Heliyon 10 (2024) e34800

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

Empowering distribution system operators: A review of distributed energy resource forecasting techniques

Nande Fose [a,*], Arvind R. Singh [a], Senthil Krishnamurthy [a], Mukovhe Ratshitanga [a], Prathaban Moodley [b]

[a] Department of Electrical , Electronics and Computer Engineering, Cape Peninsula University of Technology, Bellville Campus, South Africa
[b] South African National Energy Development Institute (SANEDI), South Africa

ARTICLE INFO

Keywords:
DER'S
DSOs
Load forecasting
Bi-directional power flow

ABSTRACT

Effective management of Distributed Energy Resources (DERs) and optimization of grid operations are crucial responsibilities of Distribution System Operators (DSOs). Hence, this comprehensive critical review aims to analyze the current state of DER forecasting practices for DSOs and their implications for achieving the SDG goals. These goals underscore the significance of clean and accessible energy, advancements in infrastructure, sustainable urban development, climate change mitigation, and collaborative partnerships. The review's core focuses on the DER forecasting techniques employed by DSOs. It explores various aspects, including data collection methods, load forecasting models, DER generation forecasting, aggregation and integration techniques, and the role of advanced technologies like machine learning and artificial intelligence. The review highlights the critical role of accurate DER forecasts in optimizing grid operations, managing energy flows, and facilitating the integration of renewable energy sources. Furthermore, the review examines the implications of DER forecasting for DSOs in achieving the SDGs. It discusses how DER forecasting facilitates the transition to affordable and clean energy, enhances industry innovation and infrastructure, builds sustainable cities and communities, drives climate action efforts, and fosters stakeholder partnerships. However, the review also identifies challenges and limitations in DER forecasting, including data availability, forecasting accuracy, uncertainty management, and regulatory barriers. It emphasizes further research and development in improved forecasting algorithms, advanced data analytics, and enhanced communication and coordination mechanisms. Finally, this comprehensive critical review highlights the importance of DER forecasting for DSOs in achieving the SDGs. Accurate forecasting can promote sustainable and clean energy practices, drive innovation, build resilient communities, mitigate climate change, and foster collaborative partnerships. The review emphasizes the necessity of advancing DER forecasting techniques and addressing associated challenges to fully realize the potential of DERs in contributing to a sustainable and inclusive future.

This comprehensive critical review aims to analyze the current state of DER forecasting practices for DSOs and their implications for achieving the SDG goals. As Distributed Energy Resources (DERs) play an increasingly significant role in the transition to sustainable energy systems, accurate forecasting techniques are essential for optimizing grid operations and facilitating the integration of renewable energy sources. By effectively managing DERs, Distribution System Operators (DSOs) contribute to the advancement of several SDGs, including affordable and clean energy (SDG 7), sustainable infrastructure (SDG 9), climate action (SDG 13), and

* Corresponding author.
E-mail addresses: nandetillax@gmail.com (N. Fose), krishnamurthys@cput.ac.za (S. Krishnamurthy).

https://doi.org/10.1016/j.heliyon.2024.e34800
Received 7 November 2023; Received in revised form 2 July 2024; Accepted 17 July 2024
Available online 22 July 2024

partnerships for the goals (SDG 17). This review explores the intersection of DER forecasting with the SDGs, highlighting how forecasting initiatives can support national and global efforts toward sustainable development by providing insights into energy demand, grid stability, and renewable energy integration. The goals and targets are derived from an analysis of current trends and the identification of potential development scenarios by 2030. Both optimistic and pessimistic projections are utilized for communicating with the general public and national governments concerning DSM network planning. Utilizing data from various nations enables the identification of effective strategies and the prediction of similar trends in other areas. Simultaneously, the magnitude of activities related to Sustainable Development Goals (SDGs) enables the improvement and efficient organization of data gathering on a global basis. This, in turn, provides a foundation for future forecasting endeavours.

1. Introduction

1.1. Background and motivation

Distributed Energy Resources (DERs) have become essential components of the modern power grid, facilitating the integration of renewable energy sources and promoting sustainable development. Significantly, DER forecasting, an essential aspect of Distribution System Operators (DSOs) operations, can accelerate nationwide efforts to achieve the Sustainable Development Goals (SDGs) by promoting clean energy, enhancing grid resilience, and mitigating climate change. A developing country's socioeconomic and infrastructural progress can be closely linked to energy generation, transmission, and consumption [1,2]. This makes South Africa's situation intriguing, as it is Africa's largest energy consumer and one of its most developed nations [3]. Situated at the southernmost tip of the continent, South Africa shares borders with Namibia, Zimbabwe, Botswana, and Lesotho, with coastlines along the South Atlantic and Indian oceans.

The country boasts abundant mineral resources like platinum, diamonds, and gold, with a population of around 56 million [4]. Among these resources, coal plays a significant role, ranking South Africa seventh in global coal reserves according to the International Energy Agency's Energy Atlas [5]. This fossil fuel accounts for over 85 % of the nation's electricity generation (Understanding Electricity, 2003) but has also positioned South Africa among the world's top ten greenhouse gas emitters [6]. Nonetheless, South Africa has undertaken initiatives to curb CO_2 emissions and promote renewable energy sources. This study examines these advancements and the challenges encountered. For a comprehensive understanding of renewable energy growth in South Africa, an exploration of the nation's power generation history is essential. Before delving into this history in section 2, it's worth noting a recent review that assessed renewable energy progress in key African economies: Nigeria, South Africa, and Egypt [7]. South Africa's renewable energy strategy was deemed the most promising among these. However, the review only briefly addressed the developments in South Africa's real estate sector. Therefore, this article builds upon the previous overview [7] by offering more comprehensive insights into the expansion of renewable energy sources in South Africa. This includes focusing on biomass, wind, and solar energy for electricity and heat generation, Fig. 1 illustrates energy consumption in SA per source.

According to the IEA, low-carbon sources are set to cover almost all the growth in global electricity demand by 2025. The entire power system has transited towards renewable energy power generation integrated with intelligent communication devices. The outlook for 2023 to 2025 shows that renewable power generation is set to increase more than all other sources combined, with an annualized growth of over 9 %. Renewables will make up over one-third of the global generation mix by 2025. This trend is supported by government pledges to increase spending on renewables as part of economic recovery plans such as the Inflation Reduction Act in the United States. The global changes in power generation sources are represented in Fig. 2 below.

The statistical reports of global and regional percentage change in electricity generation demand are represented in Fig. 3,

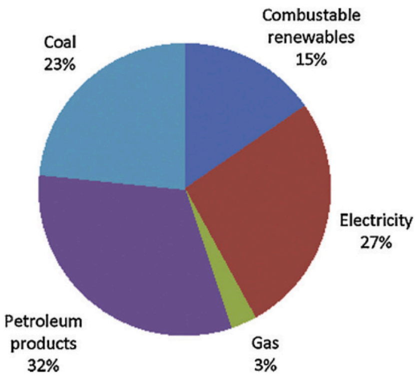

Fig. 1. Energy consumption by source in South Africa [8].

respectively. One of the prominent solutions for reducing carbon emissions and protecting the environment by integrating high participation of distributed energy sources in the traditional microgrid [1]. The conceptual representation of the energy management system integrated with DERs is represented in Fig. 4. Microgrids transfer energy and auxiliary commodities to the electric grid to balance supply and demand. DER integration into a distribution network opens up new technological, legal, and economic techniques for safe and effective operation. DERs make distribution networks more flexible and allow them to deliver transmission system services.

Isolated microgrids need upgraded distributed producing plants, feeders, and demand measurement and analysis systems to operate an active distribution system. Distributed energy sources (DERs) integration may significantly aid in the development of low-carbon pathways. Smaller power-producing facilities that are situated closer to the point of consumption are referred to as distributed energy sources. Distributed generation (DG) is defined as power generation at or near the end user. Decentralized generation often called on-site generation or distributed energy is a method of producing energy that is not only useful for generating electricity but also for co-generating it and producing heat. For the world's energy problems, DG is seen as a possible answer. Distributed generation (DG) or distributed energy systems (DES) provide various benefits over centralized power plants. Most DESs, especially those used in off-grid settings, are based on renewable energy sources, hence, this movement strongly backs DESs. DES may operate on or off the grid, using various energy sources and technologies. As a result, distributed generating systems are not only seeing tremendous expansion in installed capacity but also making quick advances in the fields of technology and regulation. Most of the world's distribution generation comes from renewable sources, rapidly improving efficiency, adaptability, and cost-competitiveness compared to traditional energy infrastructure. The amount of distributed generation contributed significantly to the increase in global installed renewable capacity from 1430 MW in 2019 to 1668 MW in 2020 [9].

These sources frequently use renewable energy techniques, which emit fewer greenhouse gases into the atmosphere than conventional fossil fuel-based power plants. DERs are grid-connected, small-scale electricity supply or demand resources. They are power production resources near load centres that can add value to the system individually or collectively. Physical DERs under 10 MW include diesel or natural gas engines, microturbines, solar arrays, tiny wind farms, battery energy storage systems, and more. Electric utilities, independent power providers, and local enterprises can run them. The utility starts and stops them like major central power plants. Integrating Renewable Energy Sources (RES) into microgrids is crucial for sustainable development. By relying on clean and renewable energy, such as solar, wind, and hydropower, microgrids with RES enable reduced greenhouse gas emissions, environmental preservation, and enhanced energy security. Additionally, this approach fosters economic growth and community empowerment while promoting global collaboration for sustainable energy solutions. By advancing technology and resource efficiency, RES integration into microgrids aligns with the principles of sustainable development and represents a pivotal pathway toward a greener and more resilient energy future. The principal incentive for electricity customers to establish a connection to the electrical grid is to have access to electric service.

In essence, the electrical grid serves to deliver electrical power to consumer installations. The process of connecting to the electrical grid, commonly referred to as interconnection or grid connectivity, offers several advantages, including enhanced power quality and dependability, as well as effective regulation of voltage and frequency, along with the reduction of harmonic distortion. Furthermore, it offers these services at a very affordable price, with minimal environmental adverse effects. Consumers who choose self-generation of energy also have several advantages, including the opportunity to utilize the power grid to enhance stability within their local consumer power system and achieve improved efficiency. In electricity, energy is transmitted through a unified electrical grid, regardless of whether the source is locally or centrally linked. It is important to note that in the absence of this grid, consumers who generate power locally would experience a much-diminished quality of service. In the absence of grid connectivity, customers who rely on self-generation may see a decrease in the dependability and functionality of their power system. However, it is possible to mitigate this issue by making significant investments in the local power system to ensure that the quality of service is comparable to that provided by the grid.

2. Introduction to the role of distribution system operator in managing DER integration

The electrical power sector is presently experiencing a paradigm shift towards a more intelligent and environmentally conscious approach. This process encompasses the integration of innovative technologies and digitalization, with its primary objective being the

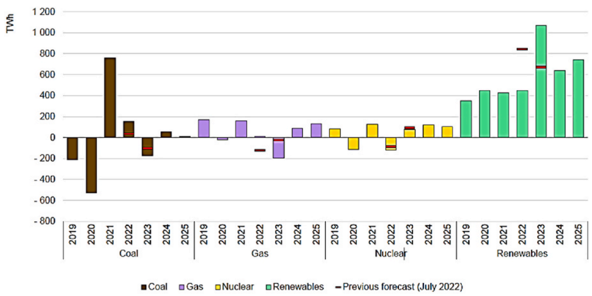

Fig. 2. Global change in electricity generation by different input sources [9].

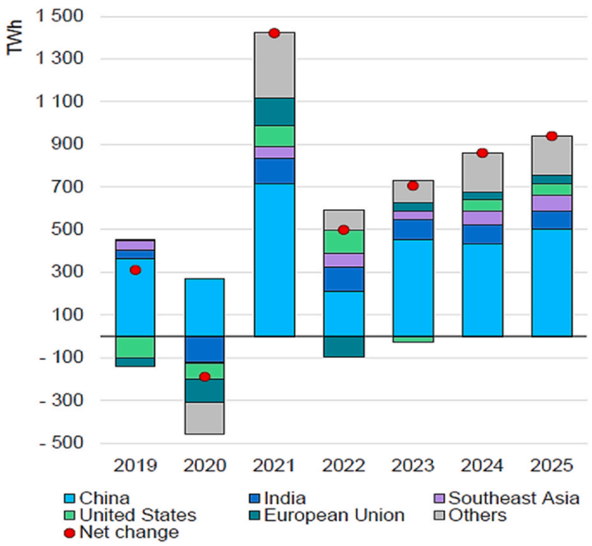

Fig. 3. Electricity demand by region, 2019–2025 [9].

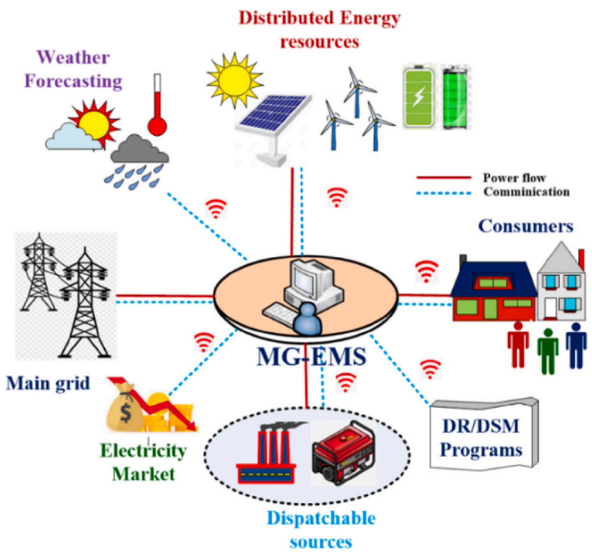

Fig. 4. Conceptual representation of microgrid energy management system integrated DERs [10].

preservation of the environment and the well-being of humanity. Adopting the DSO model can optimize the addition of renewable energy sources and expedite the process of decarbonizing the economy while enhancing the advantages experienced by customers. The management of the integration of Distributed Energy Resources (DERs) into the electrical grid is a critical function performed by Distribution System Operators (DSOs) in the contemporary energy environment. Distributed energy resources (DERs) cover various small-scale, decentralized energy sources, including solar panels, wind turbines, energy storage devices, and electric cars [11]. The increasing utilization of distributed energy resources (DERs) has led to the crucial role distributed system operators (DSOs) play in maintaining the distribution grid's stability, reliability, and efficiency. This paper provides an overview of the responsibilities and functions of Distribution System Operators (DSOs) in effectively managing the integration of Distributed Energy Resources (DERs) into the existing power grid. The expanding involvement of Distribution System Operators (DSOs) may be attributed to the rise of distributed energy resources, including distributed generation, demand-side response, and storage [12]. Therefore, distribution system operators (DSOs) may be granted the opportunity to utilize the decentralized flexibilities associated with their grid to enhance the performance of both the distribution grid itself and the end customers. If the regulatory framework permits, distributed energy resources might be operated by Distribution System Operators (DSOs) in their newly assigned responsibilities.

The utilization of distributed flexibilities serves a dual purpose, namely enhancing the efficiency of distribution networks and reducing the necessity for future grid improvements. The growing use of Distributed Energy Resources (DERs) can provide more variability and bidirectional power flow inside the electrical grid, impacting the conventional planning and management of distribution and transmission networks. Moreover, the anticipated expansion of distributed energy resources (DERs) will result in grid congestion inside the distribution system, necessitating proactive management strategies [12]. This necessitates re-evaluating the traditional responsibilities of Distribution System Operators (DSOs), who have traditionally been responsible for planning, maintaining, and managing networks and mitigating supply interruptions. To optimize the utilization of distributed energy resources (DERs) integrated into the distribution network, distribution system operators (DSOs) have the potential to enhance their involvement as proactive system operators alongside their existing responsibilities as network operators. Distribution system operators can acquire flexibility services from their network users, such as voltage support and congestion management, to delay the need for network expenditures. Furthermore, distribution system operators (DSOs) have the potential to help in the form of reactive power support to transmission system operators (TSOs). For instance, Distribution System Operators (DSOs) have the potential to collaborate with Transmission System Operators (TSOs) to establish uniform market offerings for the procurement of services through flexible mechanisms [13]. This collaboration would involve the development of standardized market goods and specifying technical requirements for participation in specialized markets. Distribution system operators (DSOs) have the potential to utilize various flexibility services, including but not limited to the management of local congestion and non-frequency ancillary services such as voltage control. On the other hand, transmission system operators (TSOs) would be exclusively responsible for providing frequency auxiliary services.

In many cases, electricity networks are managed by multiple system operators (SOs), each with distinct objectives and strategies, though this can vary depending on the nature of the grid and its operational context. Island grids, for example, often have unique management structures compared to interconnected mainland grids.". This revision acknowledges the existence of diverse management structures in electricity networks, emphasizing that such diversity is not universal across all networks. Island grids are highlighted as a specific example illustrating potential differences in management approaches. The electrical system, in contrast, does not recognize any boundaries between sovereign states but adheres only to the principles of physical physics. Hence, enhanced collaboration among system operators (SOs) has the promise of enhancing the effectiveness and reliability of grid operations on a broader scale encompassing all SOs. As seen in Fig. 6 (a), this phenomenon is observed in interactions involving transmission system operators (TSOs) and between TSOs and distribution system operators (DSOs). The latter has particular significance in light of the ongoing transformation in the energy system since wind and solar facilities,

which possess greater generating capacity and provide auxiliary services, are more frequently integrated with distribution system operators (DSOs) than traditional plants. Grid operation techniques that concentrate the data and control the system operator (SO) boundaries offer optimal coordination of variables. However, these practices fail to uphold the principles of sovereignty and data integrity for Service Organizations (SOs), who function autonomously and exercise caution when exchanging data related to their essential infrastructure. In contrast, it is recommended that grid operating approaches restrict the amount of data that has to be communicated and empower system operators to manage the flexibilities of their grids. Furthermore, there is a demand for near-optimal system states, even if the system operators have limited information about the entire system, as represented in Fig. 5 (b). The introduction section has been designed in the form of CARS framework model.

Previous research has primarily focused on centralized energy forecasting techniques, which may not adequately capture the complexities of distributed energy resources (DERs) and their impact on load forecasting at the distribution level. While centralized models offer valuable insights, they often fail to account for the dynamic nature of DERs, including their intermittent generation patterns and the influence of local environmental conditions. Understanding the challenges of load forecasting in the context of distributed energy resources is akin to navigating a rapidly changing landscape, where traditional landmarks may no longer provide reliable guidance. Just as sailors once relied on stars for navigation, distribution system operators (DSOs) must adapt their forecasting techniques to navigate the evolving energy landscape shaped by DER integration. Recent advancements in forecasting methodologies have begun to address the unique challenges posed by DERs, with researchers exploring novel approaches such as machine learning algorithms, probabilistic modelling, and data fusion techniques. These studies underscore the need for tailored forecasting solutions that account for the decentralized nature of DERs and their interactions with the grid. Despite the evolving nature of load forecasting challenges, the fundamental goal remains unchanged: to accurately predict electricity demand to ensure grid stability, reliability, and efficiency. Like their predecessors, modern forecasting techniques aim to harness data-driven insights to anticipate fluctuations in demand, optimize resource allocation, and support informed decision-making by DSOs.

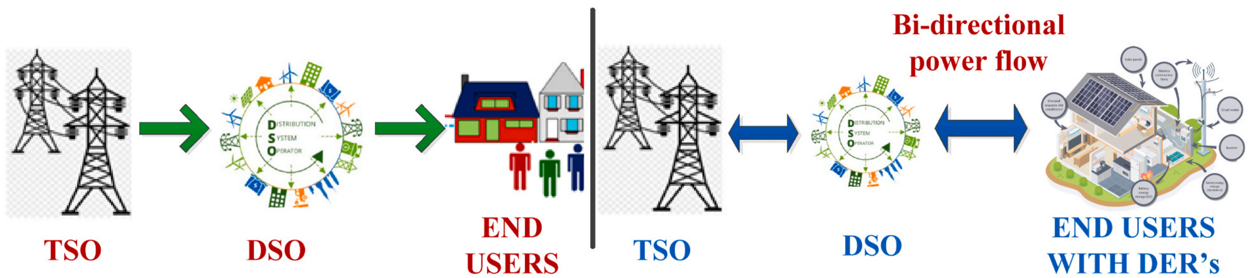

Fig. 5. Enhancing the coordination between TSO and DSO.

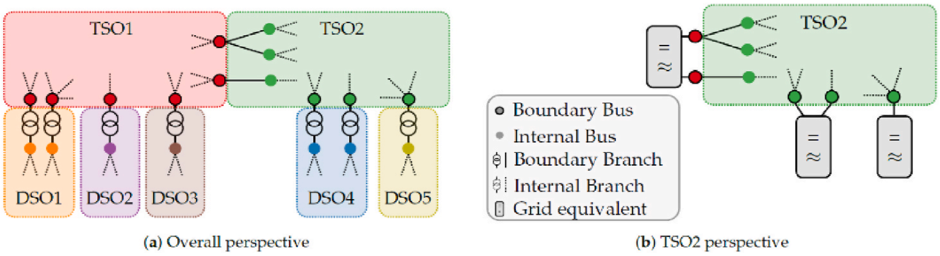

Fig. 6. Conceptual representation of TSO and DSO integration [12].

3. Distributed optimization methods coordinating multiple system operators

Within the existing body of research, several distributed optimization techniques have been put forth to strike a suitable balance between conserving sovereignty while sacrificing optimality in non-coordinated approaches and achieving centralized optimization [1]. Distributed optimization methods are employed to perform mathematical optimizations, specifically in the context of calculating optimum power flows (OPFs). As seen in Fig. 7, three distinct groups may be identified. Initially, these algorithms decompose the major optimization issue into smaller, more manageable portions. These algorithms then proceed to optimize and coordinate these individual problem components iteratively. Furthermore, using grid equivalents with iterative updates depicts the power systems of nearby System Operators (SOs) inside optimization processes. Thirdly, techniques that employ predetermined procedures without any form of repetition are referred to as sequential approaches.

Numerous mathematical decomposition approaches have been devised and continue to undergo further development. The Lagrangian relaxation (LR) approach has been suggested and empirically validated several times in electrical power systems [10]. In this context, Load Restoration (LR) is frequently utilized to optimize voltage management and reactive power dispatch, as evidenced by previous studies [4]. On the other hand, other decomposition strategies involve including variables at the system operator's boundaries in both parties' optimization problems. To properly handle these copy variables, it is necessary to incorporate and consider equality restrictions. The second group encompasses many methods, including the auxiliary problem principle (APP) [5] the predictor-corrector proximal multiplier technique (PCPM) [14] the alternating direction method of multipliers (ADMM) [15] and the augmented-Lagrangian-based alternating direction inexact Newton method (ALADIN) [6]. [16] provide comprehensive information on the applications and comparisons of electrical power systems.

4. DER forecasting techniques

Forecasting plays a crucial role in determining the accuracy and power of various methodologies, influencing the selection process. Determining the feasibility of engaging in a commercial venture may necessitate a preliminary assessment of the market's magnitude. Still, a forecast produced for budgetary considerations should exhibit a high degree of precision. Over the past 15 years, forecasting has experienced significant advancements in theoretical and practical aspects, as evidenced by the seminal review study authored by Ref. [17]. Therefore, this study encompasses many topics, from theoretical concepts to practical applications, making it both contemporary and comprehensive. The rapid progression of computing technology has facilitated the examination of increasingly extensive and intricate data collections, fostering a heightened fascination with analytics and data science. Consequently, the repertoire of forecasting methodologies that forecasters employ has expanded in quantity and complexity. Computer science has emerged as a pioneering field, driving advancements via techniques such as neural networks and various forms of machine learning [18]. These approaches have garnered significant attention and interest from experts in forecasting and decision-making domains. Additional techniques, such as Bayesian forecasting and complicated regression models, have also experienced improvements due to advancements in computation.

Furthermore, advancements have not been restricted solely to those derived from computational progress. Forecasting depends on the fundamental assumption that leveraging existing and historical information enables the generation of forecasts on future events. In

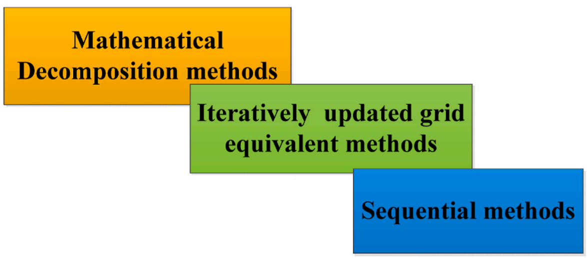

Fig. 7. Classification of grid operating approaches integrated multiple system operators through distributed optimization approaches.

the context of time series analysis, it is generally accepted that patterns observed in previous data may be effectively utilized to forecast future values. Nevertheless, it is not anticipated that the precise forecasting of future values will be achieved [19]. Among the several alternatives for predicting a future time series at a particular time, one might consider an anticipated value (referred to as a point forecast), a prediction interval, a percentile, or an entire prediction distribution. This collection of outcomes can be regarded as "the prediction" as a whole. There exist a multitude of alternative potential outcomes that may arise from a forecasting procedure. The primary aim might involve predicting an occurrence, such as equipment malfunction, and time series might have a limited impact on the forecasting procedure. Forecasting processes are most effective when they directly apply to a practical problem that must be resolved. A comprehensive study of the fundamental characteristics of the situation can further expand the theory. Consequently, the theoretical findings can contribute to enhanced practical applications. The assumption made in this introduction is that forecasting theories are formulated and established in the form of forecasting methodologies and models [20]. The generic representation of forecasting parameters integrated with different models is represented in Figs. 8 and 9 respectively. Accuracy and cost are two of the most critical metrics used to assess a simulation in the real world. Data availability, processing power, software, time to collect and evaluate, forecast horizon, and other aspects are additional considerations. For example, assuming some broad assumptions, Fig. 8 describes the rising costs and increasing accuracy of predicting against the cost of making errors in the forecast In terms of economic justification, the most advanced method is one that lies inside the zone where the total cost is least. There are many various ways to evaluate the correctness of a simulation, and each has its own set of advantages and disadvantages. While every situation is unique and has its own objectives, there are some broad guidelines for utilizing the accuracy metrics. In this context, a forecasting technique is described as a pre-established series of actions that generate forecasts for future time periods. Numerous forecasting approaches possess stochastic models that yield identical point forecasts, which does not apply to all methods. A stochastic model offers a mechanism for producing data that may be utilized to generate prediction intervals and complete prediction distributions alongside point predictions. Every stochastic model is built upon assumptions on the underlying process and its corresponding probability distributions. Even when a forecasting approach is based on a stochastic model, it is not always guaranteed that the model will be unique [21]. The simple exponential smoothing approach encompasses several stochastic models, such as state space models, which may exhibit either homoscedasticity (i.e., constant variance) or heteroscedasticity. Forecasts can be generated by a systematic procedure that integrates novel and pre-existing forecasting methodologies and models. Indeed, these more intricate procedures would also fall under forecasting methods/models.

4.1. Hyper parameter tuning significance

Hyperparameter tuning is a crucial aspect of optimizing both optimization algorithms and machine learning algorithms. In optimization algorithms, hyperparameters determine how the optimization process behaves, such as the learning rate in gradient descent or the population size in genetic algorithms. Tuning these hyperparameters can significantly impact the convergence speed and the quality of the solutions found. In machine learning algorithms, hyperparameters control the behavior and complexity of the model, such as the regularization parameter in linear regression or the depth of a decision tree. Tuning these hyperparameters is essential for improving the model's performance and generalization ability on unseen data. Both in optimization and machine learning, hyperparameter tuning involves searching for the optimal values of hyperparameters within a predefined space using techniques like grid search, random search, Bayesian optimization, or genetic algorithms.

Hyperparameter tuning methods in load forecasting involve optimizing the parameters of predictive models to improve their performance. Common techniques include grid search, random search, Bayesian optimization, and genetic algorithms. The overview of the following tuning methods are as follows:

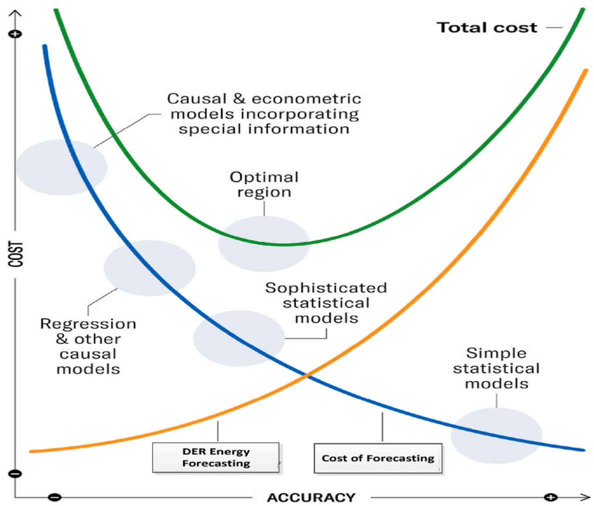

Fig. 8. Forecasting of DER energy and cost.

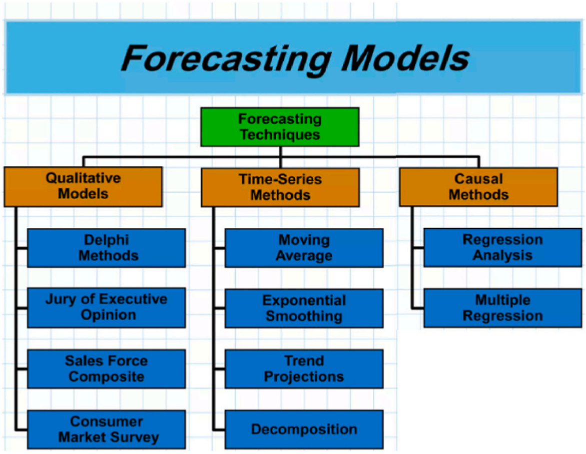

Fig. 9. Classification of Forecasting approaches.

Grid Search: This method involves evaluating model performance for different combinations of hyperparameters specified in a grid. It exhaustively searches through all possible hyperparameter combinations and selects the one that yields the best performance.

Random Search: Unlike grid search, random search samples hyperparameter combinations randomly from predefined ranges. While less computationally intensive than grid search, it can be equally effective in finding optimal hyperparameters and is particularly useful when the search space is large.

Bayesian Optimization: Bayesian optimization sequentially selects hyperparameters based on the model's performance using probabilistic models such as Gaussian processes. It efficiently explores the hyperparameter space by balancing exploration and exploitation, making it suitable for complex optimization problems.

Genetic Algorithms: Inspired by natural selection, genetic algorithms iteratively evolve a population of candidate solutions by applying selection, crossover, and mutation operations. They explore the hyperparameter space by mimicking the process of biological evolution and can handle discrete and nonlinear optimization problems effectively.

4.2. Traditional forecasting methods

Based on the evolution of time series forecasting methodologies, the prevailing prediction approaches can be categorized into three distinct groups: classical time series forecasting methods, machine learning and deep learning prediction methods, and hybrid time series forecasting methods. Traditional time series forecasting relies on mathematical and statistical modelling for future projections. Energy predictions can be categorized based on the time period, as given in Table 2 and seen in Fig. 10. Ultra-short-term electrical energy consumption forecasting (VSTCF) encompasses forecasts that predict energy consumption between 1 min and 1 h in advance. The review by Ref. [22] encompasses that very short-term energy forecasting is used for lead times ranging from a few seconds to many

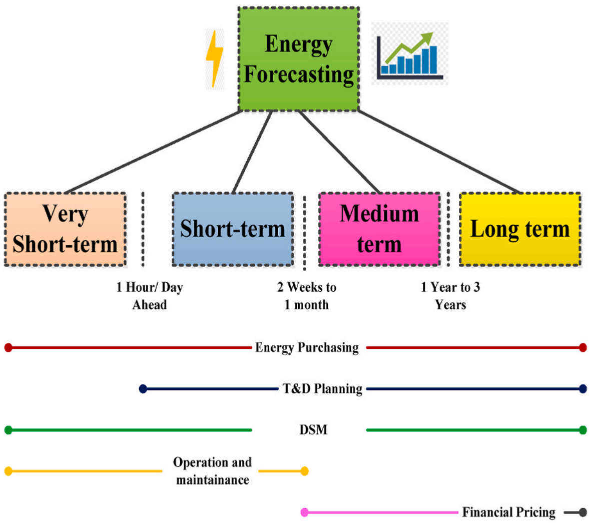

Fig. 10. Energy forecasting categorization based on time duration interval.

minutes. Instead of modelling correlations between load, time, weather, and other load influencing elements in VSTLF, we must focus on extrapolating the recently observed load pattern to the near future. There aren't many methods for anticipating very short-term load.

First and second order polynomial extrapolation, autoregressive and autoregressive moving average models, and artificial neural networks are among the techniques described. VSTLF is utilized in generation prediction, distribution schedules, contingency analysis for system security, and so on. On the other hand, short-term electrical energy consumption forecasting (STCF) pertains to predictions spanning 1 h to 2 weeks ahead. Electrical energy consumption forecasting may be categorized into two timeframes: medium-term (MTCF) and long-term (LTCF). MTCF pertains to projections made for a period ranging from 2 weeks to 3 years, while LTCF is conducted for forecasts spanning from 3 years up to 50 years. Traditional linear models primarily consist of the autoregressive (AR) model, moving average (MA) model, autoregressive moving average (ARMA) model, and autoregressive integrated moving average (ARIMA) model, as shown in Table 1.

Threshold autoregressive (TAR) models [2] constant correlated (CCC) models [2] and conditional heteroscedasticity (CHE) models [3] are examples of well-known nonlinear linear models. Moreover, numerous significant classical prediction methods are based on exponential smoothing, including Simple Exponential Smoothing (SES), Holt's linear trend approach, Holt-Winters' multiplicative method, Holt-Winters' additive method, and Holt-Winters' damped method. Currently, the ARMA model is extensively employed in time series forecasting across all disciplines, yielding significant outcomes [23]. employed the autoregressive moving average (ARMA) model to provide predictions for monthly and yearly fluctuations in the time series of stock returns for the Standard & Poor's 500 Index and the London Stock Exchange. The statistical examination of the S&P 500 index demonstrates that the Autoregressive Moving Average (ARMA) model has superior performance compared to the London Stock Exchange. Furthermore, the ARMA model exhibits the capability to accurately anticipate medium to long-term outcomes using actual historical data [24]. developed an adaptive autoregressive moving average (ARMA) model to anticipate short-term power load in power production systems. Empirical evidence demonstrates that the adaptive ARMA model outperforms the conventional ARMA model in terms of predicting accuracy for both 24-h and one-week time horizons [25]. developed a bootstrap algorithm to accurately predict the prediction interval of an ARMA (p, q) model with unknown order. This algorithm significantly improves coverage accuracy compared to methods that use pre-estimated order values.

The authors in Ref. [26] developed a closed-form equation for the finite predictor coefficients of multivariate ARMA processes. They used this expression to estimate the asymptotic behaviour of a sum in the autoregressive model fitting and sieve bootstrap. The results are novel even for univariate ARMA systems. WEISS C H proposes an extension of discrete autoregressive moving average (ARMA) models for modelling categorical time series, including nominal and ordinal ones. This method involves observation-driven regime switching, resulting in the RS-DARMA family of models. The RS-DAR(1) model is a simplified Markov chain with a simple data-generating mechanism and the ability to manage negative serial dependency. To analyze many variables at once, researchers have introduced the Granger causality [27] which builds on the existing univariate time series model. Vector autoregressive moving average (VARMA) is a multivariate promotion model that was developed as an extension of the ARMA model due to its ability to represent both the Vector autoregressive (VAR) and the Vector moving average (VMA) models. Data from a stationary time series is necessary for building the VARMA model. Time series data that is not stationary must be processed using first-order difference methods. However, the difference processing will lose the time series trend information. The Vector Error Correction Model (VECM) established by Ref. [28] is a solution to the aforementioned issues since it rigorously considers the co-integration connection between time series. The ARMA model has proven to be quite effective when predicting stationary time series. However, with real-time series data, there is hardly any static information. Because of this, the data's features severely restrict the model's usefulness and weak adaptability.

Many authors, for example [29], described how they felt about load forecasting being the key utility in network and maintenance

Table 1
Comparative analysis of model characteristics.

Model	Characteristic	PACF correlogram	ACF correlogram	Data characteristic
$AR_{(p)}$	1 y_t depends on its own past values 2. P is computed using PACF function	Spikes till pth The lag then cuts off to zero	Spikes then decay to zero	Data should be stationary in nature
$MA_{(q)}$	1 y_t depends on the error term which follows a white noise process 2. q is computed using ACF function	Spikes then decay to zero	Spikes till pth The lag then cuts off to zero	Data should be stationary in nature
$ARMA_{(p,\ q)}$	1. ARMA = AR + MA 2. Value of p and q are determined using AIC criteria	Spikes then decay to zero	Spikes then decay to zero	Data should be stationary in nature
$ARIMA_{(p,\ d,\ q)}$	1. Data is made stationary by differencing it. 2. Box-Jenkins approach is used to determine the model	Spikes then decay to zero	Spikes then decay to zero	Data should be non-stationery in nature
SARIMA(P, D,Q)	1 SARIMA transforms non-stationary data by differencing. 2 Box-Jenkins methodology selects SARIMA parameters for modeling time series dynamics.	Spikes that decay to zero,	Exhibits significant spikes at seasonal lags,	Data should ideally be made stationary through differencing before fitting a SARIMA model.

Table 2
Different types of load forecasting and their applications.

Type of Forecast	Lead Time	Application
Very short term	Few minutes	Generation, distribution schedules, and contingency analysis for system security.
Short term	Few hours	Allocating of spinning reserve, operational planning and unit commitment, and maintenance scheduling.
Midterm	Few days	Planning for seasonal peak-summer and winter.
Long term	Few months	Planning generation growth

planning for supply-side and demand-side. Future load needs are estimated for varied lead times (forecasting intervals) ranging from a few seconds to more than a year [30]. Long-term electric load forecasting is a critical component of effective and efficient planning. Overestimation of future load may necessitate investing more money to develop new power plants to meet this demand. Furthermore, an underestimate of the load may pose difficulties in feeding this load from available electric supplies, resulting in a deficiency in the system's spinning reserve, which may result in an insecure and unreliable system. The vast majority of load forecasting studies have bothered themselves with short-term ELF, although sociological, metrological, and financial factors influence time-varying system load, their effects are not frequently taken into account in short-term ELF. In literature, the purpose of load forecasting has been observed to be ISOs, electrical sales, manpower development and financing and grid formations. Classifications foreseen in the literature are not only in the area of short-term load forecasting but also in very short-term load forecasting, The four major categories of load forecasting techniques are described below. These categories are based on lead time [31].

4.2.1. Very short-term load forecasting (VSTLF)

The review by Ref. [22] encompasses that it is used for lead times ranging from a few seconds to many minutes. Instead of modelling correlations between load, time, weather, and other load-influencing elements in VSTLF, we must focus on extrapolating the recently observed load pattern to the near future. There aren't many methods for anticipating very short-term loads. First and second-order polynomial extrapolation, autoregressive and autoregressive moving average models, and artificial neural networks are among the techniques described. VSTLF is utilized in generation prediction, distribution schedules, contingency analysis for system security, and so on.

4.2.2. Short-term load forecasting (STLF)

Short-term load forecasting is of the utmost importance in commercial buildings [32], It is utilized for lead times ranging from a few minutes to several hours. Several parameters, such as time considerations, weather data, and potential client classifications, should be addressed for short-term load forecasting. Short-term load forecasting is critical in system operations because it serves as the primary source of information for all daily and weekly operations including generating commitment and scheduling. STLF is also critical for the power system's economic and reliable functioning. To obtain high forecasting accuracy and speed, it is necessary to understand the components that influence the load. These elements include the type and time of day, the meteorological conditions in the forecasting area, the season, and so on. Because most days have different load patterns, a day type is required. The time of day is a significant consideration in short-term load forecasting. It is necessary to know the predicted time of day because the degree of demand varies throughout the day. As a result, the correlations between these elements and load demand must be determined in order for projections to be feasibly precise [33].

4.2.3. Mid-term load forecasting (MTLF)

It is utilized for lead times ranging from a few days to several weeks. MTLF is used to forecast load during seasonal shifts such as peak-summer and winter.

4.2.4. Long-term load forecasting (LTLF)

According to some researchers [34], it is generally used for lead times ranging from a few months to a few years, it is used. The historical load and weather data, the number of customers in various categories, the local appliances, and their characteristics, including age, the economic and demographic data and their forecasts, the sales data for appliances, and other factors are all taken into account in the medium- and long-term forecasts. Generational growth planning is done using LTLF.

Applications of the different classes of load forecasting are listed in the table below.

However, some researchers have explored some challenges and drawbacks in electrical load forecasting as it is based on anticipated circumstances, such as the weather. Unfortunately, the weather can be unpredictable at times, and forecasts may be inaccurate when the actual weather is different from what was anticipated. The majority of seasoned utility forecasters employ manual techniques that rely on in-depth knowledge of a variety of contributing elements based on impending events or a specific dataset. The number and complexity of forecasts are growing, making it unsustainable to rely solely on manual forecasting.

4.3. Machine learning forecasting model of time series

The traditional time series forecasting models are effective in capturing linear connections within time series data and yield satisfactory results, particularly when the dataset is limited in size. However, its effectiveness is limited when used in large-scale complicated nonlinear time series. As a result, researchers place greater emphasis on the time series prediction techniques employed in machine learning or deep learning. Artificial neural networks, such as multi-layer perceptron (MLPs) networks [35] and radial basis function (RBF) networks [19] possess adaptive and self-organizing learning processes. The aforementioned models are the earliest neural network models employed for nonlinear time series forecasting across several disciplines, and they have demonstrated favorable outcomes. Furthermore, fuzzy theory, Gaussian process regression, decision trees, support vector machines, LSTM, and other similar techniques are commonly employed in the field of time series forecasting. These methods have demonstrated strong predictive capabilities, particularly when dealing with nonlinear time series data. Many of the researchers look into fuzzy time series forecasting centres on topics including the universe-splitting technique, the number of elements in a time series, the sequence of the fuzzy rules used, and the number of steps in the prediction process.

For instance [25], divide the universe of discourse through the idea of information granules to better react to the characteristics of the data [36]; used an automatic clustering algorithm to divide the universe of discourse; and [37] adjusts the fuzzy interval to improve model prediction accuracy. A high-order fuzzy time series forecasting model was presented by Ref. [38] as part of his research on model order [39]. enhanced the forecasting approach given by Ref. [40] for use in making long-term predictions. Due to the fact that each data granule slices the time series differently. The fuzzy C-means clustering algorithm is married to the dynamic temporal warping technique. Next, the fragments of varying lengths of the time series are grouped. Then, the concept of fuzzy time series prediction is implemented. Finally, a long-term prediction of the time series is constructed using the concept of fuzzy time series forecasting. Using a time series prediction model, [39]. Combining kernel mapping with an HFCM (high-order fuzzy cognitive map) creates a kernel-HFCM. The original one-dimensional time series was transformed into a multi-dimensional feature time series from which the most important characteristics were extracted.

Finally, the characteristic time series is mapped back to the expected one-dimensional time series using inverse kernel mapping. Experiments with 7 reference data sets show that this strategy effectively forecasts time series. The model presented by Ref. [38] combines fuzzy linear regression and GARCH to estimate parameters, and it is denoted by the acronym FLR-FSWGARCH. This model makes it easier to estimate reliably and forecast parameters. As a result, the model's ability to make accurate predictions is enhanced. The model's high accuracy and dependability for time series forecasting are demonstrated by trials conducted on two data sets. As a result of its distinctive features and the large range of practical applications, fuzzy time series forecasting has emerged as a prominent area of study. The study of fuzzy time series forecasting methods has significant academic significance and application value in the field of time series forecasting since it improves the accuracy and interpretability of the model.

4.3.1. How the AI tools used for forecasting

Artificial intelligence is employed in demand forecasting as a response to fluctuations in demand. Demand forecasting has conventionally been classified as a type of predictive analytics, in which historical data is utilized to analyze the process of estimating customer demand [41]. By employing AI, businesses can utilize Machine Learning algorithms to ascertain fluctuations in consumer demand with the utmost precision. Algorithms possess the capability to autonomously discern patterns, decipher intricate connections within extensive datasets, and capture indicators of demand volatility. Machine learning advances the field of demand forecasting by facilitating improved predictions through the utilization of real-time data derived from internal and external sources, including demographics, weather conditions, online evaluations, and social media [41]. Demand side management networks that leverage contemporary machine learning algorithms and external data can surpass those that are predominantly managed manually by data analysts, while also effectively responding to external fluctuations.

A substantial paradigm shift has occurred in this field since the inception of artificial intelligence (AI) and machine learning. These technologies enable a more detailed, accurate, and instantaneous evaluation of market conditions. AI in demand forecasting has the potential to drastically reduce errors in electricity supply chain networks by 50 %, according to Ref. [42]. Fundamental to demand forecasting powered by AI is the capacity to observe and analyze each disruption within the supply continuity. This entails monitoring the system of products through optimization of algorithms, discerning fluctuations and voltage dips at customer end, forecasting quality deficiencies, and identifying underperforming generating units that incur superfluous outages [43]. Furthermore, the adaptability of the system enables it to incorporate factors such as seasonality or particular periods of the year, thereby providing a dynamic forecasting model that adjusts in response to shifting market conditions.

These advantages are amplified not only by AI but also by integrating machine learning into demand side management. Conventional approaches such as ARIMA (Auto Regressive Integrated Moving Average) and exponential smoothing methods, which depend exclusively on past data, are overshadowed by the dynamic, real-time insights provided by machine learning. These algorithms engage in active learning by adapting and developing with each new dataset, rather than simply evaluating it [43]. A notable advantage of machine learning in demand forecasting is its ability to tackle challenges posed by new Distributed Energy Resources. In the dynamic world of industry, predicting future electricity sales, particularly amidst unforeseen circumstances, becomes paramount. While traditional forecasting methods may falter in the face of sudden shifts, AI offers a robust recalibration, introducing many avant-garde solutions [44]. For instance, AI aids in creating models centered on evolving market behaviors and integrating external factors like news events or economic indicators for real-time demand forecasting.

4.3.2. Impacts of AI forecasting on the DSM network planning

Integrating AI into demand forecasting in the energy and utilities sectors holds tremendous potential. Energy consumption patterns can be unpredictable, driven by seasonal fluctuations, industrial production cycles, and consumer behaviors [45]. Traditional methods of estimating demand in this area frequently rely on historical data, which may not properly capture rapid or unexpected changes. Conversely, demand predictions generated by AI-powered models can be more precise due to their ability to process immense and diverse datasets in real time, such as those pertaining to infrastructure developments, weather patterns, socioeconomic indicators, and even global events.

This accuracy guarantees that power generation and consumption are synchronized for the benefit of utilities, thereby minimizing waste and decreasing expenses. Additionally, it facilitates effective grid administration by permitting proactive modifications in the allocation of energy. By utilizing weather forecasts to forecast energy production in renewable energy sectors like solar and wind, AI can ensure that excess energy is efficiently stored or distributed [45]. AI in demand forecasting in the energy and utilities sector leads to operational efficiency, cost savings, and a more sustainable energy future.

The Impact of AI in DSM may be noted but not limited to as follows.

- AI methods have been widely employed to predict future power consumption
- Significantly improving power dispatch,
- Load scheduling, And market management

In addition to being used to detect electrical load anomalies, data-driven classification methods have been implemented in an effort to ensure the safe operation of power systems [46]. Anomaly detection research encompasses a broad spectrum of applications, including smart buildings, large-scale industrial power system monitoring, and even generic residential dwellings. On the basis of the detection activities, anomaly detection can also be classified as non-invasive health status surveillance for residents and system security detection [47]. Prior research on demand response has put forth strategies for effectively managing resources and delivering feedback to the energy market [48].

In light of the significant importance attributed to the consumption side in power systems, a multitude of recent evaluations have been undertaken to succinctly outline the implementations of AI-driven approaches in demand-side management. As an illustration, a review article on the implementation of load forecasting was published by Ref. [49]. The article examined various artificial neural network (ANN)-based load forecasting approaches, the influence of adjacent environments on predictive outcomes, and the classification of load forecasting. It identified critical parameters, including input combination, activation functions, forecast model architecture, and the training algorithm of the network, that influence the accuracy of ANN-based forecast models. The review underscored the potential of artificial intelligence (AI) methods to achieve smart grid and building objectives through precise load forecasting. It offered valuable insights into the criticality of precise load forecasting in ensuring effective energy management and improved power system planning.

A study conducted by Khan et al. [44] examined capacity forecasting, dynamic pricing, and demand-side management (DSM) in relation to the subject matter. The review examined the significance of load forecasting in power system planning and operation, as well as the manner in which dynamic pricing-based techniques and load forecasting will be integrated into future smart grids to ensure efficient DSM. In addition, a comparative analysis of forecasting methods and a discussion of the difficulties associated with load forecasting and DSM were included. These challenges were addressed through the implementation of dynamic pricing systems, appliance scheduling, and energy consumption optimization as methods to manage energy on the consumer side.

In DSM applications, AI load forecasting (LF) and dynamic pricing-based schemes offer promising solutions. DSM based on AI load forecasting has the capability to efficiently oversee the energy needs of both consumers and utilities. It facilitates the utility in reducing the supply-and-demand imbalance. Although not a novel idea, the tempo of implementation is extremely slow due to a number of obstacles.

5. Factors affecting load forecasting

A number of factors may influence the study of electrical load forecasting, according to Ref. [31] climate change, economic and demographic factors have the largest influence on the evolution of electricity consumption. The factors influencing load vary depending on the consumption unit. The level of production often determines the industrial load. The load is frequently relatively stable, and its dependence on different output levels can be estimated. However, from the perspective of the utility selling electricity, industrial units typically introduce uncertainty into estimates as their input factors vary [50]. The authors presented various challenges faced with precise anticipation of the forecast since the trends differ every time. In their paper [51] revealed that the varying input factors can be specified as follows: (a) time; (b) weather; (c) historical data; (d) economy and (e) random disturbances. It therefore appears that every predictive model's efficiency is thought to be partially dependent on the independent (input) variable to some extent. Recent studies on load forecasting highlight the significance of incorporating weather-related variables, such as temperature, humidity, and solar radiation, especially with the growing impact of climate change on energy consumption patterns. Additionally, advancements in smart grid technologies, including the proliferation of smart meters and IoT devices, offer valuable data sources for improving load forecasting models through enhanced granularity and real-time monitoring capabilities. Moreover, research emphasizes the integration of machine learning algorithms and artificial intelligence techniques to leverage big data analytics and optimize forecasting accuracy by capturing complex nonlinear relationships inherent in energy demand dynamics [51–54].

5.1. *Time*

This component influences electric load at various times of the day, including weekdays and weekends, holidays, and the seasons of the year. Seasonal impacts are responsible for long-term shifts in weather patterns [52]. Hot summer days frequently produce substantial cooling loads, which are more likely to occur in the afternoon and early evening. Cold winter evenings frequently result in high heating demands, which are more likely to occur late at night and early in the morning. Seasonal effects might include not only weather patterns but also popular vacation dates and shifts in Daylight Saving Time (DST) [53]. When DST is in effect, the electrical load profile is 1 h behind the standard time profile. In this case, the time-dependent electric load variation can reflect people's lifestyles, such as work schedules, leisure time, sleeping habits, and so on.

5.2. Weather

The most important extreme is this one. It includes factors like temperature, precipitation, wind speed, humidity, etc. The fluctuations in these variables directly cause the usage patterns of appliances like heaters, air conditioners, coolers, and so forth to change. A closer look at the review [54] shows that weather conditions substantially impact a power system's short-term electric load profile. Weather-sensitive loads, such as heating, ventilation, and air-conditioning (HVAC) equipment, will significantly influence smaller industrial/institutional power systems because these are the system's most enormous loads. HVAC equipment cycling on and off might cause electrical load profiles with seemingly random power swings. As the power system load increases, load diversity increases, the effect of load cycling is reduced, and the electric load profile becomes smoother.

5.3. Historical data

The study in Ref. [55] prevails that where a customer is concerned, these criteria include the kind of consumption (commercial, residential, agricultural, or industrial), the size of the buildings, the number of employees, and the amount of electricity consumed.

5.4. Economy

This element is significant in the deregulated market because it reflects the fluctuating electricity price. In contrast, load management policy significantly impacts the electric load growth or decline trend. Economic variables include investments in the facility's infrastructure, such as new buildings, labs, and experiments that add load to the electric grid. The site's funding profiles govern how and when equipment, processes, and experiments can be used. During system peaking, utility programs such as demand charges and demand control plans impact customers' electrical usage patterns [52]. This author further concedes that Economic considerations will not affect the STLF because they typically modify consumption patterns over a more extended period than 24 h.

5.5. Random disturbances

The shutdown or restart of massive loads, such as steel mills or wind tunnels, will result in load curve impulses. Other anomalous events that are previously recognized but have an unclear effect on the load are also classified as random disturbances. These disturbances can include substantial loads that do not follow a defined operating schedule, making prediction more complex [56].

Table 3 above offers studies reviewed for parameters used in electricity load forecasting in this literature. According to Ref. [57] the aforementioned factors can be used as inputs to the Kalman filter (KF). Because dealing with complex inputs (c., d., e.) is exceedingly challenging, the weather and time are the only variables incorporated in the KF. Typically, the KF is viewed through the lens of a discrete-time linear dynamic system.

A large proportion (50 %) of electricity demand forecasting was found to be reliant on weather characteristics, followed by the historical electricity use trend. The findings revealed that in the forecasting of electricity demand, little consideration is paid to household lifestyle. However [23], argues that at a distribution level, depending on the family's lifestyle, different residential (domestic) energy usage occurs. Thus family lifestyle cannot be discounted in the electrical load forecasting. These characteristics include life stage, family composition, housing type, age, home appliances owned and used, family income, cultural background, social life, and lifestyle patterns such as how long to stay at home and how to spend holidays, according to this research. The findings pave the way for future research on the relationship between ELF and household living [58].

6. Mathematical modelling of some commonly optimization methods

In literature, many mathematical formulations have been established for LF, but in this paper, we only looked at a few of them. In load demand forecasting, regression-based approaches have been widely employed [59]. These methods, which span from simple linear regression models to multivariate linear regression models, operate very well when the relationship between the dependent variable and the predictor factors is linear. They are typically fast, dependable, and simple to execute, with generally robust solutions.

Table 3
Variables used in electrical load forecasting.

Input parameter	No. of studies	Percentage (%)
Economic and weather	35	50
Household lifestyle	9	12.8
Historical data	23	32.9
Random disturbances	3	4.2

6.1. Multiple linear regression (MLR)

This method uses weather and non-weather variables such as temperature, humidity, day type, and customer class to predict the load at a given moment. The model can be expressed as in equation (1):

$$\mathrm{z(t)=a_0+a_1x_1(t)+\cdots+a_nx_n\ (t)+a(t)}, \tag{1}$$

where $z(t)$ is the electric load, $a(t)$ is a white noise component with zero mean and constant variance, $x_i\ (t)$ are the explanatory variables, and a_0 and a_i are regression coefficients, with $i = 1, \ldots, n$.

In general, we use correlation analysis to select this model's most common explanatory variables. The least squares technique is used to estimate the regression parameters. MLR is simple to create and maintain, but its typical sensitivity.

As stated by Ref. [60], serial correlation of weather variables disturbances might be a big issue.

6.2. Fuzzy model

According to Rouse [61] fuzzy logic is a computer technique based on "degrees of truth" rather than the well-known Boolean logic "true or false" (1 or 0). Fuzzy logic is a generalization of the Boolean logic that current computers are founded on. Fuzzy logic models were defined by Ref. [60] as a function that connects a set of input variables to a set of output variables; the input variable values do not have to be numerical. They only require transcription in a native language. For example, a weather parameter such as temperature may have "fuzzy" instances such as "low," "medium", and "high." The normalization (N) and summation ($\sum$) achieve the fuzzy mean operator. The most commonly used form of the Gaussian membership function is given in (2)

$$y=\sum\nolimits_{i=1}^{N} yi(z)\left(a_i^Tz+c_i\right) \tag{2}$$

6.3. Autoregressive model (AR)

AR model is a stochastic time-series regression model claimed to be very popular in literature [62]. Using this model, the output can be expressed as a linear combination of past outputs/measured values, as seen in equation (3).

$$A_q y_t = e_t \tag{3}$$

where qN is the shift operator i.e., $q^{\pm N} = I_{net}(t\pm N)$, $A(q)$ represents the combination of past values of output using coefficients, a_1; $a2\ldots a_n$ and e_t represents the part of the model that is omitted for more details on the models refer to Ref. [63].

6.4. Support Vector Regression (SVR)

The support Vector Regression approach is based on supervised machine learning [62]. It has been widely used in predicting finance, load, stock and other time series markets and so on. Given the training data $\{(u_1 y_1)\ldots(u_1 y_1)\}\subset U\times R$, in equation (4), U denotes the input pattern space. The purpose is to locate $f(u)$ function with the greatest departure from the actually obtained yi targets.

$$f(u)=\langle w,u\rangle+b \text{ with } w\in U, b\in R \tag{4}$$

The optimization problem is resolved using support Vector regression in equation (5):

$$\begin{aligned} &min_x\ \frac{1}{2}w^Tw+c\sum_{0<i<m}\left(\xi_i+\xi_i^*\right) \text{ given that } y(i)-\left(w^T\phi u_i+b\right) \\ &\quad\leq \varepsilon+\xi^*, \left(w^T\phi u_i+b\right)-y(ti)\leq\varepsilon+\xi^*, \xi_i, \xi_i^*\geq 0, i=1 \end{aligned} \tag{5}$$

where ϕu_i maps u_i to a higher dimensional space using a kernel

function and the training errors are subject to ε -insensitive tube

$y(i)-\left(w^T\phi u_i+b\right)\leq\varepsilon$ Cost of the error, C, width of the tube and the mapping function ϕ controls the regression quality.

6.5. Fuzzy systems

According to Rouse [61] fuzzy logic is a computer technique based on "degrees of truth" rather than the well-known Boolean logic "true or false" (1 or 0). Fuzzy logic is a generalization of the Boolean logic that current computers are founded on. Fuzzy logic models were defined by Ref. [60] as a function that connects a set of input variables to a set of output variables; the input variable values do not have to be numerical. They only require transcription in a native language. For example, a weather parameter such as temperature may have "fuzzy" instances such as "low," "medium", and "high." The normalization (N) and summation ($\sum$) achieve the fuzzy mean operator. The most commonly used form of the Gaussian membership function is given in equation (6).

$$y=\sum\nolimits_{i=1}^{N} yi(z)\left(a_i^Tz+c_i\right) \tag{6}$$

6.6. Adaptive neuro Fuzy systems

Adaptive Neuro-Fuzzy Inference System (ANFIS) was proposed by Jang Roger in 1993, which overcame the problems of neural networks and fuzzy logic by combining the self-learning ability of ANN and the logical inference ability and also robustness and ease in implementing the rules bases of fuzzy systems.

Modern neuro fuzzy systems are often represented as multilayer feed forward neural network. A neuro fuzzy system is a fuzzy system that is trained by a learning algorithm (usually) derived from neural network theory. The (heuristic) learning procedure operates on local information and causes only local modifications in the underlying fuzzy system. The learning process is not knowledge-based, but data driven. Besides, a neural-fuzzy system can always be interpreted as a system of fuzzy rules. It is possible both to create the system out of training data from scratch, and to initialize it from prior knowledge in the form of fuzzy-rules.

To present the ANFIS architecture, two fuzzy if-then rules based on a first order Sugeno model are considered.

- Rule 1: If (x is A1) and (y is B1) then (f = p1x + q1y + r1)
- Rule 2: If (x is A2) and (y is B2) then (f2 = p2x + q2y + r2)

where x and y are the inputs, Aj and Bj are the fuzzy sets, are the outputs within the fuzzy region specified by the fuzzy rule, p, q and r are the design parameters that are determined during the training process.

6.7. Artificial neural networks

Despite the advantages in comparison to classic methods, there are some drawbacks in methods based on the artificial neural networks which may affect the training process and as a result, the accuracy of the prediction may decrease. Some of these disadvantages are listed below.

- The ANNs have a black-box nature, i.e. it is not possible to completely control the way the neurons are connected in the network.
- There are no specific rules or guidelines for an optimal network design for different applications.
- The ANN approach alone is normally not enough for the modelling problems, especially in physical modelling.
- The accuracy of the results depend largely on the size of the training data.

The training of the network in some specific problems may be very hard or even impossible. The researchers have tried to overcome these disadvantages by introducing new methods or combining the ANN-based methods with other approach.

In equation (7), the fundamental topology of ANN is represented by a feed forward MLP neural network which is constituted by three types of neuron layers, namely input layer, hidden layer, and output layer:

$$\widehat{y} = b_O + \sum_{h=1}^{H} b_h g\left(\delta_{Oi} + \sum_{i=1}^{I} \delta_{hi} p_i\right) \tag{7}$$

Where;

I represents the number inputs p_i.

H is the number of hidden nodes in the network

The weights $\omega = (b, \delta)$, where $[b = b_1, \dots b_H]$ and $[\delta = \delta_1, \dots \delta_H]$, are for the hidden and output layer sequentially.

b_0 and δ_{0i} are the biases of each node, and the transfer function $g(\bullet)$ may be nonlinear and is usually either the sigmoid logistic or the hyperbolic tangent function.

6.8. Radial basis function

Recently, RBF networks have been considered as a promising alternative to MLP neural networks owing to their broad spectrum of applications and quicker learning ability. Merely one hidden layer is in existence and neurons of the hidden layer contain radial basis activation functions. RBF network is tantamount to the weighted summation of the responses of the hidden neurons and can be explained according to equation (8):

$$y_j = + \sum_{i=1}^{n} w_{ij} \phi_i(\| x - c_i \|) + b_{oj} \quad (j = 1, 2, \dots, n) \tag{8}$$

where the number of nodes in the hidden layer is n, input vector is x, the centre of the i^{th} hidden node is c_i, the weight of i^{th} node of the hidden layer is w_{ij} , the radial basis function with c_i being its centre is ϕ_i", and the bias of the j node of output layer is b_{oj}. The mapping from the input layer to the hidden layer is nonlinear, while it is linear from the hidden layer to the output layer. Herein, the radial basis function is a Gaussian function and input-to-centre distance is designated by utilizing simple Euclidean distance as indicated in equation (9) below:

$$\phi_i\ (x) = \exp\left(\frac{\|\ x - c_i\ \|^2}{\sigma^2}\right) \tag{9}$$

where $\sigma < \ 0$ is a predefined spread value of the function. It should be noted that three key parameters are decided in RBF networks which are spread, centres, and inter-network weights.

6.9. PSO- particle swarm optimization

Particle swarm optimization (PSO) [64,65]is a technique that was initially introduced by Kenndy and Eberhart in 1995 [38]. It operates on the principles of swarm intelligence. The PSO algorithm is widely employed in various domains such as machine learning, signal processing, adaptive control, and more, owing to its straightforward implementation. The PSO algorithm commences by arbitrarily initializing m particles, where each particle represents a prospective solution to the problem residing in the search space.

In this approach, the whole population is considered as a unique neighbourhood for that particle during the gaining experience process. To optimize the best search procedure, the best particle needs to share its coordinates information with other particles. In this algorithm, the i^{th} particle velocity, v_i^{k+1} , is updated according to equation (10):

$$v_i^{k+1} = wv_i^k + c_1\beta_1\left(p\phi_i - x_i^k\right) + c_2\ \beta_2\left(g\phi_i - x_i^k\right. \tag{10}$$

Where;.

β_1 and β_2 are two random numbers between zero and one;

x_i^k and v_i^k are the position and velocity of the i^{th} particle in k^{th} iteration, respectively;

$g\phi_i$ is the best position experienced by the whole particles

$g\phi_i$ represents the best personal experience of the particle

w is the inertia constant

c_1 and c_2 are constants called personal learning factor and social learning factor, respectively

To update the particle position is done according to equation (11):

$$x_i^{k+1} = x_i^k + v_i^k \tag{11}$$

Where;

x_i^{k+1} is the new particle position

x_i^k is the previous particle position

v_i^k is the new particle velocity obtained according to (10)

The velocity and position of each particle are updated according to (10) and (11) until all particles move. then the next iteration occurs, and this procedure continues until finding the best solution.

The old types of PSO algorithm had some undesirable dynamic characteristics including velocity restrictions to control the particle path. By applying the limitation on factors, the possibility of the dynamic characteristic control on the particle swarms and making a balance between local search and global search. According to these factors PSO algorithm is considered as in equation (12):

$$p = \rho_1 + \rho_{\ 2} > \ 4, \tag{12}$$

$$\theta = \frac{2\alpha}{p - 2 + \sqrt{p^2 - 4p}}$$

$$w = \theta,$$

$$c_1 = \rho_1\theta$$

$$c_2 = \rho_2\theta \tag{6}$$

Where;

ρ_1 and $\rho_{\ 2}$ are positive random numbers gained from a unified distribution, where their sum, $\rho,$ should be more than 4.

α has a value between zero and one

θ is known as restriction factor.

The PSO algorithm is known to outperform backpropagation training in terms of speed and is independent of the bias value however the optimization technique also poses some limiting constraints.

Advantages of the PSO algorithm.

- PSO algorithm is a derivative-free algorithm.
- It is easy to implementation, so it can be applied both in scientific research and engineering problems.
- It has a limited number of parameters and the impact of parameters to the solutions is small compared to other optimization techniques.

- The calculation in PSO algorithm is very simple.
- There are some techniques which ensure convergence, and the optimum value of the problem calculates easily within a short time.
- PSO is less dependent of a set of initial points than other optimization techniques.
- It is conceptually very simple.

Disadvantages of the PSO algorithm.

- PSO algorithm suffers from the partial optimism, which degrades the regulation of its speed and direction.
- Problems with non-coordinate system (for instance, in the energy field) exit.

6.10. Bayesian Optimization (BO)

The concept of Bayesian optimization (BO) was initially introduced by Pelikan et al. from the University of Illinois Urbana-Champaign (UIUC) in 1998 [66]. The algorithm determines the most favorable value of the function by calculating the probability of the output of the black box function based on the available sample points. The BO algorithm is highly efficient and particularly beneficial in situations when the evaluation cost of the objective function is high, the derivative of the independent variable is unattainable, or there are several peaks. The Bayesian optimization (BO) method consists of two fundamental elements: a probabilistic surrogate model that is constructed using prior distributions, and an acquisition function. BO is a model that operates in a sequential manner, where the posterior probability is updated using new sample points in each iteration. Simultaneously, to prevent being stuck in a suboptimal solution, Bayesian Optimization (BO) algorithms typically introduce a degree of randomness to strike a balance between random exploration and the posterior distribution. BO is a hyper-parameter estimation approach that has a strong convergence theory, as stated in reference [67].

By focusing on the optimization of hyper-parameters of models using Bayesian Optimization (BO) in equations (13) and (14), the purpose is to discover the global maximum or minimum value of a given objective function.

$$x^* = argmax\, f(x) \tag{13}$$

when $x \in X$, where

X represents the space of hyper-parameters. The objective of this article is to determine the highest possible value of the objective function.

$$\text{Let } D1 : t = (xi, yi), i = 1, 2, \cdots, \tag{14}$$

t represent the current dataset,Where yi is the dependent variable.

The model's generalization accuracy under the hyper-parameter xi is being evaluated.

The expression $D1$:t=(xi,yi),$i = 1,2,\cdots,t$ was denoted as D for simplicity. Our goal is to determine the highest possible value of the objective function within a set number of iterations. If y is considered a random observation of the generalization accuracy, it may be expressed as $y = f(x) + \varepsilon,$ where $\boldsymbol{\varepsilon}$ represents the noise that follows a normal distribution (0,2$\boldsymbol{\varepsilon}$), and is independent and identically distributed. The objective of hyper-parameter estimation is to identify the optimal value, denoted as x^*, inside the d-dimensional hyper-parameters space.

An issue with this maximum expected accuracy approach is that the actual sequential accuracy is sometimes computationally infeasible to calculate.

7. Load forecasting model strategies In SA

The ability to accurately estimate daily peak load demand is crucial for energy industry decision-makers. This aids in the creation of reliable supply schedules that are consistent throughout peak times. Effective load shifting across transmission substations, scheduling of peak station start-up times, load ow analysis, and power system security studies are all made possible by accurate short-term load projections. Without an accurate forecast, utilities can expect to incur losses in revenue.

The primary role player in South Africa's electricity is Eskom, which supplies approximately 90 % of the country's electricity demand, thus, most load forecasting studies are implemented at this utility through different stages: Gx, Tx, and Dx. However, at the distribution level, a combination of external data from municipalities and IPPs is often integrated for load forecasting.

7.1. Existing optimization methods

Load forecasting takes several forms in Eskom's Transmission and Distribution settings. A statistical method is the most commonly utilized method in Transmission for load forecasting. However, research on a neural network-based forecaster for use in the Transmission operating environment is underway [68]. Furthermore, the South African Power Systems Studies Institute, a division of Technology Services International, is in charge of a research project investigating the feasibility of utilizing a neural network-based system for long-term load forecasting. This project's results will improve decision-making in transmission planning and growth.

On the other hand, distribution currently has a Geographical-based Load Forecasting tool (GLF) within the Smallworld software environment with a forecast horizon of less than 2 years. This tool is not sufficiently utilized due to several reasons. Before the utilization of this tool, the company previously used the ARIMA model for forecasting hourly loads with the help of historical data gathered during public holidays and peak hours. All this data would then be captured on a traditional Microsoft spreadsheet, in 2007, seeking to improve its load foresting Eskom distribution adopted the above-mentioned spatial forecasting based on GLF.

To address the accuracy issue in this field and help energy management decision-makers in the country operate their everyday operations efficiently and affordably, some researchers [60] developed an STLF employing Artificial Neural Networks (ANNs) to generate extremely accurate outputs. To generate precise hourly load projections, the developed model is based on the Multilayer Perceptron (MLP), a class of feedforward ANNs that uses the backpropagation (BP) method as its training mechanism [59]. presented a Multivariate Adaptive Regression Splines (MARS) modeling approach toward daily peak electricity load forecasting in South Africa. MARS is a non-parametric multivariate regression method that generates results explained to management in high-dimensional issues with complex model structures, such as nonlinearities, interactions, and missing data.

The research institute CSIR [69] developed a methodology for forecasting annual national electricity demand in collaboration with BHP Billiton in 2003 using the Sectoral Regression Model as inputs into the Eskom Integrated Resource Plan (2017-205). This methodology was successively reused several times, including to provide forecasts used in the previous IRP and its revision. A recent study by Refs. [29,61] proposed Improving the Load Forecasting Process for a Power Distribution Network Using Hybrid AI and Deep Learning Algorithms. The study was realized using two real South African distribution redistributor customers' power consumption data, with which the techniques were found to reach superior levels of accuracy without including temperature in the formulation of their models.

7.2. *Prospects to shift the market paradigm from DS to DSO*

The distribution network planning, design, and operation approach has shifted substantially due to the proliferation of distributed energy resources (DERs), load growth, energy storage technology developments, and higher customer expectations. Thus, any forecast must be done within the proper context of its purpose. The Transmission Demand Forecast, in the context of South Africa and its energy needs [70], serves the purpose of providing the load capacity required to stimulate economic growth and to enable the planning of adequate Grid capacity for the transfer of power from generation supply to points of delivery at the customer.

South Africa is undergoing a paradigm shift, and it is critical to recognize the changes and alter forecasting methodologies going forward. Furthermore, to complicate accurate forecasting of the energy system, the shift in generation mix presenting itself globally provides additional challenges to forecasting approaches. As a result, the generation mix feeding both present and projected demand will change substantially as the country is experiencing a movement towards reduction of carbon footprint [70]. However, due to the comparatively low cost of coal in South Africa, it is still envisaged to remain a vital component of the country's baseload generation. Still, including Independent Power Producers and enhancing technology on energy sources used locally at client premises produces a bi-directional power flow, which contributes to a combination of vertically and horizontally integrated systems with new-generation sources combined to feed into the primary grid. Fig. 11 below shows the paradigm shift SA electricity utility is experiencing.

With the paradigm shift in mind, continuously revising the energy markets at the distribution level is necessary. This then changes the concentration from the Distributed Systems being fed from Transmission only to Distribution System Operators running from different supply injections. However, this shift also affects load forecasting at Tx level as they need to adopt new methods to cater to the production mix and acquire accurate results.

For DSO, presently used traditional methods will need to be integrated with new fast-tracking machine learning methods as the DERs are intermittent. Subsequently, their patterns vary daily. Assumptions on the grid connectivity must be considered in each forecasting scenario. The grid connectivity relates to the customer base being grid-connected, grid-tied (backup supply mixed with other generation sources), and off-grid solutions. The outcome of these scenarios will depend on the consumer's development and adoption rate for alternative electricity sources. Some of the assumptions, but not limited, are listed below.

- Eskom remains a going concern as the current utility for South Africa
- The socio-political environment is assumed to stabilize or improve in the medium to long term.
- Policy certainty and cost-reflective tariffs allow Eskom to function in a competitive energy market.
- Energy-efficient and Technologically advanced electricity realm is assumed.

With all these concerns assumed, DER forecasters must think ahead critically when planning forecasting methods, as there is high

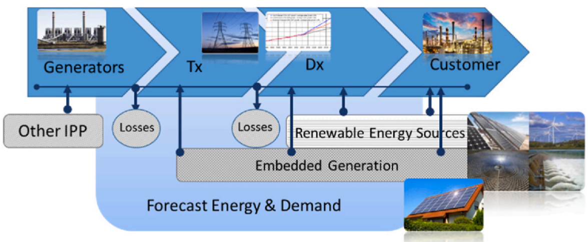

Fig. 11. Supply and demand of the future [70].

uncertainty with new developments and technological advancements. As a result, values might need to be adjusted in years to come, with insufficient data available for long-term forecasting dependability.

7.3. Impact of new technologies on load forecasting (South Africa)

The implications of emerging technologies on load forecasting in South Africa might be multifaceted, encompassing the influence of technological innovation, economic growth, response to disruptions, extreme weather events, renewable energy, electric vehicles, energy storage systems, unexpected changes in DER output and other DERs. These transformative technologies introduce a dynamic element to South Africa's power system operation, significantly influencing load forecasting accuracy [71]. Since DERs depend on weather conditions, coordinating the operation of the entire power system will become a new complex challenge. This challenge must be addressed to supply electric energy to customers at the appropriate quality level. Therefore, advanced forecasting methods are crucial for DOS to accurately forecast and accommodate the increasing penetration of DERs in South Africa's power grid. The impact of new technologies to be considered in load forecasting models are pointed out below.

- The adoption of electric vehicles (EVs) can lead to changes in electricity demand patterns, as charging infrastructure and EV penetration increase. Load forecasting models need to consider the charging behavior of EVs and their impact on the grid.
- Integration of renewable energy sources, like solar and wind, introduces variability and uncertainty in electricity generation, requiring adjustments in load forecasting models to account for their intermittent nature.
- wheeling scheme: Eskom has run a successful pilot of virtual wheeling, which will allow companies with multiple offtake sites to connect to generators using the Eskom or municipal grids.
- The development and adoption of energy storage technologies, such as batteries, can influence load forecasting by enabling the integration of intermittent renewable energy sources and providing grid stability.
- Smart grid technologies, including advanced metering infrastructure and demand response programs, enable more accurate load forecasting by providing real-time data on electricity consumption and allowing for demand-side management.

It is crucial to anticipate South Africa's load prediction with robust forecasting methodologies that account for network constraints, especially given the country's transition towards DER integrations. Many studies have addressed the role of renewable energy, economic growth, and technological innovation in predicting the load capacity factor in South Africa [29,71,72]. Additionally, another study emphasizes the nonlinear electric load and increased forecasting errors due to the unavailability of real-time metering technology for behind-the-meter solar PV generation. This study directed the need for advancements in this area [71].

The references mentioned above focus on the impact of DERs on traditional load forecasting methods. They underline the need for advanced forecasting methods, such as hybrid AI and deep learning algorithms. These methods are essential for successfully integrating DERs into the load forecasting processes of distribution system operators. Furthermore, recent studies have proposed the use of advanced forecasting algorithms, such as artificial neural networks (ANN) and XGBoost, to enhance the accuracy of load forecasting, particularly in the context of integrating renewable energy into the power grid [71,73,74]. These studies spotlight the relevance of advanced algorithms in short-term load forecasting for effective power system planning and operation. Therefore, developing a mathematical formulation to represent DER forecasting for Distribution System Operators is necessary.

7.4. Mathematical formulation of DERs forecasting for DSOs

The mathematical formulation of DERs forecasting for DSOs is essential to consider, especially with the increasing penetration of DERs in the power grid. The impact of DERs on the power system necessitates adopting AI technology and methodologies for accurate forecasting [75]. Formulating forecasting models for DERs should take into account the optimization of grid energy management, including demand-side management, to ensure the efficient and reliable operation of the distribution network [76,77]. Moreover, integrating DERs into the wholesale electricity market and providing ancillary services, such as reactive power support, necessitates advanced mathematical optimization and market-clearing mechanisms [78,79].

In the context of South Africa, the mathematical formulation of DER forecasting for DSOs might necessitate the development of probabilistic forecasting methods, optimization models for energy management, coordination mechanisms for DER aggregator's operation, and market opportunities to integrate DERs into the distribution system effectively. Therefore, the following steps are being considered.

7.4.1. Key parameters and variables

Determination of the key parameters and variables play a crucial role in the DER forecasting problem. Common parameters and variables include.

- : Time period (e.g., hourly, 15-min intervals).
- : Index for each DER unit within the distribution grid.
- : Total number of DER units.
- : Total number of time periods in the forecasting horizon.
- $R_i(t)$: Actual generation or consumption of DER unit i at time (t).
- Load (t): Total electrical load on the distribution grid at time (t).

- Weather (t): Weather-related data at time (t) (e.g., solar radiation, wind speed).

  MarketPrice (t)*:* Electricity *market prices at time (t).*

- (t): Other exogenous variables (e.g., holidays, demand response events).

7.4.2. Decision variables

Decision variables with which are the forecasted values one want to obtain need to be defined prior, these include.

- DER_F (t): Forecasted generation or consumption of DER unit i at time t.
- Grid_F (t): The total forecasted grid load at time t.

7.4.3. Parameter constraints

Constraints that must be satisfied for the DER forecasting to be practical and realistic can be defined as.

- Generation and Consumption Constraints: Define constraints for each DER unit i and time period t to ensure that forecasts are within the appropriate bounds.

  For each DER unit i and time period **:**$DER\_\mathrm{F}_i(t) \geq 0$ (Generation)$DER\_\mathrm{F}_i(t) \leq 0$ (Consumption)

- Load Forecast Constraint:

  The total forecasted grid load at time t is the sum of DER forecasts and the load forecasts as given in equation (15):

$$\mathrm{Grid\_F}\ (t) = \sum_{i-1}^{n} DER\_F_{\boldsymbol{i}}\,(t) + \mathrm{Load}\ (t) \tag{15}$$

- Equation (16) provides Resource Availability Constraints on the availability of DER resources:

$$DER\_\mathrm{Forecast}\ _i(t) \leq DER\_Capacity_i(t) \tag{16}$$

- Equation (17) gives Weather-Related Constraints:

  Integrating weather data to adjust DER forecasts:

$$DER_{\mathrm{F}_i(t)} = f(DER\_Capacityi(t), \mathrm{Weather}\ (t)) \tag{17}$$

- Equation (18) depicts Market and Economic Constraints:

  Using market prices and economic factors for decision-making:

$$DER\_\mathrm{Forecast}\ _i(t) = g(\ \mathrm{MarketPrice}\ (t), X(t)) \tag{18}$$

7.4.4. Objective function

The objective function that quantifies the optimization goal can then be specified. The choice of the objective function depends on the specific objectives. For instance.

- *Load Balancing:* Minimizing the difference between the total forecasted grid load and actual load to reduce imbalances.$min \sum_{t-1}^{N} |Grid\_F(t) - Load(t)|$.
- *Economic Optimization*: Maximizing revenue or minimizing costs by adjusting DER operations based on market prices and economic factors. $\max \sum_{t-1}^{N} \bullet \mathrm{DER\_F}\ _i(t) \bullet \mathrm{MarketPrice}\ (t)$.

7.4.5. Solution method

An appropriate modeling and solution technique, which can range from linear programming to machine learning, depending on the complexity and nature of the problem can be selected ranging from.

- *Time Series Analysis:* These methods use historical load data to forecast future loads. Common techniques include ARIMA (AutoRegressive Integrated Moving Average), exponential smoothing, and various statistical approaches.
- *Machine Learning:* Machine learning models like neural networks, support vector machines, random forests, and gradient boosting are used to capture complex load patterns and dependencies. These models can also incorporate exogenous variables like weather data.

- *Hybrid Models:* Hybrid models combine multiple forecasting methods to leverage their strengths. For instance, combining ARIMA with neural networks or other machine learning models can improve forecasting accuracy.
- *Weather-Based Models:* Load is often affected by weather conditions. These models integrate weather data, such as temperature, humidity, and wind speed, into the forecasting process to provide more accurate predictions

7.5. Mathematical model of some of the DERs

A comprehensive mathematical model integrating EV, solar PV, wind, virtual wheeling, and batteries for load forecasting is complex due to the interactions between these diverse dynamic components. Each component has its own model, and their integration in this research involves considering time-dependent and weather-dependent variables. A simplified model is presented here. Note that real-world models are typically more complex and may apply advanced simulations.

7.5.1. Electrical vehicle mathematical model

Electric vehicle charging is characterized by spatial and temporal randomness and uncertainty [80] and a large number of electric vehicle loads connected to the grid inevitably affects the security and stability of grid operation, therefore it is important to predict the charging loads of electric vehicles. Generally, the EV charging model can be divided into 3 parts [81] which models the charging stations, EV mobility in the city and spatial load as shown in Fig. 12 below.

The model assumes three separate parking lot utilization profiles: work, home, and other. Each profile depicts parking for a specific type of building, such as Home, which symbolizes residential buildings. Users adhering to one or more profiles are expected to use each parking lot. The basic assumption is that in a major metropolis, the parking lot utilization profile is similar to the usage profile of neighbouring buildings. In other words, car drivers will park near visited locations [80] particularly if parking costs are not paid.

Private car travel times on weekdays are primarily concentrated in the morning and evening peaks whereas private car travel times on holidays and leisure are evenly distributed throughout the day and charging time is also later after returning home on the same day. The arrival times of private vehicles to residential, work, and commercial locations likewise satisfy the normal distribution, as indicated in equation (19), based on an analysis of the correlation fitted data.

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{19}$$

Where x is the starting charging time

The term "state of charge," or SOC, refers to how much of the battery's rated capacity is still present. The type of electric vehicle, driving distance, and power consumption per unit mileage all affect the start charging status. The battery charge status of electric vehicles is roughly linearly proportional to driving mileage, assuming that private cars have constant average power consumption. Equation (20) illustrates the battery charge state of an electric vehicle at a given location.

$$SOC(T_j) = SOC(T_i) - \frac{\bar{I}.D_{ij}}{C} \times 100\% \tag{20}$$

where T_i is the current destination's arrival time, T_j is the next destination's arrival time, $\bar{l}$ is the average power consumption per mile, D_{ij} is the distance between destinations, and C is the battery capacity.

The status of charge of the battery is displayed in equation (21), if the private vehicle arrives at a specific location to be charged, it is deemed to be charged until it is full or arrives at the next departure moment.

$$SOC(T_j) = min\left\{SOC(T_i) + \frac{\eta.p_c.T_t.100}{C} \times 100\%\right\} \tag{21}$$

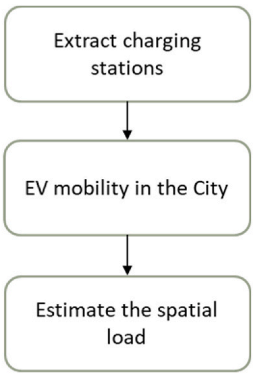

Fig. 12. A diagram depicting the three components of the EV charging model.

where p is the charging interface power, T_t is the charging time at the present location, and η is the charger's charging efficiency.

It is always presumed that the private vehicle is charged at the time of arrival and does not move until it is completely charged, that is, when it moves from position j to site i. $SOC = 1$ and that the charging power remains constant when the car is being charged.

7.5.2. Solar PV mathematical model

The present trends in photovoltaic (PV) deployments around the world indicate that PV energy will play a significant part in the energy mix in the near future. Various techniques have been put out for forecasting PV power. The techniques to provide the PV power predictions are based on two primary strategies. One is the physical method, which necessitates the availability of meteorological data in addition to previous understanding of the characteristics of PV materials and PV system information [82]. The second strategy is data-driven, meaning that operational data is needed to train or calibrate the model coefficients, which are subsequently utilized to provide predictions. Even in the event that the PV system is not yet commissioned, a physical technique can be applied, hence it is the most commonly used approach.

The charge carrier recombination mechanisms that one must take into account determine which model to choose. The one-diode model is the most widely used to simulate the working principles of a photovoltaic cell due to its simplicity. Fig. 13 below illustrates the module equivalent models for both physical and data driven approaches.

The fundamental premise that a physical model is made up of identical PV cells connected in series is typically assumed in order to develop the equivalent circuit model for a PV module [83]This presumption suggests that all of the PV cells should provide identical current and voltage under comparable circumstances (temperature and irradiation), which gives out equation (22).

$$I_M = I_{PH} - I_o\left[exp\left(\frac{V_{M+}R_S N_S I_M}{nN_S V_t}\right) - 1 - \left(\frac{V_{M+}R_S N_S I_M}{N_S R_{sh}}\right)\right] \tag{22}$$

Where the generated solar cell current, reverse saturation current, and photo-generated current are denoted by $I,\ I_o,\ and\ I_{PH}$, respectively. The solar cell voltage, series resistance, and shunt resistance are denoted by the letters $V, R_S, and\ R_{Sh}$. V_t is the thermal voltage, while n is the diode's ideal factor.

The operating solar cell temperature and sun irradiance have a direct relationship with the photo-generated current I_{PH}, one way to evaluate it is as equation (23).

$$I_{PH} = I_{SC}\frac{G}{G_{STC}} + k_1(T_C - T_{STC}) \tag{23}$$

where T_C is the cell temperature and G is the specified irradiance level. The temperature and irradiance at standard test conditions (STC) are denoted as $T_{STC}\ and\ G_{STC}$, respectively. I_{SC} is the short-circuit current at STC, and k_1 is the temperature coefficient of the current in ($A°C$).

One major flaw in the physical models covered above is that they require an excessive number of input variables, most of which are not readily available. Heuristic models are suggested in this situation to lower the quantity of inputs needed. They are based on the same concepts of generating power forecasts from irradiance and module temperature, but the number of fitting model parameters differs. Equation (24) presents such a data driven equation.

$$P_{mpp} = \left[a.G + b.G + c.G^2.\ln\left(\frac{G}{G_{STC}}\right)^2\right].(1 + y(T_C - T_{STC})) \tag{24}$$

where G is the simulated or observed irradiance and P_{mpp} is the power generated by a PV module at its maximum power point. To

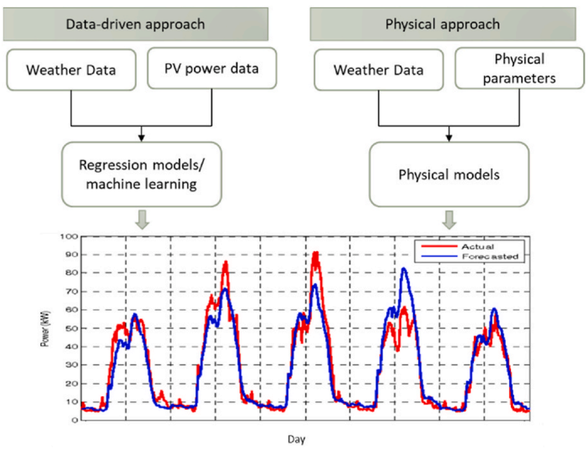

Fig. 13. Schematic of module equivalent circuits.

evaluate the cell temperature T_C, and the model fitting parameters are $a, b, c, x, and\, y$ are denoted. The temperature coefficient of power in $(\%/^\circ C)$ is represented by γ. At STC, the power and the irradiance are P_{STC} and $G_{STC} = 1000\omega/m^2$.

7.5.3. Wind mathematical model

The significant proportion of wind energy in the national power system causes numerous issues. They are caused by wind farms operating in an unstable manner due to variable wind speed. There are times when no electricity is produced during typical use [84]. There are numerous frameworks for estimating wind power demand over three different forecasting timeframes, including one step ahead, multiple step ahead, and probabilistic forecasting.

With single step ahead forecasting in equations (25) and (26) the estimation of any number today for the next day with the greatest precision and dependability is feasible [85]. Access to previous values of this quantity, data from one or more time series is possible, and other elements that influence the production of these time series.

$$WP_{t+1} = f(WP_t, \ldots, \ldots, WP_{t-d+1}) + e \tag{25}$$

With

$$t \in f\{d, \ldots, \ldots, N-1\} \tag{26}$$

According to equation (25) e represents the prediction error or noise present between the current forecasting value and the N previous observations. WP stands for wind power, T stands for target, and for many steps, the target matrix is increased with respect to each step-in advance, as shown in equations (27) and (28) respectively.

$$Single\ Step \begin{bmatrix} WP_{11} & --- & WP_{16} \\ WP_{21} & --- & WP_{26} \\ WP_{N1} & --- & WP_{N6} \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_N \end{bmatrix} \tag{27}$$

$$Second\ Step \begin{bmatrix} WP_{11} & --- & WP_{16} \\ WP_{21} & --- & WP_{26} \\ WP_{N1} & --- & WP_{N6} \end{bmatrix} \begin{bmatrix} T_2 \\ T_3 \\ T_{N+1} \end{bmatrix} \tag{28}$$

A pattern of values for a given time series is forecasted by multiple steps ahead or multiple lead hour prediction. It is a step-by-step approach that uses current prediction for deterministic next stage prediction. When the forecast period is long in the case of multi-step ahead prediction using equation (29), several anomalies such as error accumulation and data complexity prevail. It all happens as a result of the spread of bias and deviations from earlier predictions of future predictions [84]. Because of the vast forecasting horizon and the mistake involved in forecasting, this method suffers from poor performance and increasing inaccuracy due to the use of approximated values rather than actual values. As a result, choosing an input parameter function to suit a time series might be a difficult task for power system researchers.

$$K^{th} Step \begin{bmatrix} WP_{11} & --- & WP_{16} \\ WP_{21} & --- & WP_{26} \\ WP_{N1} & --- & WP_{N6} \end{bmatrix} \begin{bmatrix} T_k \\ T_{K+1} \\ T_{N+K} \end{bmatrix} \tag{29}$$

The application of probabilistic forecasting is produced by probabilistic forecast systems, which are intended to evaluate a forecast's uncertainty. A crucial component of probabilistic forecasting systems is verification [85].

7.5.4. Bess mathematical model

There hasn't been much prior research done on the implications of load forecasting on the profitability of household level BESS. Nonetheless, a large number of research with an emphasis on residential energy storage have been published [86]. examined the effects of load forecast errors in Ref. [9], comparing various error thresholds. When trying to maximize profit and profitability, the size of the BESS is crucial, sizing from an economic perspective.

BESS modeling is based on regulating the battery's state of charge (SOC), which is expressed as a percentage of full battery E_{max}. The SOC fluctuates every time the battery is charged or discharged, and it is dependent on the SOC at the preceding instant [87]. Equation (30) can be used to create this.

$$SOC_t = 100\frac{B_{eff} B_t}{E_{max}} + SOC_{t-1} \tag{30}$$

where SOC_t represents the SOC at time t and SOC_{t-1} represents the SOC at the preceding time step.

The efficiency of BESS B_{eff}, which is a mix of battery efficiency and power electronics efficiency, i.e., inverter/charger DC side is used to model the battery type. The efficiency of power electronics on the inverter/charger DC side dc is normally 99 %. Although charging and discharging losses are not equal, we can utilize average values for both processes when the battery is used cyclically. Equation (31) can be used to compute the charging and discharging efficiency η_C.

$$\eta_C = 100\frac{V_b - I_C R_b}{V_b} \tag{31}$$

where V_b is the battery's normal voltage and I_C is the charging or discharging current. Bt is obtained by multiplying charging or discharging current by charging or discharging voltage V_b, which is calculated using Equation (32).

$$V_C = V_b - I_C R_b \tag{32}$$

This control system, which is a novel structure, allows for the combination of several control targets. The control system's fundamental function is to instruct the inverter/charger when to charge and when to discharge the battery in BESS optimization.

Using equations (33) and (34), battery charging and discharging $\left(P_{\text{charge\_battery}}(t) \text{ and } P_{\text{discharge\_battery}}(t)\right)$ are influenced by the state of charge (SOC) and efficiency:

$$P_{\text{charge\_battery}}(t) = \eta_{\text{charge}} \cdot P_{\text{source}}(t) \tag{33}$$

$$P_{\text{discharge\_battery}}(t) = \frac{1}{\eta_{\text{discharge}}} \cdot P_{\text{load}}(t) \tag{34}$$

8. Critical evaluation of intermittent parameters and limitations

In this section, discusses a critical evaluation of intermittent parameters utilized in distributed energy resource (DER) forecasting techniques. Intermittent parameters, such as solar irradiance, wind speed, and cloud cover, play a significant role in renewable energy generation forecasting. However, their accurate prediction poses challenges due to inherent uncertainties and variability. We assess the methodologies employed to forecast these parameters, including numerical weather prediction models, satellite imagery analysis, and machine learning algorithms. Additionally, we discuss the sources of uncertainty associated with intermittent parameter forecasting, such as model inaccuracies and data limitations. Furthermore, we examine the impact of inaccurate intermittent parameter forecasts on DER integration and grid stability. Through this critical evaluation, we aim to provide insights into the strengths and limitations of current approaches and identify avenues for future research to improve the accuracy and reliability of intermittent parameter forecasting.

The summary of the limitations and advantages of deterministic and stochastic methods commonly used in DER forecasting. Deterministic methods, such as physics-based models and numerical simulations, offer a mechanistic understanding of system dynamics but may struggle to capture complex nonlinearities and uncertainties. On the other hand, stochastic methods, including time series analysis and machine learning algorithms, excel at capturing nonlinear relationships and uncertainties but may lack interpretability and generalization ability. We addressed how the choice between deterministic and stochastic methods depends on factors such as data availability, forecasting horizon, and desired accuracy. Additionally, we highlight the trade-offs between model complexity, computational requirements, and forecasting performance. By synthesizing the strengths and weaknesses of deterministic and stochastic methods, we provide guidance for DSOs in selecting appropriate forecasting approaches based on their specific needs and constraints.

9. Conclusion

This paper presents a comprehensive literature review on existing electrical load forecasting used in previous studies. Recent progress on emerging additional flexibility requirements in demand forecasting due to significant uncertainty and variability from increasing penetration of DERs has also been looked at. In this regard, the current paper provides existing load forecasting strategies in South Africa and also explored how the paradigm shift will influence the country's utility grid planners and forecasters. The influence of new technologies on load forecasting in South Africa encompasses technological innovation, economic growth, weather conditions, and the integration of renewable energy, electric vehicles, and energy storage systems. These transformative technologies are set to reshape the operation of South Africa's power system, exerting a profound effect on the precision of load forecasting. These technologies will change South Africa's power system operation and impact load forecasting accuracy. DSOs can effectively manage the challenges and opportunities associated with DER forecasting and integration. By embracing these technological advancements, DSOs can ensure the electrical distribution grids' reliability and stability and align with broader sustainability and adaptability objectives.

All potential scenarios involving integrating DERs for DSOs necessitate evolution and development processes including numerous factors, which are currently fascinating fields for study and research; in reality, the new issues that must be addressed concern.

- **Technical aspects**: new criteria for electrical system planning, design, control, and management;
- **Technological aspects**: evolution of components, apparatus, and systems (both hardware and software);
- **Economical and policy-regulatory aspects**: free markets, roles and responsibilities of all actors involved, connection rules, load shedding, generation curtailment, and so on;
- **Social aspects**: truly sustainable energy policies, reduction of environmental impact, et cetera;

It is envisaged that this study's conclusions will help realize SA's Just Energy Transition agenda and Sustainable Development Goals. The ability of DSOs to accurately predict and control the behaviour of distributed energy resources is hindered by a lack of basic understanding of modern methodologies for DER forecasting in distribution systems. Therefore DSOs employ various forecasting techniques to predict the behaviour of DERs. These techniques include statistical models (e.g., time series analysis, autoregressive integrated moving average models), machine learning algorithms (e.g., neural networks, random forests), and hybrid models that combine different methodologies. These models are trained using historical data and adjusted to incorporate real-time data for improved accuracy.

Data availability statement

Data and functional details presented in this study are available on request from the corresponding author.

CRediT authorship contribution statement

Nande Fose: Writing – review & editing, Writing – original draft, Investigation, Data curation, Conceptualization. **Arvind R. Singh:** Writing – review & editing, Visualization, Data curation, Conceptualization. **Senthil Krishnamurthy:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition, Formal analysis. **Mukovhe Ratshitanga:** Supervision, Project administration, Funding acquisition. **Prathaban Moodley:** Conceptualization, Review, and Funding acquisition.

Declaration of competing interest

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

Acknowledgments

Authors, with gratitude, acknowledge the financial and non-financial support for this work provided by South African National Energy Development Institute (SANEDI) JET Project RFQ0622 and the Department of Electrical, Electronic, and Computer Engineering> of the Cape Peninsula University of Technology.

References

[1] M. Bah, M. Azam, Investigating the relationship between electricity consumption and economic growth: evidence from South Africa, Renew. Sustain. Energy 80 (2017) 531–537.
[2] R. Samu, F. Bekun, M. Fahrioglu, Electricity consumption and economic growth nexus in Zimbabwe revisited: fresh evidence from Maki cointegration, Int. J. Green Energy 16 (2019) 540–550.
[3] N. Ouedraogo, Africa energy future: alternative scenarios and their implications for sustainable development strategies, Energy Pol. 106 (2017) 457–471.
[4] J. Erasmus, W. Malherbe, S. Zimmermann, A. Lorenz, M. Nachev, V. Wepener, B. Sures, N. Smit, Metal accumulation in riverine macroinvertebrates from a platinum mining region, Sci. Total Environ. 703 (2020) 134738.
[5] International Energy Agency, World Energy Statistics & Balances, IEA, Paris https://www.iea.org/data-and-statistics/data-product/world-energy-statistics-balances, Licence: Terms of Use for Non-CC Material.[Online].
[6] F. Bekun, F. Emir, S. Sarkodie, Another look at the relationship between energy consumption, carbon dioxide emissions, and economic growth in South Africa, Sci. Total Environ. 655 (2019) 759–765.
[7] A. Aliyu, B. Modu, C. Tan, A review of renewable energy development in Africa: a focus in South Africa, Egypt and Nigeria, Renew. Sustain. Energy 81 (2018) 2502–2518.
[8] O.M. Akinbami, S.R. Oke, M.O. Bodunrin, The state of renewable energy development in South Africa: an overview, Alex. Eng. J. 60 (6) (2021) 5077–5093.
[9] V. Lekavičius, A. Galinis, V. Miškinis, Long term economic impacts of energy development scenarios, the role of domestic electricity generation, Appl. Energy, Appl. Energy 235 (2019) 1254–1265.
[10] B. Verbrugge, S. Geenen, The gold commodity frontier: a fresh perspective on change and diversity in the global gold mining economy, Extract. Ind. Soc 6 (2019) 413–423.
[11] T. Mühlpfordt, X. Dai, A. Engelmann, V. Hagenmeyer, Distributed power flow and distributed optimization Formulation, solution, and open source implementation, Sustain. Energy Grids Netw 26 (2021) 100471.
[12] S. Meinecke, D.S. Stock, M. Braun, New distributed optimization method for TSO DSO coordinated grid operation preserving power system operator sovereignty, Energies 16 (12) (2023).
[13] V. Tetyana, O. Lyulyov, Y. Bilan, D. Streimikiene, Sustainable economic development and greenhouse gas emissions: the dynamic impact of renewable energy consumption, GDP, and corruption, Energies 12 (17) (2019).
[14] R.S. Kumar, L.P. Raghav, D.K. Raju, A.R. Singh, Impact of multiple demand side management programs on the optimal operation of grid connected microgrids, Appl. Energy 301 (2021).
[15] A. Srivastava, D. Steen, L.A. Tuan, O. Carlson, I. Bouloumpasis, Q.-T. Tran, L. Lemius, Development of a DSO support tool for congestion forecast, The Institute pf Enginneering and Technology 15 (23) (2021) 3345–3359.
[16] S. Jain, P. Jain, The rise of renewable energy implementation in South Africa, Energy Proc. 143 (2017) 721–726.
[17] C.W.J. Granger, Investigating causal relations by econometric models and cross spectral methods, Econometrica 37 (3) (1986) 424–438.
[18] C. Matta, N.M. Bianchesi, M. Oliveira, P. Balestrassi, L. Fabiano, "A comparative study of forecasting methods using real life econometric series data, Production 31 (2021), https://doi.org/10.1590/0103.6513.20210043.
[19] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, Computers & Mathematics with Applications, Comput. Math. Appl. 2 (1) (1976) 17–40.
[20] S. Tiwari, Electrical load forecasting methodologies and approaches, The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM 19 (2022) 1–8.
[21] K. Huarng, T. Yu, Y. Hsu, A multivariate heuristic model for fuzzy time series forecasting, IEEE Trans. Syst., Man Cybern 37 (4) (2007) 836–846.
[22] B.S. Nikhilkumar, V. Kale Pallavi, Short-term load forecasting by GRF methodology, International Journal of Technology Enhancements and Emerging Engineering Research 2 (4) (2014).
[23] I. Nti, M. Teimeh, O. Nyarko-Boateng, Electricity load forecasting: a systematic review, Journal of Electrical Systems and Inf Technol 7 (13) (2020).
[24] D. Phillips, J. Harris, M. de Wit, E. Matchan, Provenance history of detrital diamond deposits, West Coast of Namaqualand, South Africa, Miner Petrol. 112 (2018) 259–273.
[25] J. Wang, H. Lu, A hybrid system for short-term wind speed forecasting, Appl. Energy 226 (2018) 756–771.

[26] H. Nano, N. Savanna, C. Achraf, Chapter 16 - load forecasting using multiple linear regression with different calendars, in: Distributed Energy Resources in Microgrids, Academic Press, 2019, pp. 405–417.
[27] F. Nogales, F. Prieto, A. Conejo, A decomposition methodology applied to the multi area optimal power flow problem, Ann. Oper. Res. 120 (2003) 99–116.
[28] M. Granada, M. Rider, J. Mantovani, M. Shahidehpour, A decentralized approach for optimal reactive power dispatch using a Lagrangian decomposition method, Elec. Power Syst. Res. 89 (2012) 148–156.
[29] S. Motepe, A.N. Hasan, B. Twala, R. Stopforth, Power distribution networks load forecasting using deep belief networks: the South African case, in: 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 2019, pp. 507–512.
[30] J.H. Pujar, Fuzzy ideology based long term load forecasting, World Academy of Science, Engineering and Technology 40 (2010).
[31] A.K. Gangwar, F. Chishti, Electric load forecasting using genetic algorithm – a review, International Journal of Modern Engineering Research 4 (6) (2014) 2249–6645.
[32] J. Taylor, ort-term load forecasting with exponentially weighted methods, IEEE Trans. Power Syst. 27 (1) (2012) 458–464.
[33] Y. Wang, V. Ojleska, Y. Jing, Short term load forecasting: a dynamic neural network based genetic algorithm optimization, in: 14th International Power Electronics and Motion Control Conference, 2010, pp. 157–161.
[34] L. Ghods, M. Kalantar, Methods for long term electric load demand forecasting; A comprehensive investigation, in: IEEE International Conference on Industrial Technology, 2008, pp. 1–4.
[35] J. Eckstein, Parallel alternating direction multiplier decomposition of convex programs, Optim Theory Appl 80 (1994) 39–62.
[36] G. Janacek, Time series analysis forecasting and control, J. Time Ser. Anal. 3 (3) (2009) 131–133.
[37] H. Tong, K.S. Lim, Threshold autoregression, limit cycles and cyclical data, J. Roy. Stat. Soc. B 42 (3) (1980) 245–268.
[38] Y. Zheng, Predictive study of tuberculosis incidence by ARMA model combined with air pollution variables, Complexity 2020 (2020) 96–375.
[39] X. Lu, L.H. Wang, Bootstrap prediction interval for ARMA models with unknown orders, REVSTAT Stat. J. 18 (3) (2020) 96–375.
[40] A. Inoue, Closed form expression for finite predictor coefficients of multivariate ARMA, J. Multivariate Anal. 1761 (2020) 44–52.
[41] C. Dilmegani, Demand Forecasting in the Age of AI & Machine Learning, AIMultiple, 2024.
[42] Dario. Maradin, Advantages and Disadvantages of Renewable Energy Sources Utilization, Int. J. Energy Econ. Pol. 11 (2021) 176–183, https://doi.org/10.32479/ijeep.11027.
[43] B.K. Bose, Artificial intelligence techniques in smart grid and renewable energy systems—some example applications, Proccedings of the IEEE 105 (5) (2017) 2262–2273.
[44] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, A. Amira, Artificial intelligence based anomaly detection of energy consumption in buildings: a review, current trends and new perspectives, Appl. Energy 287 (2021) 116601.
[45] S.S. Ali, B.J. Choi, State-of-the-Art Artificial Intelligence Techniques for Distributed Smart Grids: A Review, Electronics 9 (6) (2020) 1–25. https://doi.org/10.3390/electronics9061030 1030.
[46] X. Wang, I. Yang, S.H. Ahn, Sample efficient home power anomaly detection in real time using semi-supervised learning, IEEE Access 7 (2019) 139712–139725.
[47] X. Wang, S.H. Ahn, " Real-time prediction and anomaly detection of electrical load in a residential community,", Appl. Energy (2020) 114145, p. 259.
[48] X. Wang, H. Wang, S.H. Ahn, Demand-side management for off-grid solar-powered microgrids: a case study of rural electrification in Tanzania, Energy 120229 (2021) 224.
[49] I. Antonopoulos, V. Robu, B. Couraud, D. Kirli, S. Norbu, A. Kiprakis, D. Flynn, S. Elizondo-Gonzalez, S. Wattam, Artificial intelligence and machine learning approaches to energy demand-side response: a systematic review, Renew. Sustain. Energy Rev. 130 (2020) 109899.
[50] S. Phimphachan, K. Chamnongthai, P. Kumhom, N. Jittiwarangkul, A. Sangswang, Energy and peak load forecast models using neural network for fast developing area, in: Proc. 2004 IEEE Communications and Information Technologies Conf, 2004, pp. 389–393.
[51] R. Mamlook, O. Badran, E. Abdulhadi, A fuzzyinferencemodelforshort-termloadforecasting, Elsevier Energy Policy 37 (2009) 1239–1248.
[52] E.L. Taylor, Short-term Electrical Load Forecasting for an Institutional/Industrial Power System Using an Artificial Neural Network" Master's Thesis, University of Tennessee, Knoxville, 2013.
[53] T. Hong, M. Gui, M. Baran, H. Willis, Modeling and forecasting hourly electric load by multiple linear regression with interactions. Power and Energy Soc. General Meeting, 2010, pp. 1–8.
[54] P. Santos, A. Martins, A. Pires, Short-term load forecasting based on ANN applied to electrical distribution substations, Universities Power Engineering Conf 1 (2004) 427–432.
[55] W. Lin, D. Wu, Residential electric load forecasting via attentive transfer of graph neural networks, in: Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21), 2021, pp. 2716–2722.
[56] Z. Osman, M. Awad, T. Mahmoud, Neural network based approach for short-term load forecasting, in: IEEE/PES Power Systems Conference and Exposition, 2009. Seattle.
[57] M.A. Hammad, B. Jereb, B. Rosi, D. Dragan, Methods and models for electric load forecasting: a comprehensive review, Sciendo Logistics & Sustainable Transport 11 (No 1) (2020) 51–76.
[58] A. Veit, C. Goebel, R. Tidke, C. Doblander, H. Jacobsen, Household electricity demand forecasting-benchmarking state-of-the-art methods (2014) 1–10. http://arxiv.org/abs/1603.00751.
[59] C. Sigauke, D. Chikobvuy, Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach, Orion 26 (2) (2010) 97–111.
[60] E.T. Ilunga, Short-Term Hourly Load Forecastingin South Africa Using Neural Networks, Faculty of Science, University of the Witwatersrand, Johannesburg, 2018.
[61] S. Motepe, A.N. Hasan, R. Stopforth, Improving load forecasting process for a power distribution network using hybrid AI and deep learning algorithms, in: IEE Access, 2019.
[62] A. Kaur, L. Nonnenmacher, C.F. Coimbra, Net load forecasting for high renewable energy penetration grids, Eergy 114 (2016) 1073–1084.
[63] L. Ljung, System identification, in: Signal Analysis and Prediction, Birkhäuser Boston, Boston, 1998, pp. 163–173.
[64] S.S. Fan, C.H. Jen, An enhanced partial search to particle swarm optimization for unconstrained optimization, Mathematics 7 (2019) 357.
[65] S.S. Fan, E. Zahara, A hybrid simplex search and particle swarm optimization for unconstrained optimization, Eur. J. Oper. Res. 181 (2007) 527–548.
[66] J. Snoek, H. Larochelle, R.P. Adams, Practical bayesian optimization of machine learning algorithms, arXiv (2012) 1–12, no. arXiv:1206.2944.
[67] Y. Li, Y. Zhang, Y. Cai, A new hyper-parameter optimization method for power load forecast based on recurrent neural networks, Algorithms 14 (6) (2021) 163.
[68] S. Naidoo, L. Jones, Feasibility of Neural Networks for Distribution Load Forecasting, Eskom, 2002.
[69] CSIR, Forecasts for Electricity Demand in South Africa (2017 – 2050) Using the Csir Sectoral Regression Model for the Integrated Resource Plan of South Africa, Department of Energy, 2017.
[70] J. Breedt, Transmission Demand Forecast Report for Planning Purposes for the Period 2022-2031, Eskom, 2021.
[71] D. Bae, B. Kwon, K. Song, Xgboost-based day-ahead load forecasting algorithm considering behind-the-meter solar generation, Energies 15 (1) (2021) 128.
[72] X. Li, V. Olanrewaju, E. Agyekum, M. El-Naggar, M. Alrashed, S. Kamel, Derterminants of load capacity factor in an emerging economy: the role of green energy consumption and technological innovation, Front. Environ. Sci. 10 (2022).
[73] S. Park, S. Park, S. Sons, Optimal vpp operation considering network constraint uncertainty of DSO, IEEE Access 11 (2023) 8523–8530, https://doi.org/10.1109/access.2023.3237692.
[74] A. Webberley, D. Gao, Study of Artificial Neural Network Based Short Term Load Forecasting, PESMG, 2013.
[75] Y. Allahvirdizadeh, H. Shayanfar, M. Moghaddam, Coordinated multi-stage expansion planning of transmission system and integrated electrical, heating, and cooling distribution systems, " let Renewable Power Generation 17 (2) (2022) 413–457.
[76] T. Abreu, T. Soares, L. Carvalho, H. Morais, T. Simao, M. Loura, Reactive power management considering stochastic optimisation under the portuguese reactive power policy applied to der in distribution networks, Energies 12 (21) (2019) 2048.

[77] M. He, F. Zhang, Y. Haung, J. Chen, J. Wang, R. Wang, A distributed demand side energy management algorithm for smart grid, Energies 12 (3) (2019).
[78] J. Prado, Q. Wang, L. Qu, J. Aguero, The next-generation retail electricity market in the context of distributed energy resources: vision and integrating framework, Energies 12 (3) (2019).
[79] H. Chen, H. Li, C. Lin, X. Jin, R. Zhang, X. Li, An integrated market solution to enable active distribution network to provide reactive power ancillary service using transmission–distribution coordination, let Energy Systems Intergration 4 (1) (2021) 98–115.
[80] X. Zhou, K. Zhang, X. Jia, J. Fan, Electric Vehicle Charging Load Demand Forecasting Model Based on Spatial and Temporal Characteristics, 2022. Zhejiang, China.
[81] M. Shepero, D. Van der Meer, J. Munkhammar, J. Widén, Residential probabilistic load forecasting: a method using Gaussian process designed for electric load data, Appl. Energy 218 (2018) 159–172.
[82] I. Kaaya, J. Ascencio-Vasquez, Photovolic Power Forecasting Methods.Solar Radiation-Measurement, Modeling and Forecasting Techniques Foir Photovoltaic Solar Energy Applications, IntechOpen, 2022, https://doi.org/10.5772/intechopen.97049.
[83] J. Ascencio-Vasquez, I. Kaaya, K. Brecl, K. Weiss, M. Topic, Global climate data processing and mapping of degradation mechanisms and degradation rates of PV modules, Energies 12 (2019) 4749, https://doi.org/10.3390/en12244749.
[84] S. Borowski, A. Sołtysiak, K. Migawa, A. Neubauer, Simple Mathematical Model to Predict the Amount of Energy Produced in Wind Turbine-Preliminary Study, 2021. Bydgoszcz, Poland.
[85] S. Saroha, S.K. Aggarwal, P. Rana, Wind Power Forecasting-Forecasting in Mathematics ,Recent Advances, New Perspectives and Applications, IntechOpen, 2020, https://doi.org/10.5772/intechopen.94550.
[86] P. Koponen, J. Ikäheimo, J. Koskela, C. Brester, H. Niska, Assessing and comparing short term load forecast performance, Energies 13 (2054) (2020).
[87] J. Koskela, A. Mutanen, P. Järventausta, Using load forecasting to control domestic battery energy storage systems, Energies 13 (3946) (2020).