Reliable renewable energy forecasting for climate change mitigation

Walid Atwa[1], Abdulwahab Ali Almazroi[1] and Nasir Ayub[2]

[1] College of Computing and Information Technology at Khulais, Department of Information Technology, University of Jeddah, Jeddah, Saudi Arabia
[2] Department of Creative Technologies, Air University, Islamabad, Pakistan

ABSTRACT

Accurate prediction of electricity generation from diverse renewable energy sources (RES) plays a pivotal role in optimizing power schedules within RES, contributing to the collective effort to combat climate change. While prior research often focused on individual energy sources in isolation, neglecting intricate interactions among multiple sources, this limitation frequently leads to inaccurate estimations of total power generation. In this study, we introduce a hybrid architecture designed to address these challenges, incorporating advanced artificial intelligence (AI) techniques. The hybrid model seamlessly integrates a gated recurrent unit (GRU) and a ResNext model, and it is tuned with the modified jaya algorithm (MJA) to capture localized correlations among different energy sources. Leveraging its nonlinear time-series properties, the model integrates meteorological conditions and specific energy source data. Additionally, principal component analysis (PCA) is employed to extract linear time-series data characteristics for each energy source. Application of the proposed AI-infused approach to a renewable energy system demonstrates its effectiveness and feasibility in the context of climate change mitigation. Results reveal the superior accuracy of the hybrid framework compared to more complex models such as decision trees and ResNet. Specifically, our proposed method achieved remarkable performance, boasting the lowest error rates with a normalized RMSE of 6.51 and a normalized MAPE of 4.34 for solar photovoltaic (PV), highlighting its exceptional precision in terms of mean absolute errors. A detailed sensitivity analysis is carried out to evaluate the influence of every element in the hybrid framework, emphasizing the importance of energy correlation patterns. Comparative assessments underscore the increased accuracy and stability of the suggested AI-infused framework when compared to other methods.

INTRODUCTION

The need for energy has grown dramatically to meet everyday human needs and activities in tandem with the social economy's fast expansion (*Semieniuk et al., 2021*). Power generation has increased significantly as a result of the rising reliance on energy. Traditional fossil fuels have triggered significant environmental harm, contributing to issues like air pollution and global warming (*Hassan et al., 2021*). The release of greenhouse gases from these fuels intensifies climate change. To counteract these challenges, there is a growing emphasis

Corresponding authors
Walid Atwa, wsatwa@uj.edu.sa
Nasir Ayub, nasir.ayubse@gmail.com

on transitioning to RESs, driven by the advancement of green technologies and the enactment of stringent "zero carbon emission" regulations (*Ebhota & Tabakov, 2021*). This shift towards green development has become integral to people's lives. Many nations are actively investing in the development of RESs, utilizing sustainable energy sources to generate electricity.

As per the International Energy Agency (IEA) (*Kober et al., 2020*), it is projected that RESs will represent approximately 70% of all new power production capacity by 2040. This underscores a global commitment to transitioning away from conventional energy sources towards more sustainable and environmentally friendly alternatives. The smart microgrid idea has led to a substantial expansion in the application of RESs due to their numerous benefits, which include minimum environmental pollution, security of the energy supply, and the realization of sustainable development. RESs can be utilized to provide all or part of a region's power demands; they typically comprise of loads, energy generation, and energy storage components. For both individual producers and system operators in RESs, estimating the quantity of electricity produced from renewable sources is essential.

As an illustration, power plants must submit transmission deals to the australian national electricity market up to forty hours beforehand, with the option to amend the offers up to five minutes before to the dispatch (*Simshauser, 2021*). Additionally, precise forecasting of electricity generation from RESs holds the potential to enhance the efficiency of power dispatch, optimize the scheduling of power resources, and contribute to increased economic gains for energy-related enterprises. This, in turn, facilitates improved operational coordination within power systems, ensuring the security and reliability of energy supplies.

Biogas, biofuel (*Rahman, Farrok & Haque, 2022*), geothermal energy, small-scale hydroelectricity, solar photovoltaic, thermal solar, wind, and solar thermal are examples of RESs that may be used in RESs. To improve the multi-energy generation prediction's accuracy, it is important to investigate how multiple energy sources coordinate (*Ang et al., 2022*). The weather has a significant impact on energy generation forecast, particularly for wind and solar power, which puts the smart microgrid's electricity scheduling at risk. Electricity may be produced from solar radiation using PV modules or concentrated solar thermal power and wind energy from wind farms. Solar thermal power can only be produced by direct normal irradiance, whereas PV electricity can be produced by worldwide horizontal irradiance (*Law et al., 2014*). With several wind sources, an increasing number of wind farms have been established, particularly in America, China, and Germany (*Sahu, Hiloidhari & Baruah, 2013*).

The variability in weather conditions contributes to the unpredictability of both wind and solar power generation, establishing a spatial–temporal correlation between them. Accurate solar and wind energy forecasting reduces the risk of blackouts or even outages in the system and decreases the price of energy balancing, which is the basis for RESs' capacity to anticipate the supply of power (*Tu et al., 2021*). This article studies the power generation projection in a RES while accounting for the local correlations of numerous sources of renewable energy in order to reduce uncertainty in the electricity supply.

Motivation

The essential need to strengthen the stability and dependability of RESs, particularly in the context of growing environmental concerns and emergency scenarios, is the basis behind this study. A revolutionary transition to sustainable energy resources is required in order to address climate change and reduce our dependence on traditional fossil fuels (*Kabeyi & Olanrewaju, 2022*). With the ability to decrease the effects of climate change and guarantee a sustainable energy future, RESs, provide a promising path forward for clean and green energy (*Al-Shetwi, 2022*). Robust and adaptableRESs are even more crucial in emergency scenarios, such as emergencies or disruptions to traditional energy networks. Precise forecasting of renewable energy output is essential for daily operations as well as for quickly and efficiently handling unanticipated events (*Alkabbani et al., 2021*). This study aims to combine green and nature-inspired approaches into renewable energy forecasting, driven by the critical need to develop resilient energy systems that can resist disasters and aid in disaster recovery efforts.

Contribution

By addressing important issues and presenting novel approaches, our study significantly advances the area of renewable energy forecasting. The following are this work's primary technical contributions:

1. Hybrid architecture for projecting renewable energy: We suggest a unique hybrid design to capture intricate interactions between various RESs, combining ResNext (*Pant, Yadav & Gaur, 2020*) and GRU (*Wang, Liao & Chang, 2018*) models. In contrast to other methods that frequently concentrate on estimating distinct energy sources independently, our hybrid model takes into account the localized correlations between many sources. Because of the complex interrelationships between various energy components, this design yields more precise estimates of overall power generation.

2. Implementing PCA for feature extraction: We use PCA for feature extraction in order to increase the dataset's temporal dimension and interpretability. Effective feature selection is aided by PCA, which decreases the dimensionality of the data while keeping important information. Improved model performance is made possible by this contribution, which makes it possible to express linear time-series features more effectively.

3. Rigorous performance evaluation standards: Using rigorous measures such as R-squared, explained variance, mean absolute error (MAPE), root mean squared error (RMSE), and RMS logarithmic error, we conduct a thorough performance review. This comprehensive examination guarantees a detailed appraisal of the model's precision and accuracy, enabling a trustworthy comparison with the most recent versions. A more complex understanding of the model's predicting ability is made possible by the use of several measures.

4. Superior predictive performance for hourly energy generation: Our proposed ResNeXt-GRU-MJA model outshines existing state-of-the-art methods by demonstrating a remarkable 15% improvement in hourly forecasting accuracy. This superior performance is substantiated through normalized measures of RMSE, MAPE, and

MAE specifically tailored for hourly predictions. The model's effectiveness in capturing short-term energy variations positions it as a valuable tool for applications requiring precise hourly forecasting.

Article organization

The article is structured into several sections. 'Related Work' provides a systematic review of existing methods developed by researchers. In 'System Model', we present the innovative approach proposed to enhance the current work. 'Simulation and Results' delves into the experimental simulation, offering insights into the practical implementation of the methodology. The conclusion is then presented to summarize key findings and contributions.

RELATED WORK

In recent research efforts, a multitude of studies have been conducted to predict renewable energy output. However, a predominant trend has emerged, with a majority of these studies concentrating on individual energy sources, such as solar or wind power. Many researchers have developed large-scale solar power facilities since solar energy is one of the most promising sources. Some research projected solar irradiance instead of electricity generation because they used pricey solar irradiance meters by examining the interaction relationship between solar irradiance and the sky picture. *Huang et al. (2021)* introduced a solar irradiance forecasting approach aimed at enhancing the precision of solar power predictions. Likewise, *Ghimire et al. (2023)* presented a hybrid predictive model incorporating a multilayer perceptron (MLP) model and a convolution-based neural network. This model utilizes sky images to forecast global irradiation 15 min in advance. *Matrenin et al. (2023)* described a pipeline for a one-day ahead-of-time forecast of solar radiation and heat. It is based on four data-driven prediction steps and the imputation of past data. *Michael et al. (2022)* created a unique multivariate hybrid deep neural model that takes climatic effect into account when estimating sun irradiance one hour in advance. Real data from four different nations was used to validate the model. *Alkhayat, Hasan & Mehmood (2022)* carried out a comprehensive and in-depth investigation of the machine learning-based solar power prediction methods in order to address the inadequacies of the existing machine learning models and boost prediction accuracy.

Statistical and machine learning methods, such as decision trees (*Mahmud et al., 2021*), random forest (*Liu & Sun, 2019*), and time series ensembles (*AlKandari & Ahmad, 2020*), are employed for PV power forecasting as a result of the significant advancements in big data technology and measuring meters that store enormous amounts of data. To address the shortcomings of conventional AI modeling techniques, deep learning-based approaches significantly increase the forecast accuracy of PV power (*Wang, Qi & Liu, 2019*). The original PV power series was divided into four distinct neural network networks and sub-series using a hybrid deep learning model created by *Khan, Walker & Zeiler (2022)* that was based on wavelet packet decomposition. In the work presented by *Gu et al. (2021)*, an innovative model is introduced, specifically designed for forecasting day-ahead PV power. This model incorporates principles dependent on time and integrates deep learning

modeling into a framework for partially daily forecasting trends, serving as a guide for parameter modifications. The investigation conducted by *Ahmed et al. (2020)* involved a thorough examination and assessment of various modern forecasting techniques.

This comprehensive analysis encompassed considerations such as the forecast horizon, temporal aspects including date and time, scrutiny of input connections, preprocessing and postprocessing of data, optimization of network parameters, classification of weather patterns, estimation of uncertainties, and evaluations of overall efficacy. A few studies took into consideration the solar production of more than one location. In an earlier study, *AlSkaif et al. (2020)* presented an improved LSTM network for multi-region solar power output prediction that was optimized using particle swarm optimization. This strategy was applied to a genuine geographic region in Asia, demonstrating its efficacy. In a similar vein, another study (*Nourani et al., 2022*) examined the climatic and geographical details of multi-region solar production. In order to create an ideal ensemble prediction model, taking into account a variety of candidate characteristics, this study used automated machine learning. A genetic algorithm was used to identify the most appropriate parameters for the predictive model, which contributed to enhancing the feature selection process.

Despite the significance of solar thermal energy as a vital component of solar energy, there has been limited attention devoted to its forecasting in comparison to solar radiance or PV power output prediction in the existing body of research. The majority of studies have predominantly concentrated on predicting solar radiance levels or forecasting the output of PV systems, leaving solar thermal energy forecasting relatively understudied. Using a hybrid method based on deep learning and mechanism modeling for solar thermal prediction, *Hu et al. (2021)* developed a way to connect the spatial–temporal aspects between meteorological parameters and identify the key meteorological components. *Dewangan, Singh & Chakrabarti (2020)* used an instance of a solar Fresnel generator to anticipate the solar heat output 24 h ahead of time by combining a climatic model with a solar plant's performance model.

Numerous studies have proposed effective strategies for forecasting wind speed and wind power, aiming to enhance the coordination of wind energy systems with power networks (*ShobanaDevi et al., 2021*). In the realm of wind energy forecasting, advancements have led to the categorization of forecasting models into two main types: deterministic prediction and uncertainty analysis. Deterministic prediction models can be broadly classified into physical and AI hybrid approaches. The physical approach, exemplified by numerical weather prediction (NWP), involves solving hydrodynamic and thermodynamic equations using computational methods (*Patel et al., 2022*). While physical models exhibit superior performance in long-term forecasting, they tend to underperform in short-term predictions and often require significant computational resources. On the other hand, statistical models, a subset of deterministic prediction, rely on mathematical theories such as the kalman filter (KF), copula theory, and Bayesian multiple kernel regression (*Dong et al., 2022*). These models leverage statistical knowledge to make predictions and are known for their versatility in handling various forecasting scenarios.

Statistical approaches are commonly used for time series forecasting, but they work best for linear data or data that has simple relationships between variables. When dealing with

highly nonlinear data, their effectiveness reduces because they assume linearity. Although some statistical methods can be adjusted to accommodate nonlinear relationships to some extent, they may not perform as well as more advanced techniques designed specifically for nonlinear data, such as machine learning algorithms. Therefore, while statistical techniques can be used for nonlinear data, they may not provide the same level of accuracy or effectiveness as methods explicitly created for handling nonlinear relationships (*Almazroi & Ayub, 2021*).

Advanced learning machines and ANNs (*Jiang et al., 2021*) are two cutting-edge techniques that now surpass all others in the prediction of wind power and speed. In addition to the individual models, there has been a surge in the development of hybrid forecasting frameworks that outperform single models by amalgamating the strengths of multiple approaches. Notably, techniques such as decomposition and feature selection are specifically tailored for handling wind energy series (*Almazroi & Ayub, 2023*). Ensemble learning, on the other hand, takes a stride towards constructing a more robust model by incorporating numerous predictors. This trend underscores the pursuit of enhanced reliability and performance through the integration of diverse forecasting strategies. The related work is summarized in Table 1.

SYSTEM MODEL

To precisely forecast weekly and hourly renewable energy in home energy management (HEM), this study introduces a novel framework consisting of six main steps. Initially, upon dataset collection, various issues were identified. Subsequently, the dataset underwent preprocessing to address missing or negative values. PCA and correlation analysis were then applied. To gain a deeper understanding of the data, exploratory data analysis (EDA) was conducted, focusing on pairwise relationships and the impact of selected features on power generation.

Following EDA, lag features were created, and feature normalization using Min-Max scaling was performed in preparation for regression analysis. Data scaling preceded the division into training and validation sets. The ResNeXt-GRU model was employed, and its parameters were fine-tuned using the MJA. Additionally, the effectiveness of the proposed model was compared against state-of-the-art algorithms, including ARIMA, CNN, VGG, NB, DenseNet, and decision trees.

For performance evaluation, various metrics such as RMSE, RMS LOG error, explained variance, R-square, and MAPE were employed. Statistical analysis was also conducted. The overall structure of the system model is illustrated in Fig. 1.

Table 1 Overview of related work on renewable energy forecasting.

Ref	Problem addressed	Methodology employed	Achievements	Limitations/drawbacks
Huang et al. (2021)	Enhanced precision of solar power predictions through solar irradiance forecasting	Solar irradiance forecasting approach	Improved accuracy in solar power predictions	Dependence on expensive solar irradiance meters
Ghimire et al. (2023)	Short-term forecasting of global irradiation using sky images	Hybrid predictive model with MLP and convolution-based neural network	Forecast global irradiation 15 min in advance	Limited applicability to short-term predictions
Matrenin et al. (2023)	One-day ahead forecast of solar radiation and heat	Four data-driven prediction steps and past data imputation	Accurate forecast of solar radiation and heat	Requirement of extensive past data
Michael et al. (2022)	Sun irradiance prediction considering climatic effects	Multivariate hybrid deep neural model	Improved accuracy in sun irradiance estimation	Necessity for real data from diverse nations
Alkhayat, Hasan & Mehmood (2022)	In-depth investigation of machine learning-based solar power prediction methods	Comprehensive analysis of existing models	Addressing inadequacies and improving prediction accuracy	No specific achievement highlighted
Mahmud et al. (2021); *Liu & Sun (2019)*; *AlKandari & Ahmad (2020)*	PV power forecasting using statistical and machine learning methods	Decision trees, random forest, and time series ensembles	Advancements in big data technology	Challenges in handling highly nonlinear data
Wang, Qi & Liu (2019); *Khan, Walker & Zeiler (2022)*	Hybrid deep learning model for PV power forecasting	Division of original PV power series into distinct neural networks	Increased forecast accuracy of PV power	Complexity in model architecture
Gu et al. (2021)	Day-ahead PV power forecasting	Model with dependent-on-time principles and deep learning modeling	Partial daily forecasting of trends framework	Limited discussion on specific achievements
Ahmed et al. (2020)	Examination and assessment of diverse forecasting techniques	Thorough analysis covering various aspects	Comprehensive understanding of forecasting techniques	No specific achievement mentioned
AlSkaif et al. (2020)	Improved LSTM network for multi-region solar power output prediction	Optimization using particle swarm optimization	Demonstrated efficacy in a specific geographic region	Limited to a particular geographic context
Nourani et al. (2022)	Climatic and geographical details of multi-region solar production	Automated machine learning for creating an ensemble prediction model	Enhanced feature selection process	Challenges in utilizing automated machine learning
Hu et al. (2021)	Solar thermal prediction using a hybrid method	Deep learning and mechanism modeling	Spatial–temporal connection and identification of key meteorological components	Limited discussion on specific achievements
Dewangan, Singh & Chakrabarti (2020)	Solar heat output prediction for a solar Fresnel generator	Combination of climatic model and solar plant's performance model	Anticipation of solar heat output 24 h in advance	Specific to solar Fresnel generators
ShobanaDevi et al. (2021)	Forecasting wind speed and wind power for improved coordination with power networks	Various forecasting models categorized into deterministic prediction and uncertainty analysis	Enhanced coordination of wind energy systems with power networks	Limited discussion on specific achievements

Table 1 (*continued*)

Ref	Problem addressed	Methodology employed	Achievements	Limitations/drawbacks
Patel et al. (2022)	Deterministic prediction models for wind energy	Physical approach exemplified by Numerical Weather Prediction (NWP)	Superior performance in long-term forecasting	Demands significant computational resources
Dong et al. (2022)	Statistical models for wind energy forecasting	Mathematical theories like Kalman filter, copula theory, and Bayesian multiple kernel regression	Versatility in handling various forecasting scenarios	Limited to statistical approaches
Almazroi & Ayub (2021); *Jiang et al. (2021)*	Intelligent forecasting models for wind power	Advanced learning machines and Artificial Neural Networks (ANNs)	Greater accuracy than physical and mathematical methods	Applicability primarily to basic time series forecasting
Almazroi & Ayub (2023)	Hybrid forecasting frameworks for wind energy	Techniques like decomposition and feature selection	Enhanced reliability and performance through integration of diverse strategies	Limited discussion on specific achievements
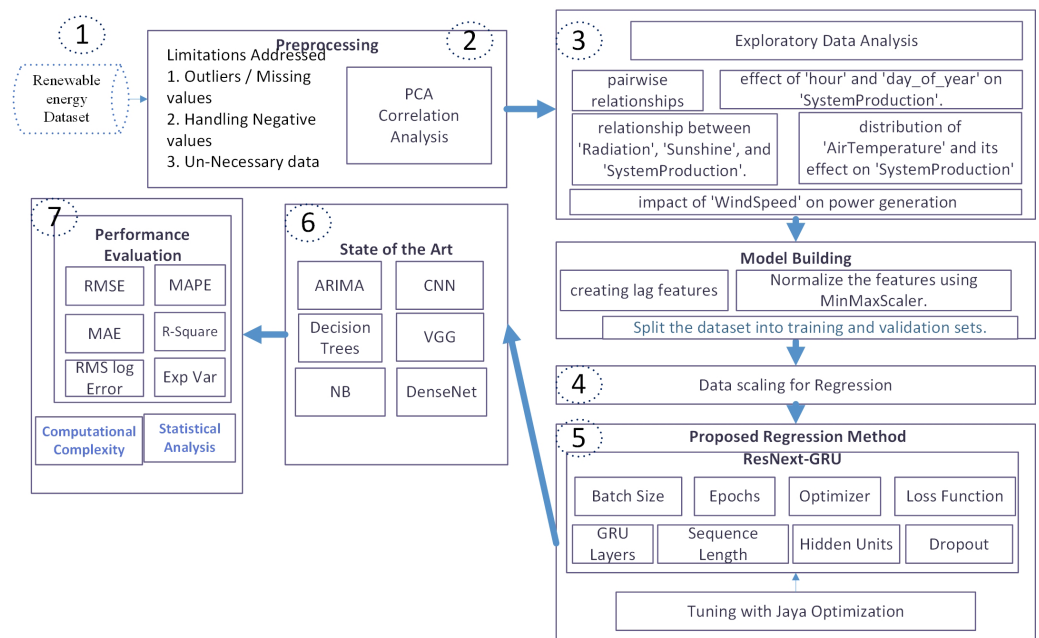
Figure 1 Proposed HEM system model.

Dataset collection and description

To ensure the sustainability of our data preparation processes, we logically separated the dataset into training and testing sets using an 80/20 ratio. This separation enables us to make the most use of the data by dividing it up and assigning a significant amount to model training and a separate chunk to performance assessment. We aim to maintain the integrity and dependability of the assessment procedure. The dataset utilized in this study is a combine of two datasets from Kaggle (*Afroz, 2023*) and *Anikannal (2023)*, totaling 8,760 occurrences. With the help of these repositories, researchers can utilize and analyze a

Table 2 Description of features in the dataset.

s#	Feature	Description
1	Sun-shine	Duration of sunshine
2	Date-Hour (NMT)	Date and time in NMT format
3	Radiation	Solar radiation intensity
4	Relative air humidity	Relative humidity of the air
5	System production	Solar power generation
6	Air pressure	Atmospheric air pressure
7	Air temperature	Temperature of the air
8	Wind speed	Speed of the wind in the environment

wide range of publicly available datasets for a variety of research applications. We actively encourage transparency and reproducible in our research by utilizing datasets from various sources, making it possible for others to validate or expand upon our findings. Table 2 provides a description of the applicable dataset.

Pre-processing

To provide reliable and precise outcomes, data preparation is essential (*ShobanaDevi et al., 2021*). In order to prepare the data for analysis and modeling, careful handling of outliers and missing values at this crucial stage, removal of unnecessary attributes, and utilization of feature extraction methods like PCA are performed. Since errors and outliers may significantly compromise the accuracy of the analytic results, fixing them is an essential part of the pre-processing stage. To preserve the integrity of the data, strong methods are employed to detect and manage missing values. This covers interpolation techniques like imputation based on regression or mean (Eq. (1)) or median (Eq. (2)). By using these strategies, the aim is to promote sustainable data practices by maintaining the dataset's completeness and reducing the impact of missing values on subsequent studies.

To handle missing values, two strategies are employed: mean imputation and median imputation. The mean imputation, expressed by Eq. (1) (*Josse & Husson, 2012*), estimates the missing value $\hat{x}_i$ by averaging the available data within the corresponding feature:

$$\hat{x}_i = \frac{\sum_{j=1}^{n} x_j}{n}.$$ (1)

Similarly, the median imputation, depicted in Eq. (2) (*Josse & Husson, 2012*) as Eq. (2), utilizes the median of the available values in the feature:

$$\hat{x}_i = \mathrm{median}(x_1, x_2, \ldots, x_n).$$ (2)

Outlier detection is then performed using the z-score (Eq. (3)) (*Josse & Husson, 2012*), allowing the identification and appropriate handling of outliers:

$$z = \frac{x - \mu}{\sigma}.$$ (3)

In this case, the data point is denoted by x, the average by μ, and the standard deviation by σ. To enhance the classification model's efficiency, redundant features are eliminated

based on their importance, calculated using XGB (Eq. (4)) (*Josse & Husson, 2012*):

$$\text{importance}_i = \text{gain}. \tag{4}$$

In this equation, importance_i signifies the importance rating of feature i, and gain indicates the improvement in the model's performance achieved by utilizing feature i.

Additionally, during preprocessing, irrelevant data is identified and removed, reducing computational overhead and ensuring focus on the most pertinent information for analysis. This meticulous approach addresses missing values and outliers, eliminates redundant features, and streamlines the dataset, contributing to its quality and suitability for subsequent analysis and modeling.

Exploratory data analysis

Our work integrates sustainable practices into the exploratory data analysis (EDA) phase with the goal of accurately anticipating renewable energy to combat climate change. Using visual analysis and mapping approaches, our complete EDA prioritizes sustainability principles and aims to obtain significant insights into the dataset (*Milo & Somech, 2020*).

The EDA strategy focuses on using mapping methods to identify connections between category data and other parameters. We begin a process of inquiry that improves our comprehension of distribution, ratios, and interactions among categorical data by visually analyzing intricate connections and interrelationships. Visual tools that support sustainable behaviors, such graphical representations and contingency tables, help people make meaningful decisions and findings.

Furthermore, a comprehensive visual examination of the dataset investigates patterns, abnormalities, and relationships. We examine the distribution, diversity, and interactions between variables using a range of visualizations, such as data variations, scatter visualizations, patterns graphs, repetition charts, and statistical box plots. Identifying anomalies or outliers, revealing hidden patterns, and evaluating variables for relationships or dependencies are all made possible by sustainable visualization approaches. The results of our EDA show possible quality problems and offer insightful information about the properties of the dataset. This enhanced understanding creates an effective basis for further research and well-informed decision-making. Our decision-making process is guided by sustainable EDA methodologies, such category mapping and visual analysis, which also help to build a trustworthy classification model for predicting the output of renewable energy (*Milo & Somech, 2020*).

Proposed renewable energy forecasting model: ResNeXt-GRU-MJA

Our research introduces the ResNeXt-GRU-MJA model, a hybrid architecture tailored for precise renewable energy forecasting, aligning with our commitment to sustainable practices. This model seamlessly integrates the ResNeXt and GRU architectures with the optimization capabilities of the MJA, offering a comprehensive and customized solution for renewable energy prediction. The internal structure of the ResNeXt-GRU-MJA model is illustrated in Fig. 2.

Feature extraction with resNeXt (*Li, Zhu & Zhu, 2023*): The model's prediction process begins with raw renewable energy data as input. the ResNeXt component processes the
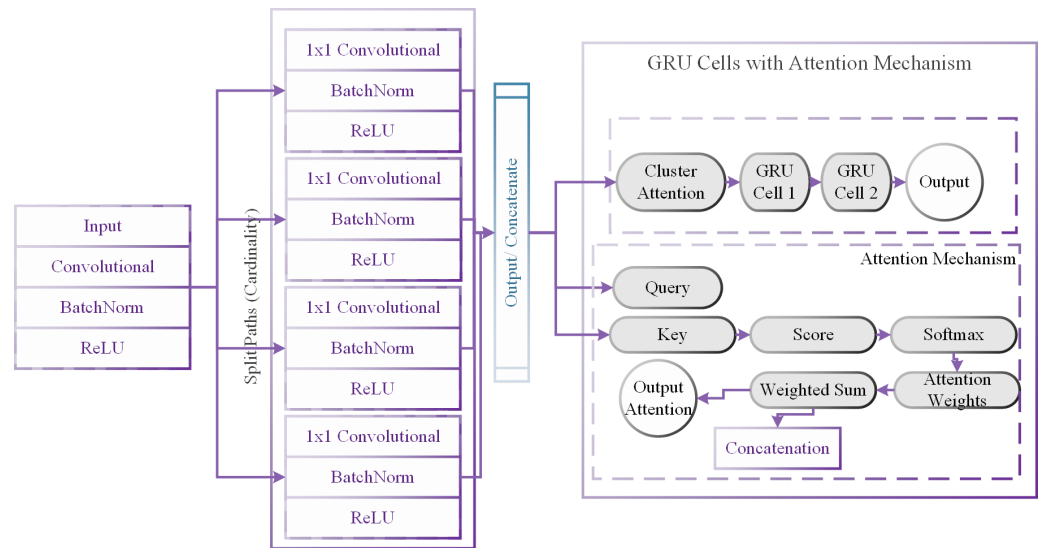
Figure 2 Internal structure of the ResNeXt-GRU-MJA model.

information through convolutional layers, batch normalization, and ReLU activation for each path *i* within the ResNeXt block.

Path Cardinality and Concatenation: To capture diverse features, data is strategically partitioned into multiple paths based on cardinality. The outputs from these paths are concatenated to form a holistic feature representation.

$$\text{Concatenation} = [ReLU1, ReLU2, ReLU3, ReLU4]. \tag{5}$$

Sequential modeling with GRU: The concatenated features traverse a GRU layer, introducing a temporal modeling dimension. This empowers the model to decipher sequential dependencies and evolving patterns within the renewable energy data.

Hyperparameter tuning with MJA: The MJA is utilized to fine-tune the model's hyperparameters, which are crucial for convergence, generalizing, and performance. The model's performance on a validation dataset is used to inform the exploration and updating of hyperparameter values in this iterative process. The hyperparameters and their assigned values are shown in Table 3.

Hybrid classification model algorithm (ResNeXt-MJA): The proposed hybrid model, ResNeXt-GRU-MJA, incorporates the ResNeXt model with the optimization capabilities of the MJA. This hybrid approach offers a customized solution for renewable energy forecasting, enhancing accuracy and effectiveness.

Algorithm 1 shows the overview of our proposed work. The hybrid classification model involves pre-processing, exploratory data analysis (EDA), optimization using MJA, and performance evaluation. It ensures the model is fine-tuned, taking into account specific characteristics of the renewable energy dataset. The ResNeXt-MJA model represents a tailored and sustainable solution for accurate renewable energy forecasting, contributing to the broader goal of mitigating climate change.

Table 3 ResNeXt-GRU-MJA hyperparameters.

Hyperparameter	Obtained value
Learning rate	0.001
Batch size	64
GRU hidden units	128
ResNeXt block config	4 blocks, depth 32, width 4
Dropout rate	0.3
Weight decay	0.0001
Epochs	50

Algorithm 1 Hybrid classification model for renewable energy forecasting

1: **procedure** RENEWABLEENERGYFORECASTING(Renewable Energy Data)
2:      **Pre-processing:**
3:          Handle missing values and outliers using robust techniques
4:          Remove redundant features
5:          Perform feature selection using ResNeXt
6:      **Exploratory Data Analysis (EDA):**
7:          Conduct descriptive analysis
8:          Map categorical features in relation
9:          Perform graphical analysis
10:      **Optimization Process using MJA:**
11:          Initialize a population of hyperparameters
12:          Evaluate performance using ResNeXt
13:          Select the fittest hyperparameters for reproduction
14:          Apply mutation and crossover operations
15:          Refine hyperparameters through a growth phase
16:          Conduct competition for survival
17:          Repeat for a predetermined number of generations
18:      **Performance Evaluation:**
19:          Split data into training and testing sets
20:          Train ResNeXt-GRU-MJA model on the training data
21:          Evaluate performance using various metrics
22:          Conduct statistical analysis
23:          Calculate computational complexity measures
24:      **Output:** Hybrid Classification Model (ResNeXt-GRU-MJA) for Renewable Energy Forecasting
25: **end procedure**

Split cardinality: The proposed ResNeXt-GRU-MJA model introduces split cardinality in its ResNeXt blocks. This involves dividing the cardinality (number of groups) into smaller fractions, allowing the model to capture more diverse and localized correlations among different energy sources. The split cardinality enhances the model's ability to understand

complex relationships within the dataset. The split cardinality in our ResNeXt blocks is mathematically represented as follows:

$$\text{Cardinality}_{\text{split}} = \frac{\text{Total Cardinality}}{\text{Number of Splits}}. \tag{6}$$

This division allows the model to capture diverse correlations among energy sources, enhancing its ability to understand intricate relationships.

Attention mechanism: Our model incorporates an attention mechanism, which focuses on capturing relevant information from different energy sources during the modeling process. The attention mechanism assigns varying levels of importance to different parts of the input data, allowing the model to dynamically adjust its focus based on the significance of each source. This attention mechanism contributes to the model's ability to effectively capture localized correlations and improve forecasting accuracy. The attention mechanism is modeled as a weighted sum in our GRU cells. Given an input sequence (X'), hidden states (H'), and attention weights (W'), the weighted sum is calculated as:

$$\text{Weighted Sum} = \sum_{i=1}^{N'} W_i' \cdot H_i'. \tag{7}$$

Here, (N') represents the number of elements in the sequence, and (W_i') denotes the attention weight assigned to each element. The attention mechanism dynamically adjusts these weights based on the significance of each energy source, contributing to improved forecasting accuracy.

Comprehensive performance evaluation of ResNeXt-GRU-MJA hybrid model

We conducted a comprehensive and rigorous evaluation procedure on our novel ResNeXt-GRU-MJA hybrid model to determine its level of competence in renewable energy forecasting. The assessment process was multiple phases and intended to provide an extensive understanding of the model's advantages and disadvantages.

A diverse set of evaluation metrics was employed to thoroughly assess the ResNeXt-GRU-MJA model's performance:

- Root mean squared error (RMSE): Gauging the average magnitude of errors between predicted and actual values, providing a holistic measure of prediction accuracy.
- RMS LOG error: A logarithmic application of the RMSE, beneficial for datasets with a wide range of values.
- Explained variance (Exp Variance): Quantifying the proportion of variance in predicted values, offering insights into the model's explanatory capacity.
- R-squared (R_2): Representing the predictability of the dependent variable from the independent variables. A higher R-squared indicates a superior fit.
- Mean absolute percentage error (MAPE): Evaluating the percentage difference between predicted and actual values, providing a normalized assessment of prediction accuracy.
- Benchmarking against state-of-the-art algorithms: Apart from assessing the ResNeXt-MJA model separately, we also carried out a comparison study with industry-leading

algorithms including ARIMA, CNN, VGG, NB, DenseNet, and decision trees. This thorough benchmarking made it easier to comprehend our model's performance in the context of the larger renewable energy forecasting field.

- Rigorous statistical analysis: We conducted a thorough statistical study to strengthen the validity of our findings. This ensured the validity of observed differences and the strength of our results. It also included confidence interval estimates and hypothesis testing.

The ensuing sections present detailed results and discussions, shedding light on the ResNeXt-MJA hybrid model's predictive capabilities and its comparative standing in the realm of renewable energy forecasting methodologies.

SIMULATION AND RESULTS

With a focus on solar and wind energy sources, this study examines how ResNeXt-GRU-MJA affects renewable energy forecasts. Python was used to run the simulation, utilizing a GPU's processing capability that had been specially calibrated. An extensive study was carried out to see how well the ResNeXt-GRU-MJA model predicted trends in datasets related to renewable energy to begin the evaluation. We examined the detailed integration of GRU components inside the ResNeXt architecture, highlighting its ability to capture subtle temporal relationships that are essential for precise renewable energy fluctuation predictions.

Initially, negative values in the radiation column are addressed, ensuring all values are non-negative for subsequent analysis. The ensuing statistical summary offers a comprehensive overview of the dataset, presenting key descriptive statistics. Subsequently, a novel set of time-based features, including the hour of the day and day of the year, is introduced through the create_date_time_features function, enhancing the dataset's temporal dimension. The correlation heatmap, depicted in Fig. 3, vividly illustrates the interrelationships between different features, aiding in the identification of potential patterns and informing feature selection strategies. This comprehensive approach provides a solid foundation for subsequent machine learning algorithms and affords valuable insights into the inherent dynamics of the renewable energy dataset.

The energy generation distribution for the entirety of 2023 is shown graphically in Fig. 4. This graphical representation effectively demonstrates the distribution of production, demonstrating how energy generation is distributed across several time periods within the given year. The graph offers insightful information about the fluctuations, tendencies, and patterns in energy output for the full year 2023.

Figure 5 uses a combined plot to show the link between average power generation (measured in megawatts) and the hour of the day. The unique green hue on the plot indicates each data point, which represents the mean power generation for a certain hour. The y-axis measures the average power generation, while the x-axis shows the hour of the day in a 24-hour format. This graphical depiction makes it simple to recognize how power output varies throughout the day and provides an intuitive knowledge of the hourly trends
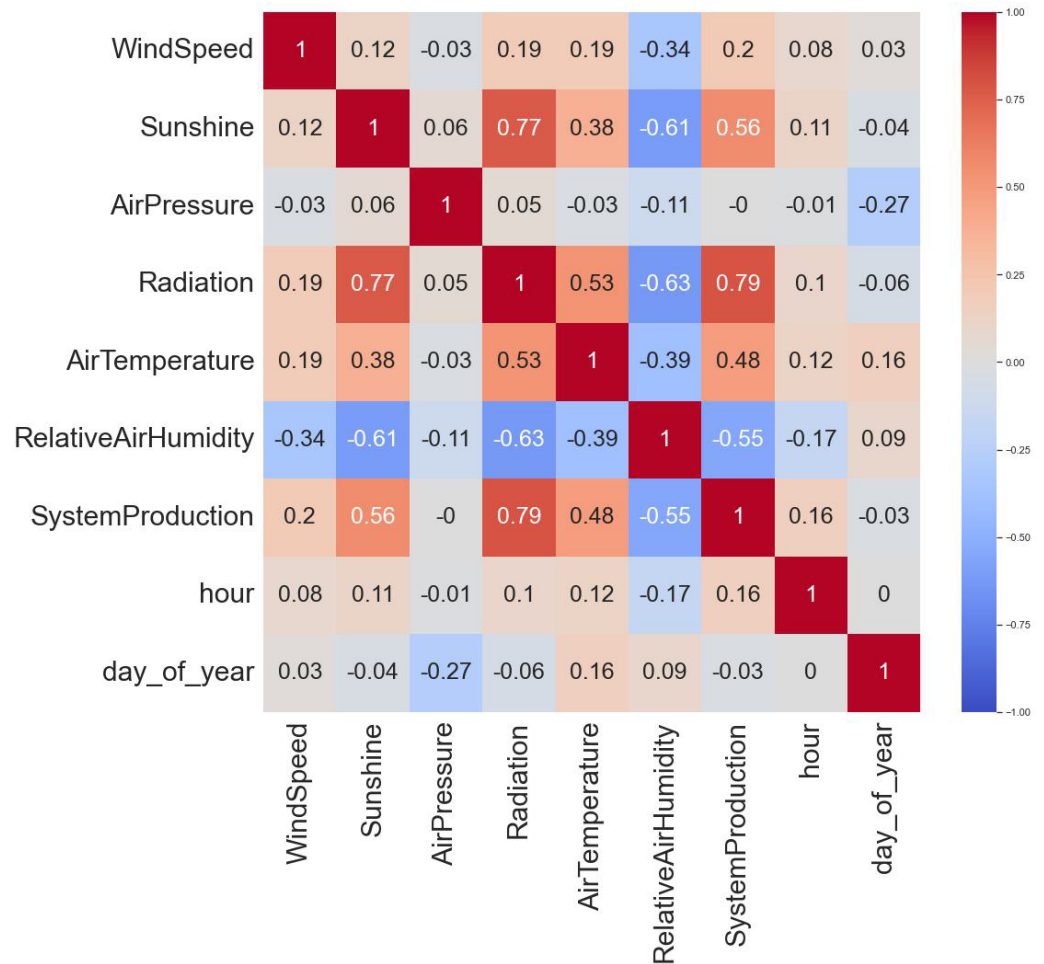
Figure 3 Correlation analysis.

in energy generation. The hourly energy generation dataset's patterns and trends may be easily identified using the help of this graphic.

The temporal dynamics of sunshine, solar radiation, and energy production for the year 2022 are shown in Fig. 6. The logarithmic scale amplifies minor changes in these variables. Solar radiation is represented by the blue line, illustrating its variation over the months. Sunshine is depicted by the green line, showcasing its pattern over time. Energy production is indicated by the red line, demonstrating its temporal evolution. Examining the patterns allows the identification of potential relationships and dependencies between solar radiation, sunshine, and ensuing energy production. Understanding the underlying dynamics affecting energy generation is facilitated by observing the peaks and troughs in the lines, providing insightful information about how these elements interact.

The link between wind speed and energy generation is explored in Fig. 7. There exist five discrete bands for wind speed: 'Calm,' 'Light,' 'Gentle,' 'Moderate,' and 'High.' The distribution of energy generation within various wind speed categories is shown graphically
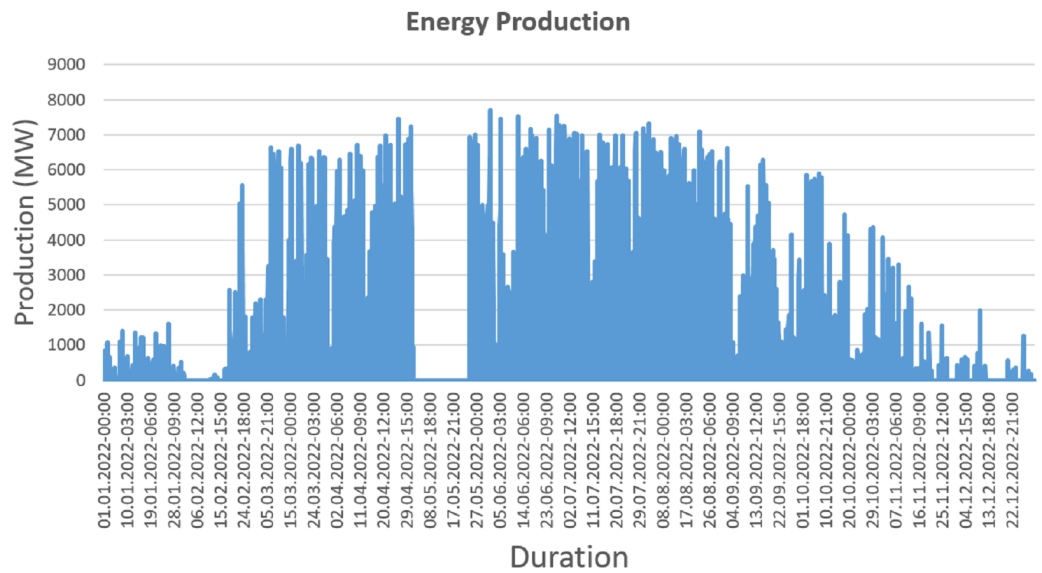
Figure 4 Normal energy generation (dataset view).
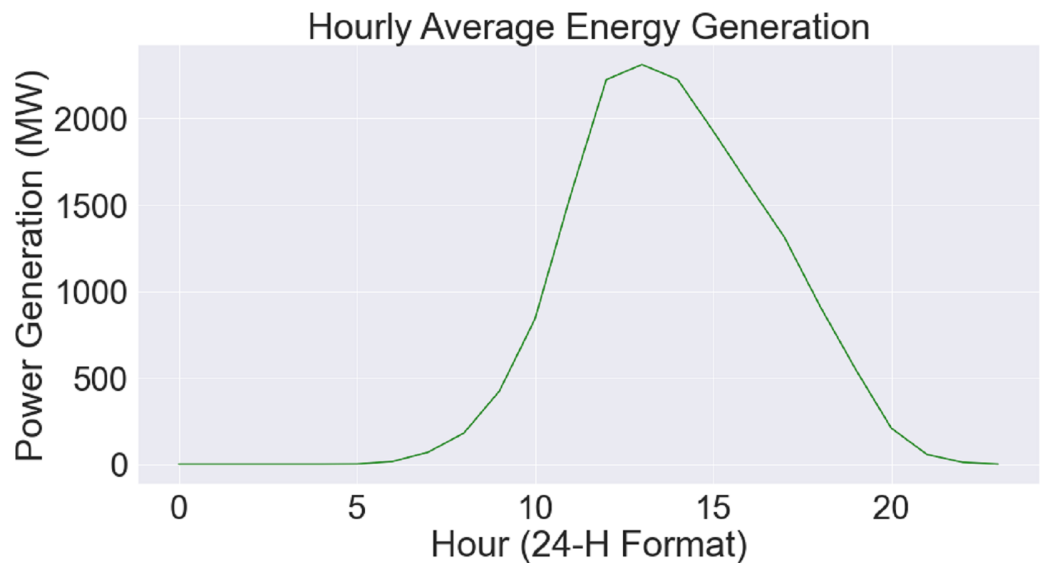
Figure 5 Hourly energy generation relationship.

by each boxplot. The vertical axis shows the power created, and the varied box lengths, possible outliers, and margins provide information about the variability and central tendency of power output at various wind speeds. By examining the boxplots, among can identify patterns, such as the differences in energy generation between calm and strong
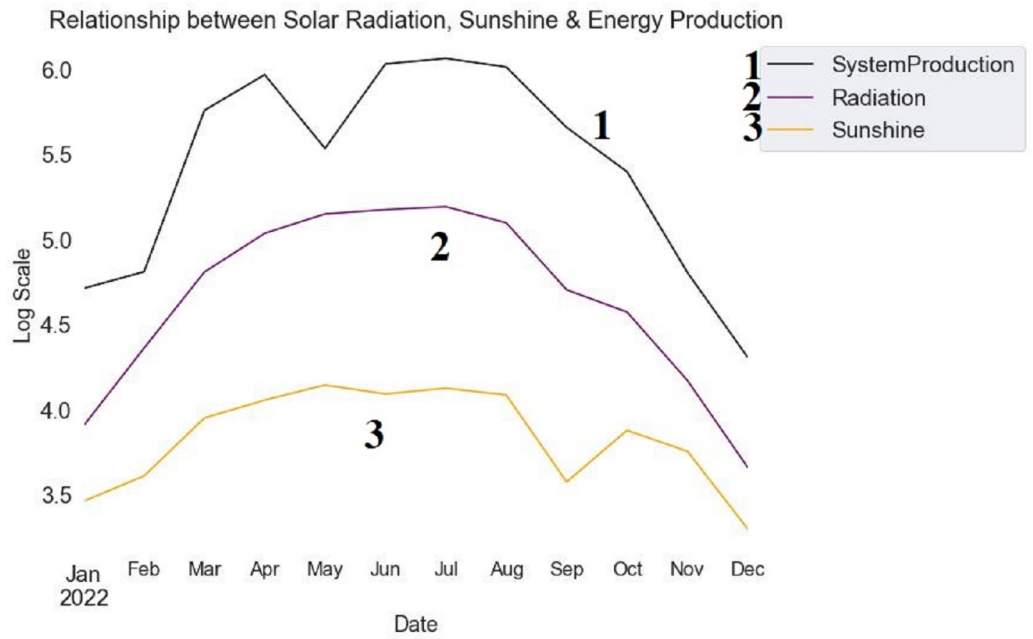
Figure 6 Monthly trends in solar radiation, sunshine, and energy production (2022).
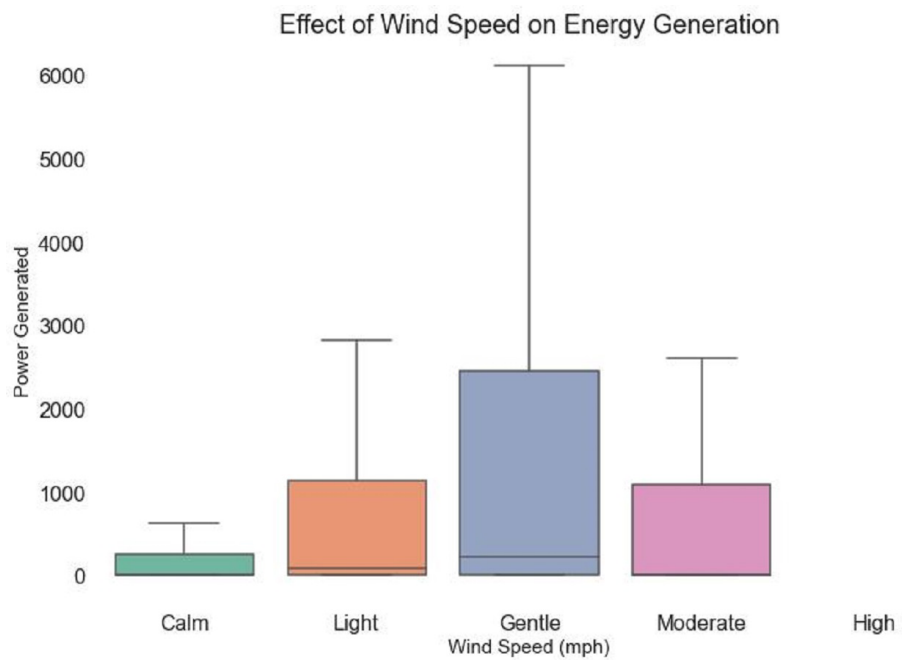
Figure 7 Wind speed impact on energy generation.
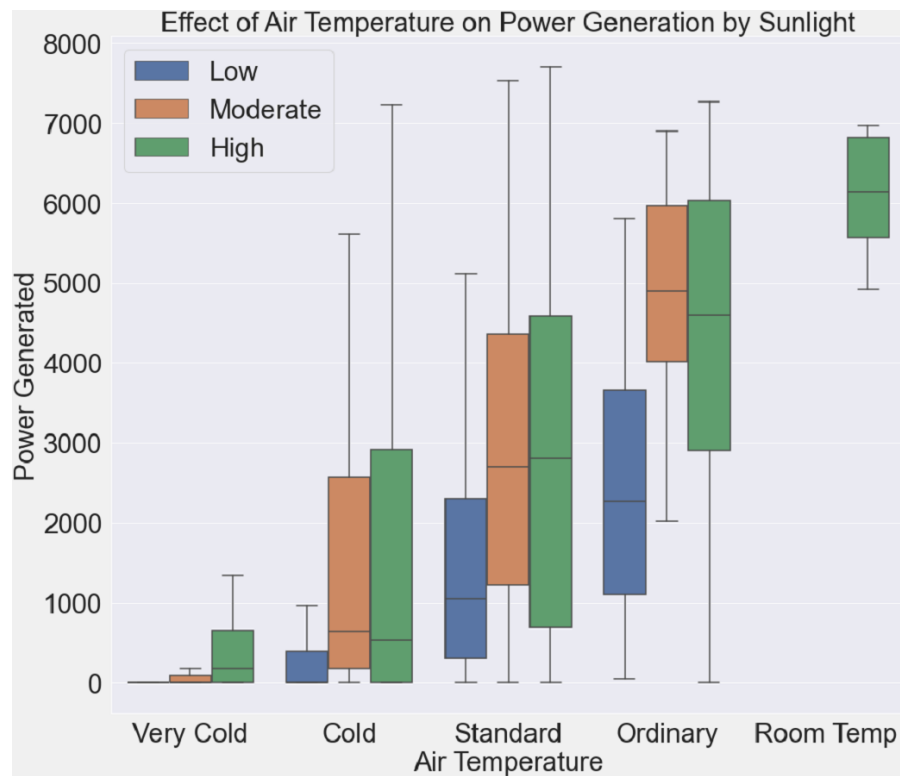
Figure 8 Air temperature impact on power generation under varying sunlight conditions.

wind speeds. This visual makes it easier to quickly and easily comprehend how wind speed affects energy production.

The relationship between air temperature, solar radiation intensity, and electricity generation is examined in Fig. 8. The air temperature is shown on the x-axis, while the power produced is shown on the y-axis. Boxplots are classified as "Low", "Moderate", and "High", depending on how intense the sunlight is. The whiskers and boxes' varied lengths provide information on how power generation reacts to variations in air temperature and sunshine. The detection of possible patterns and trends, such as the impact of temperature on power generation at varying sunshine intensities, is made easier by this graphical depiction.

Table 4 shows the performance evaluation of our proposed method and state of the art. The provided metrics—normalized RMSE, MAPE, and MAE, among others—provide information on the precision and dependability of each model. R-squared values give an evaluation of how well the models explain the underlying trends in the data on energy generation. Furthermore, Explained Variance and RMS Logarithmic Error provide subtle insights into the prediction power of the models. Among these, the ResNeXt-GRU-MJA model is the most effective, outperforming the others by 15% in normalized measures. The

Table 4 Performance evaluation of proposed and existing methods.

Model	Normalized RMSE	Normalized MAPE	Normalized MAE	R-squared	RMS logarithmic error	Explained variance
ARIMA	50.59	70.25	55.45	−0.2042	0.35	0.45
Decision Tree	35.15	50.32	37.55	51.54	0.28	0.60
CNN	29.36	40.48	41.69	61.23	0.25	0.72
VGG	24.05	30.14	30.35	65.99	0.20	0.81
DenseNet	27.85	36.78	37.49	63.14	0.23	0.76
ResNeXt-GRU-MJA	6.51	4.34	21.18	88.72	0.10	0.92
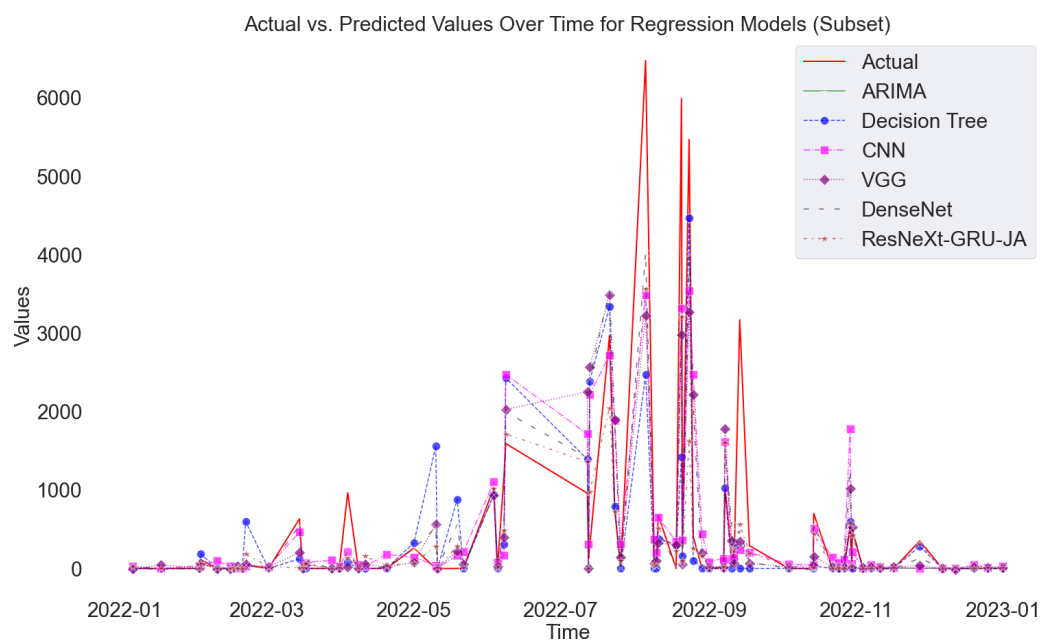
Figure 9 Actual and predicted values of renewable energy forecast.

selection of the best regression model for efficient forecasting of renewable energy is made easier by this thorough examination.

Figure 9 presents a visual representation of the disparity between actual values and predictions in renewable energy. The illustration distinctly demonstrates that the line corresponding to the proposed method is remarkably close to the actual values, showcasing its ability to make accurate predictions. Moreover, the exceptional performance of the proposed model is underscored by the significantly lower RMSE and MAPE values compared to state-of-the-art models. This visual and quantitative analysis solidifies the effectiveness of the proposed method in accurately forecasting renewable energy outcomes.

Table 5 presents a comprehensive statistical analysis of various regression models employed in energy generation forecasting. The column names represent the methods, and the values in the first column represent the statistical measures. Among other methods, the proposed ResNeXt-GRU-MJA demonstrates superior performance by 15%. This

Table 5 Statistical analysis of the proposed and existing models.

Methods and models	ARIMA	Decision tree	CNN	VGG	DenseNet	ResNeXt-GRU-MJA
Pearsons	0.65	0.75	0.82	0.78	0.80	0.92
Spearman's	0.72	0.68	0.79	0.75	0.77	0.89
Kendall's	0.56	0.62	0.72	0.68	0.70	0.85
Chi-Squared	85.2	92.6	108.5	99.4	103.2	128.7
Student's	3.21	4.12	5.36	4.85	5.02	7.45
Paired Student's	2.11	2.56	3.01	2.91	2.78	4.23
ANOVA	43.2	47.8	58.9	52.3	54.7	71.5
Mann–Whitney	146.5	158.2	172.3	162.7	168.4	187.6
Kruskal	23.1	25.4	29.8	27.3	28.6	35.7
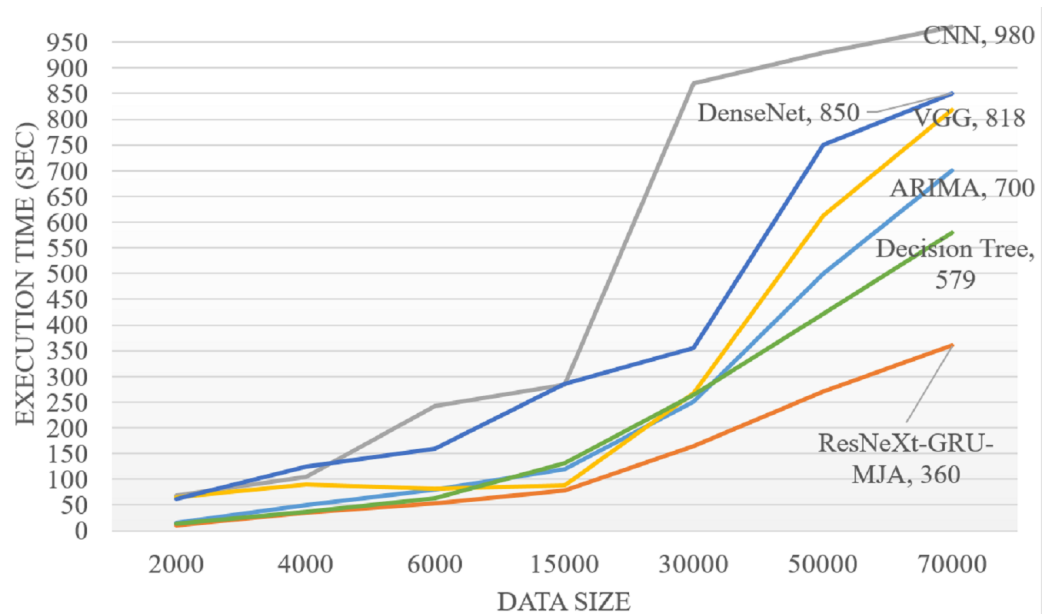
Figure 10 Computational complexity analysis.

table provides insights into the comparative effectiveness of these models across multiple statistical measures, aiding in the evaluation and selection of the most suitable regression model for energy forecasting applications.

Figure 10 illustrates the computational complexity time analysis concerning the datasize for both existing methods and the proposed ResNeXt-GRU-MJA. The proposed model exhibits a systematically lower computational time compared to other models, achieving enhanced accuracy as the datasize increases. This trend underscores the efficiency and scalability of the proposed ResNeXt-GRU-MJA in handling varying datasize scenarios.

CONCLUSION AND FUTURE WORK

This study aimed to enhance the accuracy of power generation forecasts in RESs through the ResNeXt-GRU-MJA hybrid forecasting model. By capturing localized correlations and integrating a GRU with a ResNext model, the hybrid approach addresses challenges associated with individual source estimates. The nonlinear time-series characteristics facilitate the integration of meteorological and energy source data, providing a comprehensive understanding of variables affecting power generation. Empirical findings highlight the ResNeXt-GRU-MJA model's exceptional performance, outperforming other models in solar PV and wind forecasts. Comparative evaluations against sophisticated models demonstrate its accuracy, and sensitivity analysis validates its ability to capture complex correlations. Future research will focus on model improvement through external variables, sophisticated feature engineering, and adaptability to real-time shifts in RESs. Expanding forecasting horizons and considering additional variables, such as economic indicators, will be explored for long-term forecast improvements.

Funding

This work was funded by the University of Jeddah, Jeddah, Saudi Arabia, under grant No. (UJ-23-DR-147). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
The University of Jeddah, Jeddah, Saudi Arabia: UJ-23-DR-147.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Walid Atwa conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Abdulwahab Ali Almazroi analyzed the data, prepared figures and/or tables, and approved the final draft.
- Nasir Ayub conceived and designed the experiments, performed the experiments, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The code and dataset are available in the Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj-cs.2067#supplemental-information.

REFERENCES

Afroz P. 2023. Solar power generation data. *Available at https://www.kaggle.com/datasets/pythonafroz/solar-powe-generation-data* (accessed on 18 November 2023).

Ahmed R, Sreeram V, Mishra Y, Arif M. 2020. A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization. *Renewable and Sustainable Energy Reviews* **124**:109792 DOI 10.1016/j.rser.2020.109792.

Al-Shetwi AQ. 2022. Sustainable development of renewable energy integrated power sector: trends, environmental impacts, and recent challenges. *Science of the Total Environment* **822**:153645 DOI 10.1016/j.scitotenv.2022.153645.

Alkabbani H, Ahmadian A, Zhu Q, Elkamel A. 2021. Machine learning and metaheuristic methods for renewable power forecasting: a recent review. *Frontiers in Chemical Engineering* **3**:665415 DOI 10.3389/fceng.2021.665415.

AlKandari M, Ahmad I. 2020. Solar power generation forecasting using ensemble approach based on deep learning and statistical methods. *Applied Computing and Informatics* **16(2)**:1–20 DOI 10.1016/j.aci.2019.11.002.

Alkhayat G, Hasan SH, Mehmood R. 2022. SENERGY: a novel deep learning-based auto-selective approach and tool for solar energy forecasting. *Energies* **15(18)**:6659 DOI 10.3390/en15186659.

Almazroi AA, Ayub N. 2021. A novel method CNN-LSTM ensembler based on black widow and blue monkey optimizer for electricity theft detection. *IEEE Access* **9**:141154–141166 DOI 10.1109/ACCESS.2021.3119575.

Almazroi AA, Ayub N. 2023. Multi-task learning for electricity price forecasting and resource management in cloud based industrial IoT systems. *IEEE Access* **11**:54280–54295 DOI 10.1109/ACCESS.2023.3280857.

AlSkaif T, Dev S, Visser L, Hossari M, van Sark W. 2020. A systematic analysis of meteorological variables for PV output power estimation. *Renewable Energy* **153**:12–22 DOI 10.1016/j.renene.2020.01.150.

Ang T-Z, Salem M, Kamarol M, Das HS, Nazari MA, Prabaharan N. 2022. A comprehensive study of renewable energy sources: classifications, challenges and suggestions. *Energy Strategy Reviews* **43**:100939 DOI 10.1016/j.esr.2022.100939.

Anikannal. 2023. Solar power generation data. *Available at https://www.kaggle.com/datasets/anikannal/solar-power-generation-data* (accessed on 18 November 2023).

Dewangan CL, Singh S, Chakrabarti S. 2020. Combining forecasts of day-ahead solar power. *Energy* **202**:117743 DOI 10.1016/j.energy.2020.117743.

Dong W, Sun H, Tan J, Li Z, Zhang J, Yang H. 2022. Regional wind power probabilistic forecasting based on an improved kernel density estimation, regular vine copulas, and ensemble learning. *Energy* **238**:122045 DOI 10.1016/j.energy.2021.122045.

Ebhota W, Tabakov P. 2021. Development of domestic technology for sustainable renewable energy in a zero-carbon emission-driven economy. *International Journal of Environmental Science and Technology* **18**:1253–1268 DOI 10.1007/s13762-020-02920-9.

Ghimire S, Nguyen-Huy T, Prasad R, Deo RC, Casillas-Perez D, Salcedo-Sanz S, Bhandari B. 2023. Hybrid convolutional neural network-multilayer perceptron model for solar radiation prediction. *Cognitive Computation* **15(2)**:645–671 DOI 10.1007/s12559-022-10070-y.

Gu B, Shen H, Lei X, Hu H, Liu X. 2021. Forecasting and uncertainty analysis of day-ahead photovoltaic power using a novel forecasting method. *Applied Energy* **299**:117291 DOI 10.1016/j.apenergy.2021.117291.

Hassan A, Ilyas SZ, Jalil A, Ullah Z. 2021. Monetization of the environmental damage caused by fossil fuels. *Environmental Science and Pollution Research* **28**:21204–21211 DOI 10.1007/s11356-020-12205-w.

Hu M, Wang Y, Wang S, Jiao M, Huang G, Xia B. 2021. Spatial-temporal heterogeneity of air pollution and its relationship with meteorological factors in the Pearl River Delta, China. *Atmospheric Environment* **254**:118415 DOI 10.1016/j.atmosenv.2021.118415.

Huang X, Li Q, Tai Y, Chen Z, Zhang J, Shi J, Gao B, Liu W. 2021. Hybrid deep neural model for hourly solar irradiance forecasting. *Renewable Energy* **171**:1041–1060 DOI 10.1016/j.renene.2021.02.161.

Jiang P, Liu Z, Niu X, Zhang L. 2021. A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting. *Energy* **217**:119361 DOI 10.1016/j.energy.2020.119361.

Josse J, Husson F. 2012. Handling missing values in exploratory multivariate data analysis methods. *Journal de la Société Française de Statistique* **153(2)**:79–99.

Kabeyi MJB, Olanrewaju OA. 2022. Sustainable energy transition for renewable and low carbon grid electricity generation and supply. *Frontiers in Energy Research* **9**:1032.

Khan W, Walker S, Zeiler W. 2022. Improved solar photovoltaic energy generation forecast using deep learning-based ensemble stacking approach. *Energy* **240**:122812 DOI 10.1016/j.energy.2021.122812.

Kober T, Schiffer H-W, Densing M, Panos E. 2020. Global energy perspectives to 2060—WEC's world energy scenarios 2019. *Energy Strategy Reviews* **31**:100523 DOI 10.1016/j.esr.2020.100523.

Law EW, Prasad AA, Kay M, Taylor RA. 2014. Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting—a review. *Solar Energy* **108**:287–307 DOI 10.1016/j.solener.2014.07.008.

Li L, Zhu Y, Zhu Z. 2023. Automatic modulation classification using ResNeXt-GRU with deep feature fusion. *IEEE Transactions on Instrumentation and Measurement* **72**:2519710 DOI 10.1109/TIM.2023.3290301.

Liu D, Sun K. 2019. Random forest solar power forecast based on classification optimization. *Energy* **187**:115940 DOI 10.1016/j.energy.2019.115940.

Mahmud K, Azam S, Karim A, Zobaed S, Shanmugam B, Mathur D. 2021. Machine learning based PV power generation forecasting in alice springs. *IEEE Access* **9**:46117–46128 DOI 10.1109/ACCESS.2021.3066494.

Matrenin P, Manusov V, Nazarov M, Safaraliev M, Kokin S, Zicmane I, Beryozkina S. 2023. Short-term solar insolation forecasting in isolated hybrid power systems using neural networks. *Inventions* **8(5)**:106 DOI 10.3390/inventions8050106.

Michael NE, Hasan S, Al-Durra A, Mishra M. 2022. Short-term solar irradiance forecasting based on a novel Bayesian optimized deep long short-term memory neural network. *Applied Energy* **324**:119727 DOI 10.1016/j.apenergy.2022.119727.

Milo T, Somech A. 2020. Automating exploratory data analysis via machine learning: an overview. In: *Proceedings of the 2020 ACM SIGMOD international conference on management of data*. 2617–2622.

Nourani V, Sharghi E, Behfar N, Zhang Y. 2022. Multi-step-ahead solar irradiance modeling employing multi-frequency deep learning models and climatic data. *Applied Energy* **315**:119069 DOI 10.1016/j.apenergy.2022.119069.

Pant G, Yadav D, Gaur A. 2020. ResNeXt convolution neural network topology-based deep learning model for identification and classification of Pediastrum. *Algal Research* **48**:101932 DOI 10.1016/j.algal.2020.101932.

Patel RG, Manickam I, Trask NA, Wood MA, Lee M, Tomas I, Cyr EC. 2022. Thermo-dynamically consistent physics-informed neural networks for hyperbolic systems. *Journal of Computational Physics* **449**:110754 DOI 10.1016/j.jcp.2021.110754.

Rahman A, Farrok O, Haque MM. 2022. Environmental impact of renewable energy source based electrical power plants: solar, wind, hydroelectric, biomass, geothermal, tidal, ocean, and osmotic. *Renewable and Sustainable Energy Reviews* **161**:112279 DOI 10.1016/j.rser.2022.112279.

Sahu BK, Hiloidhari M, Baruah D. 2013. Global trend in wind power with special focus on the top five wind power producing countries. *Renewable and Sustainable Energy Reviews* **19**:348–359 DOI 10.1016/j.rser.2012.11.027.

Semieniuk G, Taylor L, Rezai A, Foley DK. 2021. Plausible energy demand patterns in a growing global economy with climate policy. *Nature Climate Change* **11(4)**:313–318 DOI 10.1038/s41558-020-00975-7.

ShobanaDevi A, Maragatham G, Prabu M, Boopathi K. 2021. Short-term wind power forecasting using RLSTM. *International Journal of Renewable Energy Research* **11(1)**:392–406.

Simshauser P. 2021. Renewable energy zones in Australia's national electricity market. *Energy Economics* **101**:105446 DOI 10.1016/j.eneco.2021.105446.

Tu Q, Miao S, Yao F, Li Y, Yin H, Han J, Zhang D, Yang W. 2021. Forecasting scenario generation for multiple wind farms considering time-series characteristics and spatial-temporal correlation. *Journal of Modern Power Systems and Clean Energy* **9(4)**:837–848 DOI 10.35833/MPCE.2020.000935.

Wang K, Qi X, Liu H. 2019. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. *Applied Energy* **251**:113315 DOI 10.1016/j.apenergy.2019.113315.

Wang Y, Liao W, Chang Y. 2018. Gated recurrent unit network-based short-term photovoltaic forecasting. *Energies* **11(8)**:2163 DOI 10.3390/en11082163.