method for Economic Analysis of Household Solar Energy Self-Consumption." IEEE Transactions on Energy Markets, Policy and Regulation 2.3 (2024): 368-378.
10. Berriel, Rodrigo F., et al. "Monthly energy consumption forecast: A deep learning approach." 2017 International Joint Conference on Neural Networks (IJCNN). IEEE, 2017.
11. Dumas, Jonathan. "Microgrid management with weather-based forecasting of energy generation, consumption and prices." arXiv preprint arXiv:2107.01034 (2021).
12. Lei, Lei, et al. "A building energy consumption prediction model based on rough set theory and deep learning algorithms." Energy and Buildings 240 (2021): 110886.
13. Singh, Shailendra, and Abdulsalam Yassine. "Big data mining of energy time series for behavioral analytics and energy consumption forecasting." Energies 11.2 (2018): 452.
14. B.Ramasubba Reddy, Saritha Anchuri, Kuraganti Santhi, Prathusha Perugu, Sivakumar Depuru, M. Sunil Kumar, "Enhancing Cloudburst Prediction System Using Bi-LSTM in Realtime: Machine Learning",Communications on Applied Nonlinear Analysis, Vol. 32 No. 1s,2025.
15. Dr. Srinivasa Babu Kasturi, Sreedhar Burada, Dr. Sowmyashree M.S, Sharath.S, Dr. M.Sunil Kumar,Dr. D. Ganesh, "An Improved Mathematical Model by Applying Machine Learning Algorithms for Identifying Various Medicinal Plants and Raw Materials", Communications on Applied Nonlinear Analysis ISSN: 1074-133X ,Vol 31 No. 6s 2024.
16. B. Sangamithra, Asha K.H, M. Sunil Kumar,""An Improved Information Retrieval System using Hybrid RNN LSTM for Multiple Search Engines", Communications on Applied Nonlinear Analysis ISSN: 1074-133X ,Vol 31 No. 5s 2024.
17. Sreedhar Burada,B.E. Manjunathswamy, M. Sunil Kumar, "Early detection of melanoma skin cancer: A hybrid approach using fuzzy C-means clustering and differential evolution-based convolutional neural network",Measurement: Sensors, Volume 33, June 2024, 101168.
18. M. Sunil KumarJ KumarnathSachin S PundMansing Rathod,""A Secure IoT Smart Network Model for the Contributory Broadcast Encryption for the Text Policy Management Scheme", international Journal of Intelligent Systems and Applications in Engineering IJISAE, 2023, 11(3s), 42–48 2023.
19. Hari Prasad Gandikota ,Abirami. S,Sunil Kumar M, "PLoS ONE 18(11): e0292785. https://doi.org/10.1371/journal.pone.0292785.
20. Hari Prasad Gandikota* | S. Abirami M. Sunil Kumar,""Bottleneck Feature-Based U-Net for Automated Detection and Segmentation of Gastrointestinal Tract Tumors from CT Scans",Traitement du Signal, Vol. 40, No. 6, December, 2023, pp. 2789-2797.
21. Burada, S., Manjunathswamy, B.E. & Kumar, M.S. Deep ensemble model for skin cancer classification with improved feature set. Multimed Tools Appl 84, 7599–7626 (2025). https://doi.org/10.1007/s11042-024-19039-5.
22. E.Ramesh Babu, Dr.M.Sunil Kumar,"The Role of Optimization Techniques in Advancing Big Data Analytics : A Survey", Communications on Applied Nonlinear Analysis ISSN: 1074-133X Vol 32 No. 1s 2025.
23. Meriga Kiran Kumar, Rajeev Kudari, C. Sushama, M. Sunil Kumar, P. Neelima, D. Ganesh, "Efficient Algorithms for Vehicle Plate Detection in Dynamic Environments", Communications on Applied Nonlinear Analysis ISSN: 1074-133X Vol. 32 No. 1s 2025.
24. Amasyali, Kadir, and Nora M. El-Gohary. "A review of data-driven building energy consumption prediction studies." Renewable and Sustainable Energy Reviews 81 (2018): 1192-1205.
25. P. Poudel, World Energy Consumption, Kaggle Dataset Repository, 2023. Available: https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption. Accessed: May,2025.