FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 8 No. 6, December, 2024, pp 285 - 298
DOI: https://doi.org/10.33003/fjs-2024-0806-2786

EFFICIENT METHOD FOR FORECASTING SOLAR IRRADIANCE - A REVIEW

*[1,2]Odejobi Olusegun A., [1]Alawode, Kehinde O., [1]Lawal, Muyideen O.

[1]Department of Electrical Electronics Engineering, Osun State University, Osogbo, Nigeria
[2]Prototype Engineering Development Institute, Ilesa, Nigeria

*Corresponding authors' email: oodejobi@pedi.gov.ng; segodej@gmail.com

ABSTRACT

Efficient solar irradiance forecasting is essential for optimizing solar energy systems and integrating renewable energy sources into power grids. This review aims to evaluate the effectiveness of various forecasting methods to inform energy management and grid integration strategies. It compares physical models, statistical approaches, machine learning techniques, and hybrid models, using specific criteria such as accuracy, computational efficiency, and data requirements. Physical models like Numerical Weather Prediction (NWP) provide detailed atmospheric simulations but are computationally intensive. Statistical models, such as ARIMA, are simpler yet struggle with non-linear data. Machine learning methods, particularly Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks, effectively capture complex data relationships but require substantial datasets and computing power. Hybrid models, which combine physical and machine learning approaches, achieve high accuracy and are suitable for real-time applications despite increased computational costs. One of the key findings indicates that hybrid models offer high accuracy but demand significant computational resources and offer the best balance between accuracy and computational efficiency**.** However, challenges such as data quality, geographic and temporal variability, and model complexity persist. Emerging technologies like artificial intelligence, big data, and quantum computing present promising solutions for enhanced irradiance forecasting. The review highlighted the models' limitations and strengths to facilitate informed decision making and concluded with recommendation of the adoption of hybrid models, investment in data acquisition and sharing technologies, balancing model complexity with practicality as steps towards improved irradiance forecasting for grid integration and stability to ensure sustainable yet cost-effective energy solutions.

Keywords: Forecasting, Irradiation forecasting, Machine learning, Solar prediction, Efficient forecasting models

INTRODUCTION

The world is transitioning towards more sustainable energy systems, solar energy stands out as one of the most abundant and renewable sources of power that is capable of playing a vital role in the decarbonization of energy systems and contributes significantly to efforts aimed at combating climate change. Unlike fossil alternatives, solar energy is clean, sustainable, and available in almost all regions globally. The increasing deployment of solar PV is expected to provide around 32% of the world's total electricity demand by 2050, necessitating an installed capacity of approximately 14 terawatts peak (TWp) and annual installations of 630 GW each year by that time (Lennon *et al*., 2022), helping to reduce greenhouse gas emissions and enhance energy security (Sobrina *et al*., 2018). However, integrating solar power into existing power grids poses challenges due to its intermittent nature (Zha and Lu, 2014; Avwioroko *et al*., 2024). Solar irradiance, which refers to the power of solar radiation received per unit area, fluctuates based on weather conditions, time of day, and geographic location. This variability affects the stability and efficiency of solar energy systems. The need to ensure that electricity generation matches demand requires precise control and prediction of solar energy availability. Therefore, accurate solar irradiance forecasting is essential for optimizing the performance of solar PV systems, reducing operational costs, and maintaining the reliability of power grids (Sweeney *et al*., 2019).

The Need for Accurate Solar Irradiance Forecasting

Forecasting solar irradiance is a critical component of managing solar energy systems. Accurate predictions allow grid operators to plan energy dispatch and ensure a stable supply of electricity, even when solar energy is not immediately available. This is particularly important in regions where solar power constitutes a significant portion of the energy mix, as accurate forecasting helps balance supply and demand, reducing the need for costly backup power from fossil fuels (Zwane *et al*., 2022). Furthermore, precise irradiance forecasts enable solar power systems to integrate more smoothly with smart grids, facilitating effective energy storage, improving power quality, and enhancing overall grid stability (Wang *et al*, 2020).

The benefits of solar irradiance forecasting extend beyond grid management. For energy market participants, accurate forecasting allows for better decision-making in energy trading and risk management. As described by Sweeney *et al*. (2019), the importance of forecast accuracy has increased with the rise of renewable energy sources like solar and wind, making forecasting a vital tool for both energy producers and consumers (Sweeney *et.al*., 2019). With the growing need for renewable energy and the variability of solar power, the development of efficient forecasting methods is essential for the successful integration of solar energy into the power grid.

The Paper's Structure

This review paper is structured to provide an evaluation of the most efficient and recent methods for forecasting solar irradiance. The paper is organized as follows: Section 2 covers the basics of solar irradiance, including its significance to solar energy systems and the factors influencing its variability, such as geographic location, atmospheric conditions, and seasonal changes. Section 3 examines various forecasting models, ranging from physical models like Numerical Weather Prediction (NWP) to statistical models,

machine learning techniques, and hybrid approaches that combine both physical and data-driven models.
In Section 4, the performance of different forecasting methods is compared, focusing on accuracy, computational efficiency, and data requirements. The strengths and weaknesses of each method will be discussed, as well as the trade-offs between model complexity and practicality. Section 5 highlights the challenges and limitations associated with solar irradiance forecasting, such as uncertainty in weather predictions, data quality issues, and the complexities of model development. In Section 6, the paper tries to identify emerging techniques and research directions aimed at improving solar forecasting accuracy. Finally, Section 7 summarizes the key findings and provides recommendations for researchers and practitioners, emphasizing the growing importance of accurate solar irradiance forecasting for optimizing solar energy systems and ensuring grid stability.

Aim of the Review

The primary goal of this review is to highlight the efficient methods for forecasting solar irradiance, with a particular focus on accuracy, computational efficiency, and practical applicability. By examining various forecasting techniques, from traditional physical models to advanced machine learning approaches, this paper aims to provide a guide to selecting the appropriate methods for different solar energy applications. This review, in addition, will explore how emerging technologies, such as the Internet of Things (IoT), big data, and cloud-based systems, can revolutionize the field of solar forecasting. These advancements have the potential to significantly improve the accuracy of solar irradiance forecasts and enhance the reliability of solar energy systems (Wang *et al*., 2020). Ultimately, this review aims to provide valuable insights for researchers, engineers, and decision-makers involved in the development and implementation of solar energy technologies.

Solar Irradiance Basics

The total solar irradiance (TSI) refers to the sum of the solar electromagnetic radiation energy of all wavebands reaching the top of the earth's atmosphere per unit area in unit time at the average distance between the sun and the earth (Biktash, 2017). Solar irradiance is defined as the power per unit area produced by the Sun in the form of electromagnetic radiation. It is a critical parameter in solar energy systems because it determines the amount of solar energy that can be converted into electricity by photovoltaic (PV) panels. Measured in watts per square meter (W/m^2), solar irradiance fluctuates depending on several factors, including geographic location, atmospheric conditions, and time of day (Sobrina *et al.*, 2018). Understanding solar irradiance is essential for optimizing the performance of solar power systems, as it directly influences energy output. Zwane *et al.*, (2022) indicated that Systems designed for solar energy production, such as PV panels and concentrated solar power plants, rely on accurate irradiance measurements to estimate energy potential and manage energy supply.
Solar irradiance also plays a crucial role in the economic feasibility of solar energy projects. By predicting the amount of sunlight that will reach a specific area, energy planners can determine the most suitable locations for solar farms and optimize their design. For instance, according to Heinemann *et al*., (2006), high irradiance levels make certain regions, such as deserts and equatorial areas, ideal for large-scale solar energy generation. Moreover, accurate irradiance data helps in planning energy storage systems, ensuring that surplus energy generated during sunny periods can be stored for later use, thereby improving grid stability (Radovan *et al*., 2021; Sobrina *et al*., 2022)

Solar Irradiance and Photovoltaic Systems

Photovoltaic systems rely on solar irradiance as the key input that determines the amount of electricity generated by PV panels. When sunlight strikes a PV panel, the energy from the irradiance is converted into electrical energy through the photovoltaic effect. The efficiency of this process is influenced by the intensity of the solar irradiance (Lei *et al*., 2022)., the angle of incidence (Lu, 2023), and the temperature of the PV cells (Darian and Ghorreshi, 2020). For optimal performance, PV systems are designed to maximize their exposure to sunlight, often by adjusting the tilt and orientation of the panels to capture the most irradiance throughout the day and across seasons (Zwane *et al*., 2022).
Solar irradiance forecasting is vital for predicting the power output of PV systems. By knowing the expected levels of irradiance, energy providers can plan energy dispatch, optimize the use of energy storage systems, and ensure that supply meets demand. This is particularly important in regions where solar energy forms a significant portion of the energy mix, as fluctuations in irradiance can lead to instability in the grid if not properly managed (Yang *et al*., 2018). Furthermore, understanding irradiance patterns helps energy developers in designing PV systems that can operate efficiently even under varying weather conditions.

Factors Affecting Solar Irradiance

Several factors influence solar irradiance, affecting the amount of solar energy available for power generation. These factors include:

i. *Geographic Location:* Solar irradiance varies significantly depending on a location’s latitude and altitude. Regions near the equator receive more direct sunlight year-round compared to regions at higher latitudes, where the Sun’s angle is lower, and irradiance levels are generally weaker (Zhang *et al*., 2019)
ii. *Time of Year:* The Earth's axial tilt causes seasonal variations in solar irradiance. During summer, regions in the northern and southern hemispheres experience higher irradiance as they tilt closer to the Sun, while during winter, the irradiance levels decrease as the tilt moves away from the Sun (Melhem and Shaker, 2023; Heinemann *et al*., 2006; Yang *et al*., 2018).
iii. *Weather Conditions:* Cloud cover, humidity, and atmospheric particles, such as dust and aerosols, can scatter and absorb sunlight, reducing the amount of irradiance that reaches the Earth's surface. These weather patterns are highly variable and are one of the main challenges in predicting solar irradiance with accuracy (Govender and Sivakumar, 2019).
iv. *Atmospheric Influences:* The thickness of the atmosphere between the Sun and the Earth's surface affect solar irradiance. In regions with higher altitudes, the atmosphere is thinner, resulting in less scattering and absorption of sunlight. Conversely, in regions closer to sea level, the atmosphere is denser, leading to more scattering and reduced irradiance (Lunche *et al*., 2016).
v. *Albedo Effect:* The albedo effect refers to the reflection of solar radiation from the Earth's surface. Surfaces like snow, water, and sand have different reflectivity levels, which can influence the amount of irradiance that is absorbed by solar panels. High-reflectivity surfaces, such as snow, can increase the overall irradiance received by solar panels in certain conditions (Nguyen *et al*., 2017; Sobrina *et al*., 2018).

Figure 1 illustrates the flow of solar radiation through the atmosphere, highlighting how clouds and aerosols affect irradiance levels.

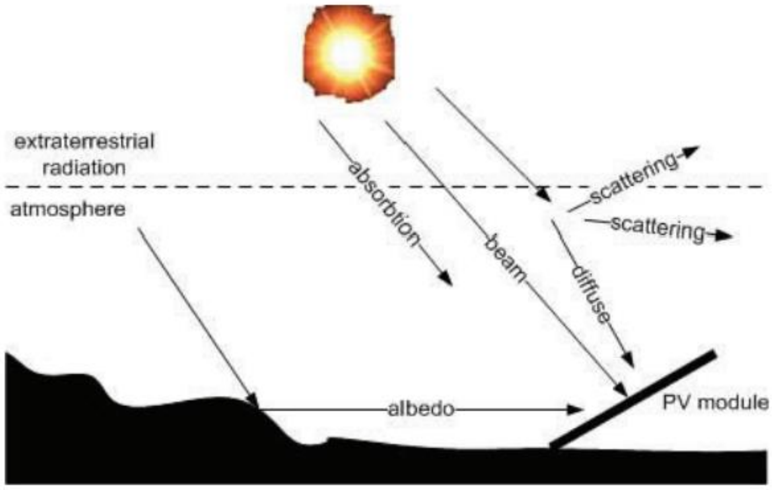

Figure 1: The flow of solar radiation through the atmosphere (Ibrahim *et al*., 2012)

Solar Irradiance Forecasting Methods
Forecasting solar irradiance is crucial for enhancing solar energy production, particularly in relation to grid integration and energy demand management. Over the years, numerous methods have been developed, each presenting varying degrees of accuracy, computational efficiency, and data prerequisites. These methods can be generally classified into four categories: physical models, statistical models, machine learning-based models, and hybrid models. Each category possesses distinct benefits and drawbacks, which will be examined in this section.

Physical Models
Physical models rely on the laws of physics to predict solar irradiance by simulating atmospheric conditions and solar radiation processes. These models include Numerical Weather Prediction (NWP) models and radiative transfer models.

Numerical Weather Prediction (NWP) Models
NWP models are among the most widely used physical methods for solar irradiance forecasting. These models use atmospheric data, such as temperature, wind speed, humidity, and cloud cover, to simulate weather conditions and predict solar irradiance (Ibrahim *et. al*., 2022. NWP models typically operate over various time scales, from hours to several days, making them suitable for short-term and medium-term forecasting (Du *et al*., 2018; Trapero *et al*., 2015; Hashimoto and Yoshimoto, 2023)
The basic equation governing NWP models is derived from the laws of fluid dynamics and thermodynamics, specifically through the application of partial differential equations (PDEs). These equations describe the motion and thermodynamic processes of the atmosphere, allowing for the simulation of weather phenomena. For instance, NWP models utilize non-linear differential equations that encapsulate the dynamics of atmospheric flow, which are essential for accurate forecasting (Sutikno, 2024; Trojáková *et al*., 2019; Zhang *et al*., 2014).

$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u = -\frac{1}{\rho}\nabla p + g + F \qquad (1)$$

Where:
u is the velocity field, t is time, ρ is air density, p is pressure, g is gravitational force, and F represents frictional forces.

NWP model face significant limitations, particularly regarding high computational costs and the necessity for extensive data inputs. The computational demands of NWP models arise from their complex algorithms and the need for high-resolution grids, which can restrict their operational feasibility and scalability (Rojas-Campos *et al*., 2022). Furthermore, the accuracy of NWP forecasts is often compromised by the quality and quantity of input data, as insufficient data can lead to systematic biases and errors in predictions (Tian, 2024; Seo and Cha, 2023). The assimilation of meteorological data is crucial for enhancing model accuracy, yet this process also contributes to the overall computational burden (Hastuti *et al*., 2023; Liu *et al*., 2013). Consequently, while NWP models are invaluable tools for weather forecasting, their operational efficiency is hindered by these inherent challenges (Pathak *et al*., 2022).

Radiative Transfer Models
Radiative transfer models focus on simulating the interaction of solar radiation with the Earth's atmosphere. These models account for the absorption, scattering, and reflection of solar radiation by atmospheric gases, clouds, and aerosols. The SMARTS (Simple Model of the Atmospheric Radiative Transfer of Sunshine) model is one example of a radiative transfer model used in solar irradiance forecasting (Bouchouicha and Bachari, 2023; Lunche *et al*., 2016). Santa Barbara DISORT Atmospheric Radiative Transfer (SBDART) model is another radiative transfer model, they are highly accurate, particularly for short-term forecasts where atmospheric conditions play a dominant role. Their development exemplifies the ongoing efforts to enhance the accuracy of radiative transfer simulations, thereby improving climate modeling and solar energy applications (Gupta, 2023; Jahani *et al*., 2021). However, they require precise atmospheric data, which may not always be available in real-time (Heinemann *et. al.*, 2006).

Statistical Models
Statistical models serve as essential tools for forecasting solar irradiance by leveraging historical data. Among these are time series models, and classical regression models which can be employed to establish relationships between solar irradiance and various predictor variables, such as meteorological factors. These were reviewed in this section.

Time Series Models (ARIMA, SARIMA)

Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) are popular time series models used in solar irradiance forecasting. ARIMA models are widely recognized for their effectiveness in predicting solar irradiance by leveraging historical data, assuming a linear relationship between past and future values. This linearity is crucial for accurate forecasting, as demonstrated in various studies (Alsharif *et al*., 2019; Santos *et al*., 2022; Chodakowska *et al*., 2023). However, the inherent seasonality of solar radiation necessitates the use of Seasonal ARIMA (SARIMA) models, which incorporate seasonal patterns to enhance prediction accuracy. SARIMA models have been shown to outperform traditional ARIMA models in contexts where seasonal variability significantly influences solar irradiance, such as in solar energy forecasting (Belmahdi *et al*., 2023; Fara *et al*., 2021).

The general equation for ARIMA is:

$$Y_t = \alpha + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \qquad (2)$$

where:

Y_t is the predicted irradiance, α is a constant, ϕ_i are the autoregressive parameters, θ_j are the moving average parameters, ϵ_t is the error term, and p , q are the autoregressive and moving average terms, respectively.

Although ARIMA and SARIMA models are simple and computationally efficient, they may struggle to capture complex non-linear relationships in solar irradiance data (Akhter *et al*., 2016).

Classical Regression Models

Classical regression models, offer foundational techniques such as Linear and Logistic Regression, each suited to different types of data and research needs. Linear Regression is used for modeling linear relationships between variables, while Logistic Regression, employing the logit function, is ideal for binary outcome variables (Gupta *et al*., 2017; Zhang, 2024). These basic models are essential for straightforward data analysis, while more specialized methods extend their utility in complex scenarios. The mathematical representation of a classical linear regression model is expressed in Equation 3.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \epsilon_i \qquad (3)$$

where Y_i represents the dependent variable for observation i, X_{1i} , X_{2i} ,…, X_{ki} are the independent variables, β_0 is the intercept, β_1, β_2,..., β_k are the coefficients, and ϵ_i is the error term accounting for unobserved factors.

In high-dimensional settings, where the number of predictors exceeds the number of observations, regularization techniques such as Lasso, Ridge, and Elastic Net become crucial. The Least Absolute Shrinkage and Selection Operator (Lasso) applies an L1 penalty to induce sparsity in the model, effectively selecting a subset of predictors by shrinking some coefficients to zero (Emmert-Streib and Dehmer, 2019; Xu, 2024). Ridge regression, on the other hand, employs an L2 penalty, which helps mitigate issues of multicollinearity without necessarily performing variable selection (Nur, 2023; Xin and Khalid, 2018). Elastic Net combines both Lasso and Ridge penalties, providing a balanced approach that can handle correlated predictors effectively (Kayanan and Wijekoon, 2020; Watagoda *et al*., 2021).

Stepwise regression, including both forward and backward selection, is another method used to refine models by iteratively adding or removing predictors based on statistical criteria (Rady and Mahmoud, 2018; Ajeel and Hashem, 2020). Generalized Linear Models (GLMs) extend traditional linear models to accommodate various types of response variables, such as Poisson and Binomial distributions, making them versatile for different research contexts (Rady and Mahmoud, 2018). Quantile Regression offers a robust alternative by estimating the conditional median or other quantiles of the response variable, which is particularly useful in cases of heteroscedasticity (Al-Sharoot, 2023; Tang *et al*., 2020).

Weighted Least Squares (WLS) is usually employed when dealing with heteroscedasticity, allowing for more accurate estimates by giving different weights to different observations based on their variance (Rady and Mahmoud, 2018; He *et al*., 2019). Ordinary Least Squares (OLS) remains a widely used method due to its simplicity and interpretability, although it can be sensitive to outliers and multicollinearity (Rady and Mahmoud, 2018; Xin and Khalid, 2018). Partial Least Squares (PLS) and Principal Component Regression (PCR) are techniques that reduce dimensionality by transforming predictors into a smaller set of uncorrelated components, which can enhance model performance in high-dimensional data (Emmert-Streib and Dehmer, 2019; Al-Sharoot, 2023).

Classical regression models have been widely utilized to predict solar irradiance by correlating it with various meteorological factors such as temperature, humidity, and cloud cover. Some authors opined that these models are straightforward to implement and interpret, making them accessible for initial analyses of solar irradiance data (Soni *et al*., 2011; Hejase and Assi, 2012; Aggarwal and Saini, 2014). However, they often fall short in capturing the complex, non-linear dynamics inherent in solar irradiance variations, primarily due to the stochastic nature of cloud cover and other atmospheric conditions (Cha *et al*., 2021). The choice of a regression model is contingent upon the specific characteristics of the data and the research questions at hand. Classical regression models provide a robust framework for understanding relationships between variables, while a more advanced techniques like Lasso, Ridge, and Elastic Net offer solutions for high-dimensional challenges, ensuring that researchers can derive meaningful insights from their analyses.

Machine Learning-Based Models

Machine learning models have become increasingly prominent in solar irradiance forecasting due to their capability to process extensive datasets and model complex non-linear relationships. Techniques such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) are widely utilized for this purpose, demonstrating significant efficacy in predicting solar irradiance components like global and diffuse radiation (Abdel-Nasser *et al*., 2021; Boubaker *et al*., 2021). Moreover, advanced methods, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have shown superior performance in capturing intricate patterns in solar data, thereby enhancing forecasting accuracy (Mukhtar *et al*., 2022; Alali *et al*., 2023).

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are effective for forecasting solar irradiance due to their ability to capture intricate, non-linear relationships between various inputs, such as meteorological data, and outputs like solar irradiance levels. The architecture of ANNs includes multiple layers of interconnected neurons, where each neuron computes a weighted sum of its inputs followed by a non-linear activation function, allowing for complex modeling capabilities (Belmahdi *et al*., 2023; Voyant *et al*., 2017). ANNs consist of multiple layers of neurons, with each neuron performing a weighted sum of its inputs followed by a non-linear activation function. The general equation for an ANN is expressed in Equation 4:

$$f(x) = \sigma(\sum_{i=1}^{n} w_i x_i + b) \quad (4)$$

where:
x_i are the input variables, w_i are the weights, b is the bias term, and σ is the activation function (Ahmed *et al*., 2020).

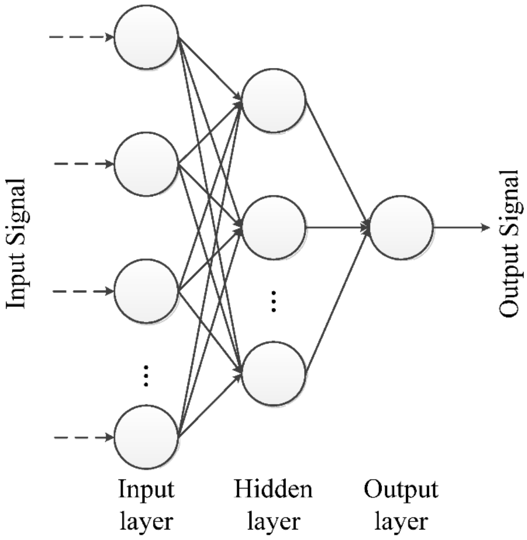

Figure 2: Basic artificial neural network structure (Zhu *et al*., 2015)

Studies have shown that incorporating relevant meteorological parameters significantly enhances the accuracy of solar irradiance predictions, with specific factors such as the clearness index and relative air mass being identified as critical inputs (Belmahdi *et al*., 2023; Cha *et al*., 2021). Moreover, the integration of satellite data and numerical weather prediction models with ANNs has been demonstrated to improve forecasting precision, particularly in dynamic environments where conditions change rapidly (Hashimoto and Yoshimoto, 2023; Du *et al*., 2018). This multifaceted approach underscores the potential of ANNs in optimizing solar energy management and enhancing the reliability of renewable energy systems (Santos *et al*., 2022). Application and implementation of ANN is however resource intensive.

Support Vector Machines (SVM)

Support Vector Machines (SVM) are a prominent machine learning technique utilized for solar irradiance forecasting. This method operates by identifying a hyperplane that effectively distinguishes between various data classes. In regression applications, SVM is adept at predicting continuous variables, such as solar irradiance levels, by minimizing prediction errors while maximizing the margin between the predicted and actual values (Cha *et al*., 2021; Voyant *et al*., 2017). SVM models demonstrate significant efficacy with small to medium-sized datasets; however, they may encounter challenges when applied to larger datasets due to computational complexity and overfitting risks (Voyant *et al*., 2017; Wang *et al*., 2016). Recent studies have highlighted the versatility of SVM in conjunction with other machine learning algorithms, enhancing the accuracy of solar irradiance predictions (Cha *et al*., 2021; Santos *et al*., 2022).

Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)

Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are pivotal in time-series forecasting, particularly for solar irradiance prediction. CNNs excel in feature extraction from spatial data, such as satellite imagery, enabling effective analysis of solar patterns (Belmahdi *et al*., 2023; Hashimoto and Yoshimoto, 2023). Conversely, RNNs, specifically Long Short-Term Memory (LSTM) networks, are adept at managing sequential data, making them suitable for capturing long-term dependencies in solar irradiance time series (Rojas-Campos *et al*., 2022). Research indicates that LSTMs can outperform traditional forecasting methods by leveraging their memory capabilities to enhance prediction accuracy (Santos *et al*., 2022; Du *et al*., 2018). Figure 3 shows basic LSTM unit.

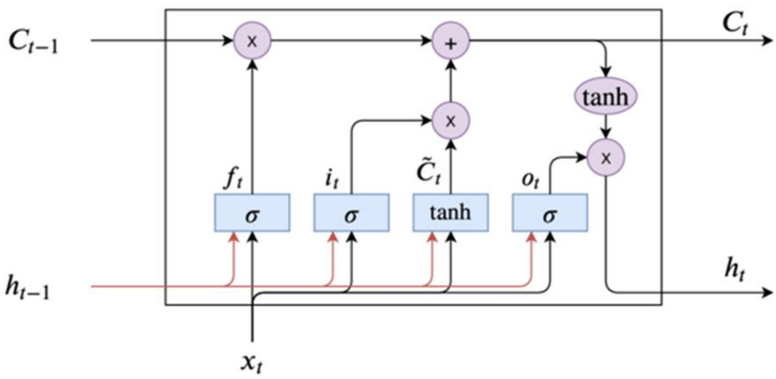

Figure 3: Long-Short Term Memory (LSTM) unit (Smilevski, 2020)

The mathematical expression of the LSTM unit shown in Figure 3 is expressed by Equation 5 to Equation 10

$$f_t = \sigma\left(W_{fh}h_{t-1} + W_{fx}x_t + b_f\right), \quad (5)$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i), \quad (6)$$

$$\tilde{c}_t = \tanh(W_{\tilde{c}h}h_{t-1} + W_{\tilde{c}x}x_t + b_{\tilde{c}}), \quad (7)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t, \quad (8)$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o), \quad (9)$$

$$h_t = o_t \cdot \tanh(c_t). \quad (10)$$

c_t denotes the cell state of the LSTM. W_i, $W_{\tilde{c}}$ and W_o are the weights. Jozefowicz *et. al.*, (2015) discovered that increasing the forget gate bias b_f increases the performance of the LSTM unit.

It is however germane to indicate that CNNs often struggle with capturing long-term temporal dependencies and require significant computational resources and data, limiting their practicality in certain scenarios. In RNNs, the opaque nature of the models makes it difficult to understand how and why they arrive at their predictions. This is particularly problematic in scenarios demanding high safety standards (Liu *et. al.*, 2024). A prominent limitation of LSTMs is their high computational demands and the need for extensive training data, making real-time applications more challenging (Shrestha, 2023; Akhter *et. al.*, 2016).

Hybrid Models

Hybrid models combine the strengths of physical models and machine learning approaches to improve forecasting accuracy. These models use physical models to capture the fundamental physics of solar irradiance, while machine learning algorithms are used to fine-tune the predictions based on historical data. Recent research has shown that hybrid models are highly effective for forecasting solar irradiance. These models, which integrate historical weather data, sky imaging, and physical principles, have demonstrated improved accuracy over different forecasting timeframes. For instance, Almarzooqi *et al*. (2024) introduced a hybrid system that utilizes truncated-regularized kernel ridge regression, surpassing traditional forecasting techniques in both short- and medium-term predictions. Karout *et al*. (2023) combined clear-sky direct normal irradiance (DNI) forecasts with machine learning processing of sky images, yielding superior results under varying sky conditions. Wang *et al*. (2023) developed a hybrid ensemble model using XGBoost, incorporating historical data and sky images, which outperformed standard methods for short-term forecasting. These findings emphasize the value of hybrid models towards the improvement of the precision of solar irradiance forecasts. Another example is the use of NWP models in conjunction with ANNs to improve short-term solar irradiance predictions.

It is evident that hybrid approaches outperform standalone models (Santos *et al*., 2022; Chodakowska *et al*., 2023). However, the implementation of these models often demands substantial computational resources and extensive datasets, which can complicate their practical application (Pathak *et al*., 2022; Rojas-Campos *et al*., 2022).

Comparison of Forecasting Methods

The performance of solar irradiance forecasting methods is evaluated based on several critical factors, including accuracy, computational efficiency, data requirements, and their respective advantages and disadvantages. In this section, we compare the physical, statistical, machine learning, and hybrid approaches based on these criteria.

Evaluation Metrics

The effectiveness of forecasting models is typically measured using standard metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics quantify the difference between predicted and actual irradiance values and are widely used to evaluate the accuracy of forecasting methods.

Root Mean Square Error

This metric emphasizes larger errors by squaring the differences between predicted and observed values, making it sensitive to outliers. The equation for RMSE is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2} \quad (11)$$

where:

Y_i is the actual irradiance, $\widehat{Y}_i$ is the predicted irradiance, and n is the number of observations.

Mean Absolute Error

MAE is the average of the absolute errors between the predicted and observed values, making it more robust to outliers. The equation for MAE is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \widehat{Y}_i| \quad (12)$$

Mean Absolute Percentage Error

This metric expresses forecast error as a percentage of the actual values, providing an intuitive measure of accuracy. MAPE is calculated as:

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n} \left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \quad (13)$$

These metrics are critical in comparing the relative accuracy of the various forecasting models.

Accuracy

Accuracy in solar irradiance forecasting refers to the degree to which predicted values align with actual observed values over a specified timeframe. Accuracy in solar irradiance forecasting is typically assessed using several key metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics quantitatively measure the discrepancies between predicted and actual irradiance values, enabling developers to evaluate model performance effectively. For instance, studies have shown that RMSE and MAE are essential for comparing the accuracy of various forecasting models, including those based on machine learning and statistical approaches (Belmahdi *et al*., 2023; Yang *et al*., 2018). The closer to zero the values of an error metric is, the more accurate the model. Hybrid models like CNN + LSTM can achieve a lower RMSE and MAE values, making them more accurate. These models leverage the strengths of both convolutional layers (for spatial data analysis) and recurrent layers (for time-series forecasting), resulting in improved performance over standalone models like ARIMA or ANN.

Computational Efficiency

While accuracy is crucial, computational efficiency is also an important consideration, particularly for real-time applications. Physical models like Numerical Weather Prediction (NWP), while highly detailed, are computationally expensive due to the need for simulating atmospheric processes over large geographic areas. This makes NWP models less suitable for real-time solar forecasting (Sutikno, 2024). In contrast, statistical models like ARIMA are computationally efficient, leveraging simple linear relationships in data, but they often fail to capture the non-linear dynamics of solar irradiance effectively (Santos *et al*., 2022; Chodakowska *et al*., 2023).

Machine learning models, particularly ANNs and LSTMs, offer a balance between accuracy and computational complexity. While they require more processing power than traditional statistical models, advances in computational hardware and algorithms have made them more viable for real-time forecasting (Belmahdi *et al*., 2023; Cha *et al*., 2021).

Hybrid models, which combine physical and machine learning approaches, tend to be the most computationally intensive, but they also deliver the highest accuracy (Hashimoto and Yoshimoto, 2023; Fara *et al*., 2021).

Data Requirements

Data availability is a key factor in the choice of forecasting models. Physical models like NWP rely heavily on meteorological data, including temperature, wind speed, and cloud cover, which can be difficult to obtain in real-time for certain regions (Pathak *et al*., 2022; Trojáková *et al*., 2019). Radiative transfer models also require detailed atmospheric data, such as aerosol concentrations and humidity levels.

Statistical models, such as ARIMA, require historical irradiance data but do not depend on real-time meteorological inputs. This makes them suitable for locations with limited meteorological infrastructure, but they may struggle to capture sudden changes in weather patterns (Santos *et al*., 2022; Chodakowska *et al*., 2023).

Machine learning models require large datasets for training, including both historical irradiance data and meteorological variables. These models can also integrate satellite data and real-time weather information, making them highly flexible but data-intensive (Voyant *et al*., 2017; Cha *et al*., 2021). The hybrid models, which combine physical and machine learning approaches, require even more extensive datasets, as they need both real-time meteorological inputs and historical irradiance data (Belmahdi *et al*., 2023; Fara *et al*., 2021).

Strengths and Weaknesses of Models

Each forecasting method has its strengths and weaknesses, depending on the application and available resources. Physical models are highly detailed but computationally demanding, while statistical models are simpler and more efficient yet may struggle with non-linear relationships. Machine learning models can capture complex patterns in data but often require extensive resources and expertise. Hybrid models combine the strengths of multiple approaches for improved accuracy but are complex to implement. Ensemble methods enhance stability and generalization but may lack interpretability. Table 2 provides a detailed comparison of the strengths and weaknesses of various forecasting approaches for broader insights.

Comparative Analysis of Models

Solar irradiance forecasting models present varying levels of accuracy, data requirements, computational efficiency, and applicability to specific use scenarios. Deep learning and hybrid models, such as LSTM, CNN, and CNN-LSTM hybrids, offer the highest accuracy (Sansine *et. al.*, 2023). However, these models come with substantial trade-offs, requiring extensive historical and real-time meteorological datasets and significant computational power, which limits their real-time applicability in resource-constrained settings. In contrast, statistical models like ARIMA and linear regression have lower data requirements and are highly computationally efficient, making them suitable for rapid forecasting in data-limited environments. However, their simplicity often results in less accurate predictions compared to machine learning and hybrid approaches. Physical models, such as NWP, provide reasonable medium-term accuracy and simulate atmospheric processes in detail but require extensive data and high computational resources, impacting their practicality for real-time use.

Table 1: Strength and weaknesses of the forecasting methods

Model Type	Model Name	Strengths	Weaknesses
Physical Models	Numerical Weather Prediction (NWP)	- Highly detailed; incorporates complex meteorological processes	- Computationally intensive; requires extensive atmospheric data; limited real-time application
	Radiative Transfer Models (RTM)	- Accurate in modeling solar radiation interactions in the atmosphere; Excellent for clear-sky irradiance forecasting	- Requires precise atmospheric data; not suitable for short-term forecasts or real-time prediction; - Limited by the availability of real-time atmospheric parameters
Statistical Models	ARIMA, SARIMA	- Simple and computationally efficient	- Struggles with non-linear relationships; less accurate for short-term forecasts
	Multiple Linear Regression (MLR)	- Easy to implement and interpret - Requires fewer data points than machine learning models	- Assumes linear relationships, which may not capture non-linear behavior in solar irradiance data - Performs poorly when irradiance is influenced by complex, dynamic weather patterns
Machine Learning & Deep Learning Models	Artificial Neural Networks (ANN)	- Can model complex, non-linear relationships in the data	- Requires large datasets for training; computationally intensive
	Support Vector Machines (SVM)	- Good generalization with small datasets	- Limited ability to handle very large datasets; requires careful tuning of hyperparameters
	Long Short-Term Memory (LSTM)	- Excellent for time-series forecasting; remembers long-term dependencies	- Requires large amounts of data; computationally intensive

	Convolutional Neural Networks (CNN)	- Effective at extracting spatial patterns from satellite images - High accuracy in capturing both spatial and temporal data	- Requires large computational resources and data; complex to implement. - Prone to overfitting in smaller datasets
Hybrid Models	Physical + Machine Learning (e.g., CNN + LSTM)	- Combines physical insights with data-driven accuracy - Highest accuracy in short-term and medium-term forecasts	Requires extensive computational resources and data; challenging for real-time application Complex model architecture that is difficult to interpret
Ensemble Methods	Boosting, Bagging, Random Forests	- Improves model stability by combining multiple weak learners - Reduces variance and improves generalization	- Computationally demanding; hyperparameter tuning is critical - May be less interpretable compared to single models

The trade-offs between model complexity and real-time applicability are crucial, especially in resource-limited settings. While complex models like deep learning and hybrid approaches excel in accuracy and adaptability, their demand for data and computational power can make them impractical for smaller-scale or under-resourced applications. Table 3 highlights these distinctions, showing how different models perform based on key factors. Physical models are more suited for medium- to long-term planning where computational delays are acceptable. In contrast, statistical models offer simplicity, fast computation, and minimal data requirements, making them valuable for quick forecasting and feasibility studies, albeit with a compromise in predictive power. Machine learning models such as ANN and LSTM balance complexity and utility, providing good accuracy for short-term predictions in moderately resourced environments. Ultimately, selecting the appropriate forecasting model involves navigating the trade-offs between accuracy, complexity, data availability, and the need for real-time responses to ensure the chosen approach aligns with the specific operational and resource constraints of the application.

Summary of Model Comparison

The choice of solar irradiance forecasting methods depends on factors like accuracy, computational resources, and data availability. Physical models provide detailed forecasts but are computationally intensive, which limits their use in real-time applications (Hashimoto and Yoshimoto, 2023). Statistical models like ARIMA are easier to implement, though they may lack the necessary precision for more critical scenarios (Cha *et al*., 2021; Fara *et al*., 2021). Machine learning and hybrid models offer a balance between accuracy and efficiency by leveraging large datasets, including satellite and weather prediction data, to enhance performance (Belmahdi *et al*., 2023; Voyant *et al*., 2017; Wang *et al*., 2020). Despite their higher accuracy, these models require substantial computational power and training data, making them challenging to implement in resource-constrained settings (Belmahdi *et al*., 2023; Voyant *et al*., 2017).

Challenges and Limitations

While significant progress has been made in developing models for solar irradiance forecasting, several challenges and limitations still hinder the effectiveness and efficiency of these methods. These challenges arise from the variability in weather patterns, the complexity of data, and the need for balancing model accuracy with computational feasibility. This section outlines the primary challenges associated with solar irradiance forecasting and the limitations faced by the current forecasting models.

Uncertainty in Weather Predictions

One of the key challenges in solar irradiance forecasting is the uncertainty in weather predictions, particularly with cloud cover and atmospheric conditions. NWP models, although effective for large-scale simulations, often struggle with localized weather phenomena and the rapid changes in cloud cover, which significantly affect irradiance levels (Hashimoto and Yoshimoto, 2023; Tian, 2024). This issue is especially problematic for short-term forecasts, where even small inaccuracies can lead to large deviations in irradiance estimates (Radovan *et al*., 2021). Machine learning models, while capable of capturing non-linear relationships, also face limitations due to the chaotic and unpredictable nature of weather systems. These models rely on historical data, which may not fully capture future variability (Cha *et al*., 2021).

Table 3: Comparative Overview of Forecasting Models: Performance Metrics and Use Scenarios

Model Type	Model Name	Accuracy	Data Requirements	Computational Efficiency	Typical Use Scenario
Physical Model	Numerical Weather Prediction (NWP)	Medium-High (accurate for medium-term)	Very High (extensive meteorological data, real-time inputs)	Low (high computational cost)	Medium- to long-term grid operation planning, energy dispatch
	Radiative Transfer Models	High for short-term (clear conditions)	High (detailed atmospheric data, aerosols, cloud cover)	Low (computationally intensive)	Short-term solar power output prediction in specific weather conditions

Statistical Model	ARIMA	Medium	Low (basic historical irradiance data)	High (low computational cost)	Basic short-term forecasts in data-limited areas, initial system planning
	SARIMA	Medium	Low-Medium (historical data with seasonal components)	High (low computational cost)	Seasonal solar generation forecasting, early-stage planning
	Linear Regression	Low-Medium	Low (minimal historical data)	Very High (simple to compute)	Preliminary analysis, educational tools, simple solar forecasting
	Multiple Linear Regression	Medium	Medium (historical data with multiple input variables)	High	Simple multi-variable forecasting for basic operational insights
Machine Learning Model	Artificial Neural Networks (ANN)	High (non-linear data handling)	High (large datasets with historical and weather inputs)	Medium-High (training-intensive)	Adaptive short-term forecasting with changing weather patterns
	Support Vector Machines (SVM)	Medium-High	Medium (moderate-sized datasets with meteorological variables)	Medium (efficient with proper kernel selection)	Short-term predictions in moderately resourced environments
	Long Short-Term Memory (LSTM)	Very High	Very High (large training datasets, sequential time-series data)	Low-Medium (computationally demanding)	Real-time solar power prediction, handling diurnal and sequential data
	Convolutional Neural Networks (CNN)	High (spatial data processing)	High (satellite imagery, weather data)	Medium (processing-intensive)	Forecasting spatial variations in irradiance, cloud tracking
	Gradient Boosted Trees	High	Medium-High (historical data, weather variables)	Medium (efficient with boosting techniques)	Enhanced medium-term forecasts with structured data
	Random Forests	Medium-High	Medium (historical and meteorological data)	High (efficient but with many trees)	Forecasting with non-linear data variability, robust modeling
Hybrid Model	CNN + LSTM Hybrid	Very High (best performance)	Very High (extensive combined meteorological and historical data)	Low (most demanding computationally)	High-accuracy short-term forecasting, advanced grid management
	NWP + ANN	High	High (combined real-time and historical data)	Medium (depends on input complexity)	Comprehensive medium- to long-term solar generation predictions
	Ensemble Learning (e.g., Bagging, Boosting)	Very High	High (large, diverse datasets)	Low-Medium (depending on ensemble size)	Aggregated forecasts for high precision, uncertainty reduction
Deep Learning Model	Deep Belief Networks (DBN)	High	Very High (extensive datasets)	Low (due to complex training phases)	Advanced non-linear forecasting with extensive data modeling
	Recurrent Neural Networks (RNNs)	High	High (sequential data with weather attributes)	Medium-High (computationally complex)	Time-dependent solar power predictions, handling variability

Although combining NWP and data-driven techniques has enhanced forecasting, the unpredictability of weather patterns remains a significant obstacle (Yang *et al*., 2018; Fara *et al*., 2021).

Data Availability and Quality

Accurate solar irradiance forecasting heavily relies on high-quality data, including real-time meteorological information, historical records, and satellite observations. However, in many developing regions, the scarcity or poor quality of such data hinders forecasting accuracy (Belmahdi *et al*., 2023; Chodakowska *et al*., 2023). Machine learning models, in particular, require clean, well-labelled datasets, and incomplete data reduces their performance (Voyant *et al*., 2017). Additionally, acquiring high-resolution satellite imagery and real-time weather data is especially challenging in rural or remote areas with limited infrastructure (Hashimoto and Yoshimoto, 2023; Radovan *et al*., 2021). Innovative solutions, such as ground-based sky imagers, are needed to improve data quality and forecasting reliability (Du *et al*., 2018; Lu, 2023).

Model Complexity vs Practicality

As solar irradiance forecasting models evolve, a clear trade-off arises between accuracy and computational feasibility. Hybrid models, which combine Numerical Weather Prediction (NWP) with machine learning techniques like Artificial Neural Networks (ANN) or Long Short-Term Memory (LSTM) networks, deliver superior accuracy in capturing complex meteorological patterns. However, their high computational requirements limit their practicality in resource-constrained environments such as developing regions or small-scale solar operations (Hashimoto and Yoshimoto, 2023; Belmahdi *et al*., 2023). On the other hand, simpler statistical models like ARIMA and linear regression offer better computational efficiency but often lack the accuracy needed for modeling the complex dynamics of solar irradiance (Alsharif *et al*., 2019; Chodakowska *et al*., 2023; Fara *et al*., 2021). The challenge lies in balancing model complexity with the need for real-time predictions, particularly in scenarios requiring rapid data processing (Sobri *et al*., 2018; Wang *et al*., 2020).

Geographic and Temporal Resolution

The accuracy and applicability of solar irradiance forecasting models are heavily influenced by their geographic and temporal resolution. Localized factors such as topography, vegetation, and urban infrastructure require models capable of capturing these variations for more precise forecasts, particularly in distributed solar energy systems (Chodakowska *et al*., 2023; Du *et al*., 2018). Temporal resolution is equally important, with short-term forecasts critical for real-time grid management and medium-term forecasts supporting energy dispatch (Sweeney *et al*., 2019). While NWP models perform well for medium-term forecasting, they face challenges in short-term accuracy due to rapid atmospheric changes (Rojas-Campos *et al*., 2022). To enhance both spatial and temporal resolution, satellite imagery and ground-based data are utilized, though their integration can add complexity and may not always be available in real-time (Du *et al*., 2018). Ground-based sky imagers offer improved data for intra-hour forecasts, addressing some limitations of satellite data (Lu, 2023).

Model Interpretability

Interpretability refers to the ability to understand how a model arrives at its predictions or forecasts. It is the extent to which humans, especially operators and decision-makers, can trust, explain, and justify the outcomes produced by a model. Machine learning models, particularly deep learning techniques like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, demonstrate high accuracy in solar irradiance forecasting but face challenges due to their lack of interpretability. This "black box" nature complicates trust and decision-making in critical applications such as energy grid management, where transparency is essential (Tian, 2024; Voyant *et al*., 2017; Belmahdi *et al*., 2023). To improve interpretability, researchers are developing hybrid models that integrate physical principles with machine learning to enhance understanding, though these systems still pose challenges for non-expert users due to their complexity (Wang *et al*., 2020; Sobri *et al*., 2018). Despite advancements, interpretability remains a significant concern in the field (Tian, 2024; Wang *et al*., 2020).

Climate Change and Long-Term Forecasting

Climate change introduces significant uncertainties in long-term solar irradiance forecasting, particularly due to changes in atmospheric conditions like increased cloud cover and extreme weather events. These changes can disrupt historical patterns that machine learning models rely on, reducing the accuracy of predictions (Voyant *et al*., 2017). Traditional models may struggle to adapt to these evolving conditions, leading to potential inaccuracies in solar energy assessments (Radovan *et al.*, 2021). Integrating climate models with solar irradiance forecasting holds promise for improving accuracy, though the development of such models remains in its early stages (Belmahdi *et al*., 2023). Ongoing research focuses on refining these models to better account for climate-induced variability in solar energy generation (Yang *et al*., 2018).

Future Directions

The field of solar irradiance forecasting continues to evolve, driven by advancements in technology, data availability, and the need for improved accuracy in renewable energy systems. As the demand for solar energy grows, so too does the need for more efficient and precise forecasting methods. This section explores emerging techniques and research directions that hold promise for overcoming the challenges discussed in Section 5.

Emerging Techniques

Recent advancements in computational technologies, particularly quantum computing and artificial intelligence (AI), are revolutionizing solar irradiance forecasting. Quantum computing offers the potential to solve complex optimization problems more efficiently than classical methods, improving real-time forecasting as quantum algorithms advance (Sweeney *et al.*, 2019). In AI, machine learning techniques such as deep learning, reinforcement learning, and transfer learning are enhancing accuracy. Deep learning uncovers patterns from large datasets, reinforcement learning adapts models in response to real-time feedback, and transfer learning allows models trained in data-rich environments to be applied in data-scarce regions, improving forecasts in those areas (Wang *et al*., 2020; Radovan *et al*., 2021; Santos *et al*., 2022). These innovations are set to make solar irradiance forecasting more reliable and efficient, crucial for integrating solar energy into power grids (Yang *et al*., 2018).

Integration with Smart Grids

Integrating solar irradiance forecasting with smart grid technologies significantly improves the efficiency and reliability of renewable energy systems. Smart grids leverage real-time data to optimize energy distribution and maintain grid stability, making them ideal for managing the intermittency of solar energy (Avwioroko, 2024; Zhao and Lu, 2014). By incorporating accurate solar forecasts, energy providers can dynamically manage distributed energy resources (DERs) and demand response systems, enhancing grid stability and reducing reliance on fossil fuels during low solar generation periods (Santos *et al*., 2022; Sweeney *et al*., 2019). Additionally, blockchain technology facilitates secure and transparent energy transactions, further supporting solar integration within smart grids and improving energy trading practices (Sweeney *et al*., 2019; Cha *et al*., 2021).

Improving Data Acquisition and Real-Time Forecasting

Advancements in data acquisition technologies are key to improving the accuracy of solar irradiance forecasting. Recent research highlights the growing importance of IoT and ground-based sensors in solar forecasting and monitoring. Satellite-based systems are being enhanced with higher-resolution imaging and frequent updates, allowing for better monitoring of atmospheric conditions like cloud cover and aerosols, which significantly impact solar irradiance (Hashimoto and Yoshimoto, 2023; Radovan *et al*., 2021). Ground-based sensors and IoT devices further contribute by enabling real-time data collection across multiple locations, providing granular insights that boost the performance of machine learning models, particularly for short-term forecasting (Cha *et al*., 2021; Wang *et al*., 2020). The integration of big data technologies allows for the processing of large datasets, combining historical and real-time data with cloud computing to provide scalable, accurate solar forecasts (Fara *et al*., 2021; Belmahdi *et al*., 2023). These advancements are vital for enhancing the precision of solar energy predictions, supporting the transition to renewable energy (Chodakowska *et al*., 2023; Sobri *et al*., 2018).

Advances in Hybrid Models

Hybrid models that integrate physical and machine learning approaches are becoming essential in solar irradiance forecasting, combining the atmospheric simulations of Numerical Weather Prediction (NWP) with the pattern recognition capabilities of machine learning to enhance short-term forecast accuracy (Pathak *et al*., 2022). Recent advancements include the use of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), where CNNs process spatial data like satellite imagery, and RNNs handle time-series data, optimizing both spatial and temporal predictions (Yang *et al*., 2018; Fara *et al*., 2021). Additionally, ensemble learning techniques are being employed within hybrid models to improve prediction robustness in regions with variable weather conditions by aggregating multiple forecasting outputs (Santos *et al*., 2022; Belmahdi *et al*., 2023). This comprehensive approach significantly enhances both the accuracy and reliability of solar irradiance forecasting.

Interdisciplinary Research and Collaboration

The future of solar irradiance forecasting is set, as it has started to benefit greatly from interdisciplinary collaboration, particularly among experts in meteorology, data science, computer science, and energy systems engineering. By combining advanced meteorological models with artificial intelligence (AI) techniques, researchers can enhance both the accuracy and interpretability of forecasts (Tian, 2024; Wang *et al*., 2020). Collaboration between computer scientists and energy engineers is essential for making these models scalable and practically applicable, allowing them to be more accessible to energy providers (Radovan *et al*., 2021; Santos *et al*., 2022). As climate change continues to impact weather patterns, adapting forecasting models to account for these shifts is increasingly important. Incorporating climate projections into solar forecasting models will help ensure the resilience of renewable energy systems in a changing environment (Soni *et al*., 2011; Sweeney *et al*., 2019). This interdisciplinary approach will be crucial for improving forecasting accuracy and ensuring the sustainable integration of solar energy into power grids (Zhao and Lu, 2014; Chodakowska *et al*., 2023).

CONCLUSION

Solar irradiance forecasting is a critical component of optimizing solar energy systems and integrating renewable energy sources into power grids. Over the years, various methods—ranging from physical models like Numerical Weather Prediction (NWP) to advanced machine learning-based models—have been developed to improve the accuracy and efficiency of these forecasts. However, despite the significant progress, challenges such as data availability, model complexity, geographic and temporal resolution, and uncertainty in weather predictions persist. Future research in solar irradiance forecasting should focus on integrating emerging technologies like IOT, quantum computing, artificial intelligence, and big data to improve both the accuracy and speed of forecasts. Additionally, interdisciplinary collaboration will be essential in addressing the remaining challenges, such as climate change's impact on irradiance patterns and the need for scalable, real-time forecasting solutions. Ultimately, accurate solar irradiance forecasting is crucial for the growth of solar energy and the broader transformation towards a low-carbon future.

Summary of Key Findings

Machine learning models, particularly Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks, have shown considerable potential in solar irradiance forecasting. Hybrid models that combine physical meteorological simulations with machine learning techniques further enhance accuracy by leveraging the strengths of both approaches (Belmahdi *et al*., 2023; Voyant *et al*., 2017). However, these models often require significant computational resources and large datasets, complicating their real-time application (Wang *et al*., 2020; Rojas-Campos *et al*., 2022). On the other hand, traditional statistical models like ARIMA are computationally efficient but lack the accuracy needed to capture the non-linear dynamics of solar irradiance (Alsharif *et al*., 2019). Physical models like Numerical Weather Prediction (NWP) excel in medium- to long-term forecasting but struggle with short-term variability due to the unpredictability of weather patterns (Hashimoto and Yoshimoto, 2023; Sweeney *et al*., 2019). Thus, while machine learning and hybrid models offer advancements in accuracy, their computational demands and forecasting horizons remain critical trade-offs.

RECOMMENDATIONS

Advancing solar energy forecasting techniques requires several strategic steps. The findings from this review implied that adopting advanced solar irradiance forecasting methods can significantly enhance grid integration and stability. Adopting hybrid models is crucial, as they combine physical principles and machine learning to improve accuracy in areas like grid management. Investing in data acquisition technologies, such as IoT sensors and high-resolution satellite imagery, along with improved data-sharing infrastructures, is essential, particularly for regions with limited meteorological data. Balancing model complexity with practicality is also important; while complex models like CNN-LSTM hybrids offer high accuracy, they costly to implement while simpler models like ARIMA may be more suitable for smaller operators and cheaper. Improving the interpretability of machine learning models is key to building trust in forecast outputs, especially for critical applications like energy grid management. Finally, it is paramount for researchers and practitioners in the choice of a model for forecasting to engage in the balancing of the application need and model implementation costs to ensure a sustainable and cost-effective solution supporting the global transition toward a sustainable and reliable energy future.

REFERENCES

Aggarwal, S. and Saini, L. M. (2014). Solar energy prediction using linear and non-linear regularization models: a study on AMS (American meteorological society) 2013–14 solar energy prediction contest. Energy, 78, 247-256. https://doi.org/10.1016/j.energy.2014.10.012

Ahmed, R., Mekhilef, S., Shah, N. M., Mokhlis, H., and Fathi, S. (2020). Machine learning and metaheuristic techniques for PV power forecasting: A review. *Journal of Cleaner Production, 253,* 119670. https://doi.org/10.1016/j.jclepro.2020.119670

Ahmed, R., Sreeram, V., Mishra, Y., and Arif, M. D. (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.

Ajeel, S. M. and Hashem, H. (2020). Comparison some robust regularization methods in linear regression via simulation study. Academic Journal of Nawroz University, 9(2), 244. https://doi.org/10.25007/ajnu.v9n2a818

Akhter, M. N., Mekhilef, S., Mokhlis, H., and Shah, N. M. (2016). Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. *Renewable and Sustainable Energy Reviews, 61*, 384–397. https://doi.org/10.1016/j.rser.2016.04.024

Alsharif, M. H., Younes, M. K., and Kim, J. (2019). Time series arima model for prediction of daily and monthly average global solar radiation: the case study of Seoul, south Korea. Symmetry, 11(2), 240. https://doi.org/10.3390/sym11020240

Al-Sharoot, H., M., K. Mohammed, F., and N. Mayali, H. (2023). Lasso quantile principal component regression. Journal of Al-Qadisiyah for Computer Science and Mathematics, 15(4). https://doi.org/10.29304/jqcsm.2023.15.41356

Avwioroko, A., Ibegbulam, C., Afriyie, I., and Fesomade, A. T. (2024). Smart grid integration of solar and biomass energy sources. European Journal of Computer Science and Information Technology, 12(3), 1-14. https://doi.org/10.37745/ejcsit.2013/vol12n3114

Belmahdi, B., Louzazni, M., Marzband, M., and Bouardi, A. E. (2023). Global solar radiation forecasting based on hybrid model with combinations of meteorological parameters: Morocco case study. Forecasting, 5(1), 172-195. https://doi.org/10.3390/forecast5010009

Biktash, L. (2017). Long-term global temperature variations under total solar irradiance, cosmic rays, and volcanic activity. Journal of Advanced Research, 8(4), 329-332. https://doi.org/10.1016/j.jare.2017.03.002

Bouchouicha, K. and Bachari, N. E. I. (2023). Simulation of a clear sky satellite image in water vapor and infrared satellite m.s.g channel's. Journal of Renewable Energies, 15(4). https://doi.org/10.54966/jreen.v15i4.355

Cha, J., Kim, M. K., Lee, S., and Kim, K. S. (2021). Investigation of applicability of impact factors to estimate solar irradiance: comparative analysis using machine learning algorithms. Applied Sciences, 11(18), 8533. https://doi.org/10.3390/app11188533

Chodakowska, E., Nazarko, J., Nazarko, Ł., Ahmad, H. S., Abendeh, R. M., and Alawneh, R. (2023). Arima models in solar radiation forecasting in different geographic locations. Energies, 16(13), 5029. https://doi.org/10.3390/en16135029

Conor, S., Bessa, R. J., Browell, J., and Pinson, P. (2019). The future of forecasting for renewable energy. *WIREs Energy and Environment*, e365. https://doi.org/10.1002/wene.365

Darian, M. M. M. and Ghorreshi, A. M. (2020). Comparison of the effect of temperature parameter on tracking and fixed photovoltaic systems: a case study in Tehran, Iran. Scientia Iranica, 0(0), 0-0. https://doi.org/10.24200/sci.2020.55173.4102

Dazhi, Y., Kleissl, J., Gueymard, C. A., Pedro, H. T. C., and Coimbra, C. F. M. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy, 168*, 60-101. https://doi.org/10.1016/j.solener.2017.11.023

Du, J., Min, Q., Zhang, P., Guo, J., Yang, J., and Yin, B. (2018). Short-term solar irradiance forecasts using sky images and radiative transfer model. Energies, 11(5), 1107. https://doi.org/10.3390/en11051107

Emmert-Streib, F. and Dehmer, M. (2019). High-dimensional lasso-based computational regression models: regularization, shrinkage, and selection. Machine Learning and Knowledge Extraction, 1(1), 359-383. https://doi.org/10.3390/make1010021

Fara, L., Diaconu, A., Crăciunescu, D., and Fara, S. (2021). Forecasting of energy production for photovoltaic systems based on ARIMA and ANN advanced models. International Journal of Photoenergy, 2021, 1-19. https://doi.org/10.1155/2021/6777488

Gallego-Castillo, C., Cuerva-Tejero, A., and Lopez-Garcia, O. (2015). Wind power ramp forecasting: A review. *Renewable Energy, 83*, 940-953. https://doi.org/10.1016/j.renene.2015.05.045

Govender, P. and Sivakumar, V. (2019). Investigating diffuse irradiance variation under different cloud conditions in Durban, using k-means clustering. Journal of Energy in Southern Africa, 30(3), 22-32. https://doi.org/10.17159/2413-3051/2019/v30i3a6314

Gupta, A., Sharma, A., & Goel, A. (2017). Review of regression analysis models. International Journal of Engineering Research And, V6(08). https://doi.org/10.17577/ijertv6is080060

He, Y., Zhao, Y., & Tsui, K. L. (2019). Exploring influencing factors on transit ridership from a local perspective. Smart and

Resilient Transportation, 1(1), 2-16. https://doi.org/10.1108/srt-06-2019-0002

Hejase, H. and Assi, A. (2012). Time-series regression model for prediction of mean daily global solar radiation in al-ain, UAE. ISRN Renewable Energy, 2012, 1-11. https://doi.org/10.5402/2012/412471

Hashimoto, A. and Yoshimoto, K. (2023). Development of a short-term solar irradiance forecasting using satellite image in combination with numerical weather prediction model. Electrical Engineering in Japan, 216(3). https://doi.org/10.1002/eej.23432

Hastuti, M. I., Min, K., and Lee, J. (2023). Improving radar data assimilation forecast using advanced remote sensing data. Remote Sensing, 15(11), 2760. https://doi.org/10.3390/rs15112760

Heinemann, D., Lorenz, E., and Girodo, M. (2006). Forecasting of solar radiation. *Solar energy resource management for electricity generation from local level to global scale. Nova Science Publishers, New York*, 83-94.

Heinemann, D., Lorenz, E., and Kramer, O. (2020). Numerical weather prediction-based forecasting methods for solar power generation. *Renewable Energy*, *145*, 2530–2536. https://doi.org/10.1016/j.renene.2020.01.115

Ibrahim, M., Muhammed, A., Ahmad, M. A., & Adamu, A. (2022). An Improved Accuracy for The Forecasting of Power Generation Over A Long-Term Horizon. *FUDMA JOURNAL OF SCIENCES*, *6*(6), 318-324.

Ibrahim, S., Daut, I., Irwan, Y. M., Irwanto, M., Gomesh, N., and Farhana, Z. (2012). Linear regression model in estimating solar radiation in Perlis. *Energy Procedia*, *18*, 1402-1412.

Jozefowicz, R., Zaremba, W., and Sutskever, I. (2015). An empirical exploration of recurrent network architectures. In Proceedings of the International Conference on International Conference on Machine Learning (pp. 2342–2350). New York: ACM

Kashyap, A., Raza, A., Mekhilef, S., Mokhlis, H., and Shah, N. M. (2015). Solar radiation forecasting using multiple parameters neural networks: A review. *Renewable and Sustainable Energy Reviews*, *65*, 811–830. https://doi.org/10.1016/j.rser.2016.07.048

Kayanan, M. and Wijekoon, P. (2020). Variable selection via biased estimators in the linear regression model. Open Journal of Statistics, 10(01), 113-126. https://doi.org/10.4236/ojs.2020.101009

Lei, W., He, Q., Yang, L., and Jiao, H. (2022). Solar photovoltaic cell parameter identification based on improved honey badger algorithm. Sustainability, 14(14), 8897. https://doi.org/10.3390/su14148897

Lennon, A., Lunardi, M. M., Hallam, B., and Dias, P. (2022). The aluminium demand risk of terawatt photovoltaics for net zero emissions by 2050. Nature Sustainability, 5(4), 357-363. https://doi.org/10.1038/s41893-021-00838-9

Liu, J., Bray, M., and Han, D. (2013). A study on WRF radar data assimilation for hydrological rainfall prediction. Hydrology and Earth System Sciences, 17(8), 3095-3110. https://doi.org/10.5194/hess-17-3095-2013

Liu, J., Zhang, Q., and Macián-Juan, R. (2024). Enhancing interpretability in neural networks for nuclear power plant fault diagnosis: A comprehensive analysis and improvement approach. *Progress in Nuclear Energy*, *174*, 105287.

Lu, Z. (2023). Photovoltaic power forecasting approach based on ground-based cloud images in hazy weather. Sustainability, 15(23), 16233. https://doi.org/10.3390/su152316233

Lunche, W., Kisi, O., Zounemat-Kermani, M., Salazar, G. A., and Gong, W. (2016). Solar radiation prediction using different techniques: Model evaluation and comparison. *Renewable and Sustainable Energy Reviews, 61*, 384-397. https://doi.org/10.1016/j.rser.2016.04.024

Melhem, R. and Shaker, Y. (2023). Optimum tilt angle and solar radiation of photovoltaic modules for gulf collaboration council countries. International Journal of Energy Research, 2023, 1-17. https://doi.org/10.1155/2023/8381696

Nguyen, K. C., Katzfey, J., Riedl, J., and Troccoli, A. (2017). Potential impacts of solar arrays on regional climate and on array efficiency. International Journal of Climatology, 37(11), 4053-4064. https://doi.org/10.1002/joc.4995

Nur, A. R., Jaya, A. K., and Siswanto, S. (2023). Comparative analysis of ridge, lasso, and elastic net regularization approaches in handling multicollinearity for infant mortality data in south sulawesi. Jurnal Matematika, Statistika Dan Komputasi, 20(2), 311-319. https://doi.org/10.20956/j.v20i2.31632

Oladunjoye, O. O., Olasoji, Y. O., Adedeji, K. B., Oladunjoye, O. A., and Olebu, C. (2022). A solar energy control system for on-grid energy storage device. European Journal of Electrical Engineering and Computer Science, 6(3), 1-6. https://doi.org/10.24018/ejece.2022.6.3.429

Ortega, J. L., and Aguillo, I. F. (2014). Mapping academic collaboration in Europe using Google Scholar citations. *Journal of Informetrics, 8(3)*, 654–662. https://doi.org/10.1016/j.joi.2014.04.005

Pathak, J., Subramanian, S., Harrington, P. d. B., Raja, S. K. S., Chattopadhyay, A., Mardani, M., … and Anandkumar, A. (2022). Fourcastnet: a global data-driven high-resolution weather model using adaptive fourier neural operators. https://doi.org/10.48550/arxiv.2202.11214

Radovan, A., Šunde, V., Kučak, D., and Ban, Ž. (2021). Solar irradiance forecast based on cloud movement prediction. Energies, 14(13), 3775. https://doi.org/10.3390/en14133775

Rady, E. A. and Mahmoud, H. (2018). Modeling of biological data based on regression methods. Zagazig Veterinary Journal, 46(2), 146-153. https://doi.org/10.21608/zvjz.2018.14387

Rojas-Campos, A., Langguth, M., Wittenbrink, M., and Pipa, G. (2022). Deep learning models for generation of precipitation maps based on numerical weather prediction.. https://doi.org/10.5194/egusphere-2022-648

Sansine, V., Ortega, P., Hissel, D., and Ferrucci, F. (2023). Hybrid deep learning model for mean hourly irradiance probabilistic forecasting. *Atmosphere*, *14*(7), 1192.

Santos, D. S. d. O., Neto, P. S. G. d. M., Oliveira, J. F. L. d., Siqueira, H. V., Barchi, T. M., Lima, A. R., … and Marinho, M. H. N. (2022). Solar irradiance forecasting using dynamic ensemble selection. Applied Sciences, 12(7), 3510. https://doi.org/10.3390/app12073510

Seo, Y. A. and Cha, J. (2023). Precipitation probability prediction through NWP bias correction for south korea using random forest. International Journal on Advanced Science, Engineering and Information Technology, 13(3), 935-942. https://doi.org/10.18517/ijaseit.13.3.18224

Smilevski, M. (2020). Applying recent advances in Visual Question Answering to Record Linkage. *arXiv preprint arXiv:2007.05881*.

Sobrina, S., Koohi-Kamali, S., and Rahim, N. A. (2018). Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management*, *156*, 459-497. https://doi.org/10.1016/j.enconman.2017.11.019

Soni, V. K., Pandithurai, G., and Pai, D. S. (2011). Evaluation of long-term changes of solar radiation in india. International Journal of Climatology, 32(4), 540-551. https://doi.org/10.1002/joc.2294

Sutikno, S., Cahyoko, F. D., Putra, F. W., Makmur, E. E. S., Hanggoro, W., Taufik, M. R., … and Aza, V. (2024). Calibration Indonesian-numerical weather prediction using geostatistical output perturbation. Jurnal Meteorologi Dan Geofisika, 24(2), 105-113. https://doi.org/10.31172/jmg.v24i2.1037

Sweeney, C., Bessa, R. J., Browell, J., and Pinson, P. (2019). The future of forecasting for renewable energy: Perspectives and challenges. *WIREs Energy and Environment*, *8(3)*, e365. https://doi.org/10.1002/wene.365

Tang, Q., Zhang, H., & Gong, S. (2020). Bayesian regularized quantile regression analysis based on asymmetric laplace distribution. Journal of Applied Mathematics and Physics, 08(01), 70-84. https://doi.org/10.4236/jamp.2020.81006

Tian, Y., Zhao, Y., Li, J., Chen, B., Deng, L., and Wen, D. (2024). East asia atmospheric river forecast with a deep learning method: gan-unet. Journal of Geophysical Research: Atmospheres, 129(5). https://doi.org/10.1029/2023jd039311

Trapero, J. R., Kourentzes, N., and Martín, A. (2015). Short-term solar irradiation forecasting based on dynamic harmonic regression. Energy, 84, 289-295. https://doi.org/10.1016/j.energy.2015.02.100

Trojáková, A., Mile, M., and Tudor, M. (2019). Observation preprocessing system for rc lace (oplace). Advances in Science and Research, 16, 223-228. https://doi.org/10.5194/asr-16-223-2019

Wang, H., Liu, Y., Zhou, B., and Li, C. (2020). Taxonomy research of artificial intelligence for deterministic solar power forecasting. *Energy Conversion and Management*, *214*, 112909. https://doi.org/10.1016/j.enconman.2020.112909

Watagoda, L. C. R. P., Arnholt, A. T., & Don, H. S. R. A. (2021). hrlr regression. RMS: Research in Mathematics &Amp; Statistics, 8(1). https://doi.org/10.1080/27658449.2021.1921904

Xin, S. J. and Khalid, K. (2018). Modelling house price using ridge regression and lasso regression. International Journal of Engineering &Amp; Technology, 7(4.30), 498. https://doi.org/10.14419/ijet.v7i4.30.22378

Xu, M. (2024). Sales prediction based on lasso regression. Highlights in Science, Engineering and Technology, 88, 343-349. https://doi.org/10.54097/p9hyrk70

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., and Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, *105*, 569-582. https://doi.org/10.1016/j.renene.2016.12.095

Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T. C., and Coimbra, C. F. M. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy*, *168*, 60-101. https://doi.org/10.1016/j.solener.2017.11.023

Zhang, L., Wang, S., Yu, Z., He, C., and Jin, X. (2014). Development of an instant correction and display system of numerical weather prediction products in china. Chinese Geographical Science, 24(6), 682-693. https://doi.org/10.1007/s11769-014-0672-7

Zhang, T., Cheng, C., and Gao, P. (2019). Permutation entropy-based analysis of temperature complexity spatial-temporal variation and its driving factors in china. Entropy, 21(10), 1001. https://doi.org/10.3390/e21101001

Zhang, Y., and Wang, J. (2019). A review of solar irradiance forecasting methods. Renewable and Sustainable Energy Reviews, 101, 1-12.

Zhang, Z., Zhang, T., Zhang, R., Zhu, X., Wu, X., Tan, S., … & Jian, Z. (2024). Predicting colorectal cancer risk: a novel approach using anemia and blood test markers. Frontiers in Oncology, 14. https://doi.org/10.3389/fonc.2024.1347058

Zhao, L. and Lu, F. (2014). Study on effects of solar generation on power grid. Advanced Materials Research, 986-987, 560-563. https://doi.org/10.4028/www.scientific.net/amr.986-987.560

Zhu, H., Li, X., Sun, Q., Nie, L., Yao, J., and Zhao, G. (2015). A power prediction method for photovoltaic power plant based on wavelet decomposition and artificial neural networks. *Energies*, *9*(1), 11.

Zwane, N., Tazvinga, H., Botai, C., Murambadoro, M., Botai, J., De Wit, J., ... and Mabhaudhi, T. (2022). A bibliometric analysis of solar energy forecasting studies in Africa. *Energies*, *15*(15), 5520.