International Journal of Electrical and Electronics Engineering Studies, 10 (1), 1-18, 2024
ISSN 2056-581X (Print),
ISSN 2056-5828(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Advanced Forecasting Techniques and Grid Management Strategies

Tuanima Pepple
Bachelor of Electrical and Electronics Engineering
Department of Electrical and Electronics Engineering
University of Port Harcourt

doi: https://doi.org/10.37745/ijeees.13/vol10n1118 Published July 31, 2024

Citation: Pepple T. (2024) Advanced Forecasting Techniques and Grid Management Strategies, *International Journal of Electrical and Electronics Engineering Studies*, 10 (1), 1-18

Abstract: *Energy forecasting is crucial for addressing challenges in data-rich smart grid (SG) systems, encompassing applications such as demand-side management, load shedding, and optimal dispatch. Achieving efficient forecasting with minimal prediction error remains a significant challenge due to the inherent uncertainty in SG data. This paper provides a comprehensive, application-focused review of advanced forecasting methods for SG systems, highlighting recent advancements in probabilistic deep learning (PDL).The review extensively examines traditional point forecasting methods, including statistical, machine learning (ML), and deep learning (DL) techniques, evaluating their suitability for energy forecasting. Additionally, the importance of hybrid approaches and data preprocessing techniques in enhancing forecasting performance is discussed.A comparative case study utilizing the Victorian electricity consumption in Australia and American Electric Power (AEP) datasets is conducted to assess the performance of deterministic and probabilistic forecasting methods. The analysis reveals that DL methods, with appropriate hyper-parameter tuning, exhibit superior efficacy when dealing with larger sample sizes and nonlinear patterns. Moreover, PDL methods demonstrate at least a 60% reduction in prediction errors compared to other benchmark DL methods. However, the increased execution time for PDL methods, due to the large sample space, necessitates a balance between computational performance and forecasting accuracy.*
Keywords: advanced forecasting techniques, grid, management, strategies

INTRODUCTION

Energy forecasting plays a pivotal role in planning, investment, decision-making, and addressing operational and management challenges in modern power systems, commonly referred to as smart grid (SG) systems. The advent of smart meters and advanced metering infrastructure (AMI) in SG has significantly increased the bidirectional flow of energy and data between the grid and end-users. Consequently, numerous data analytics applications, such as energy forecasting, have emerged in the SG domain. These applications are highly beneficial for scheduling generation, implementing demand response strategies, and ensuring financial benefits through optimal bidding in the energy market.

Traditionally, statistical methods such as autoregressive integrated moving average (ARIMA) have been extensively used for forecasting energy demand and generation. However, with the breakthrough of smart metering and the resulting high volume of data generation, statistical methods face scalability issues and cannot effectively analyze complex nonlinear data features.

In recent years, artificial intelligence (AI)-based methods, including machine learning (ML) and deep learning (DL) algorithms, have gained significant attention for their ability to generate accurate forecasts in SG systems. Sequence-based DL algorithms, such as recurrent neural networks (RNN) and long short-term memory (LSTM) models, have proven to be effective in handling nonlinear energy data features with long sequences. Additionally, SG data generation experiences stochastic uncertainties due to renewable generation intermittencies and variations in energy consumption behaviors. Addressing parametric or model uncertainties is another challenge for traditional deterministic forecasting methods, which provide point forecasts. In such cases, probabilistic methods, which generate prediction intervals (PIs), are more effective in managing uncertainties compared to point forecasting methods.

Recently, probabilistic deep learning (PDL) has emerged as a more efficient approach for forecasting by integrating deep neural networks and Bayesian inference. However, this approach requires further investigation and offers substantial potential in modern power systems and forecasting applications.

Assembling multiple methods into a hybrid approach can significantly improve forecasting accuracy. However, the increased model complexity is a potential drawback, necessitating a tradeoff between accuracy and computational complexity. Data preprocessing plays a crucial role in enhancing model performance and minimizing forecasting error. Dimensionality reduction (DR) and feature extraction are key approaches often adopted for effective data preprocessing. Techniques such as principal component analysis (PCA) and singular value decomposition (SVD) have been widely used to address high-order dimensionality challenges in SG data. In the DL domain, auto-encoders (AE) implemented with convolutional layers have been used as a feature extraction scheme to filter multiple dimensions in consumption data.

While numerous research papers have surveyed ML methods for solar irradiance forecasting and reviewed state-of-the-art DL methods for renewable energy forecasting, none have thoroughly investigated advanced DL methods such as RNN and LSTM for energy forecasting applications, nor compared statistical, ML, and DL methods. Furthermore, no review paper has yet considered recent developments in probabilistic forecasting, particularly the contributions of probabilistic deep learning for energy forecasting applications in SG systems. This motivates a holistic review

International Journal of Electrical and Electronics Engineering Studies, 10 (1), 1-18, 2024
ISSN 2056-581X (Print),
ISSN 2056-5828(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

of state-of-the-art forecasting methods, along with recent developments in the PDL domain, in an application-oriented manner.

Specifically, this paper makes the following contributions:

- A comprehensive review of statistical, ML, DL, probabilistic, and hybrid forecasting methods, along with their applications in SG systems across different time horizons. The existing preprocessing techniques to aid energy forecasting performance are also discussed.

- Implementation of various statistical, ML, DL, and PDL techniques on multiple energy consumption datasets for load and renewable generation forecasting to compare forecasting accuracies. The impact of high variability and data sizes on different methods is also evaluated. The analysis shows that RNN and LSTM can achieve higher accuracies with the least prediction errors under deterministic forecasting methods with larger dataset sizes, especially in the presence of high variability. However, accuracy can vary with the choice of activation function and hyper-parameter tuning, which needs to be appropriately selected for a given dataset.

- Among probabilistic forecasting techniques, Bayesian bidirectional LSTM (BLSTM) is observed to outperform deterministic methods by exhibiting the least error and tighter prediction intervals for different energy forecasting datasets. Although PDL methods can improve forecasting accuracy, they have a higher execution time compared to deterministic methods, necessitating a tradeoff between model performance and computational cost.

The rest of the paper is organized as follows: Section 2 discusses the applications of forecasting in SG systems. Section 3 outlines the time horizon-based categorization of forecasting methods. Section 4 describes the taxonomy of energy forecasting methods. Section 5 elaborates on different deterministic forecasting methods. Contributions related to probabilistic forecasting methods are described in Section 6. Several hybrid methods are presented in Section 7. Data preprocessing methods are briefly discussed in Section 8. Section 9 presents a comparative case study of energy forecasting methods on two different datasets. Finally, Section 10 concludes the paper and discusses future directions.

APPLICATIONS OF FORECASTINGIN ENERGY SYSTEMS

Energy forecasting with high accuracy and precision is essential for grid planning and reliability, offering a wide range of applications such as load shedding, optimal dispatch, and peak load shaving. The data generated from smart grid (SG) systems typically pertains to energy consumption by residential, commercial, and industrial consumers, as illustrated in Figure 1. Additionally, data is generated from distributed energy resources (DERs) like solar and wind power plants. This data is utilized by forecasting methods to predict energy generation, demand, and prices within SG systems.

The authors in [25] surveyed various challenges and trends in demand forecasting across different time horizons and regions. Similarly, in [26], the authors compared existing statistical methods for demand forecasting, emphasizing the importance of achieving the lowest root-mean-square error (RMSE). Further, Kong et al. in [27] employed deep learning (DL) methods to forecast short-term energy consumption and demand for individual households using highly granular data.

With increasing incentives for demand response programs, accurately forecasting electricity demand has become crucial. Additionally, forecasting the flexibility with which users can adjust their energy demand has emerged as a significant challenge. Existing research has focused on developing more accurate models for demand response calculation through feature engineering [28]. The authors in [29] incorporated renewable generation forecasts to improve the accuracy of energy optimization algorithms for demand response. On the other hand, the impact of load and price forecasts on optimal demand management has been elaborated in [30]. Recently, demand flexibility prediction using machine learning (ML) models has been considered in [31], taking into account the demand response potentials of electric vehicles and hot water systems.

Energy generation using renewable-based DERs is more intermittent due to exogenous factors such as weather changes and user behaviors, making high forecasting accuracy a challenging task. In this regard, [33] used vector autoregressive (VAR) models to forecast solar irradiance, temperature, and wind speed for 61 locations in the United States. Similarly, Messner et al. in [34] used the VAR method to forecast wind power generation based on high-dimensional data.

Forecasting price spikes is another critical application, as it addresses the biggest risk factor in the energy market. Recent zero or negative pricing, as observed in the Australian national electricity market (NEM), can negatively impact generators [35]. Conversely, positive high price spikes can yield higher profits for generators, especially during peak generation hours [36]. Yang et al. surveyed the latest trends in decision-making for electricity retailers using consumed load price forecasting [37]. Furthermore, [38] utilized a set of relevance vector machines to predict prices for individual time periods ahead of time and implemented a micro-genetic algorithm to optimize linear regression ensemble coefficients for aggregated price forecasts. Toubeau et al. in [39] utilized probabilistic deep learning (PDL) methods to forecast wind and photovoltaic (PV) generation, subsequently predicting electricity prices generated from renewable DERs. Moreover, the locational marginal price for optimal scheduling of energy storage systems is determined using artificial neural networks (ANN) in [40].

International Journal of Electrical and Electronics Engineering Studies, 10 (1), 1-18, 2024
ISSN 2056-581X (Print),
ISSN 2056-5828(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

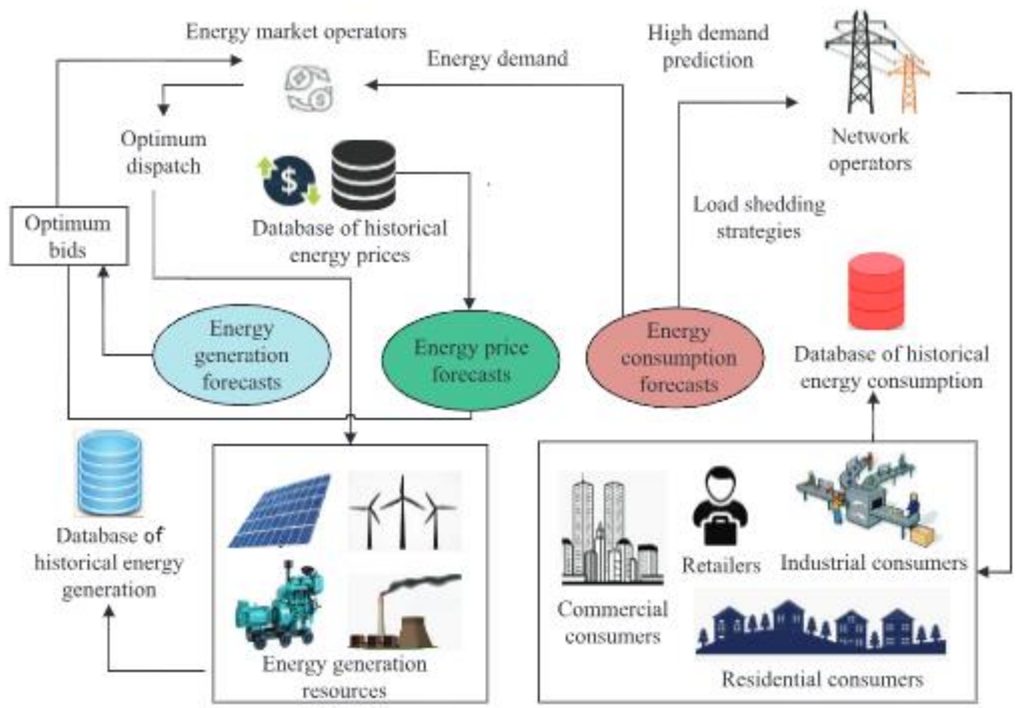


FIGURE 1 Overview of forecasting applications for energy systems

Energy forecasting is a multifaceted field that leverages various statistical, ML, DL, and PDL methods to address the complexities of modern power systems. The integration of advanced forecasting techniques with robust data preprocessing and hybrid approaches can significantly enhance forecasting accuracy and reliability, ultimately contributing to more efficient and resilient smart grid operations.

TAXONOMY OF THE ENERGYFORECASTING TECHNIQUES

This section elaborates on the categorization of state-of-the-art methods and recent advancements in energy forecasting systems, along with their associated literature and applications. We have employed various criteria to classify energy forecasting methods, including time horizon, representation of the forecasting output (point vs. distribution), and model performance (using error metrics). These criteria effectively capture the different dimensions of forecasting problems, such as the forecast time horizon, forecast variability, and model learning mechanisms.

International Journal of Electrical and Electronics Engineering Studies, 10 (1), 1-18, 2024

ISSN 2056-581X (Print),

ISSN 2056-5828(Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

Based on these criteria, energy forecasting methods can be broadly classified into very short-term forecasting (VSTF), short-term forecasting (STF), medium-term forecasting (MTF), and long-term forecasting (LTF), as detailed in the previous section. Regarding the type of forecasting model output, methods can be grouped into deterministic and probabilistic categories. Deterministic methods generate point forecasts, while probabilistic methods produce forecasts in terms of prediction intervals.

In terms of how forecasting models are trained to predict future demand and generation, methods can be categorized as statistical, artificial intelligence (AI)-based (including machine learning (ML) and deep learning (DL)), quantile regression-based, and the recent probabilistic deep learning (PDL) methods, as illustrated in Figure 3. Statistical methods are traditional and simpler, whereas learning-based methods, involving ML, DL, and their variants, are considered more accurate but also more complex.

It is important to note that the classification of these methods may have some overlaps among different groups. For instance, statistical methods, typically explained under deterministic techniques, can also be formulated as probabilistic methods and are discussed under probabilistic techniques. Similarly, ML and DL methods can be combined with probabilistic methods and are covered under PDL methods. Furthermore, hybrid methods and data preprocessing techniques are categorized under both deterministic and probabilistic methods, as they can be utilized in either context.The following sections provide a detailed discussion of state-of-the-art and advanced forecasting methods, along with their literature and applications for SG systems, in accordance with the categorization reflected in Figure 3.

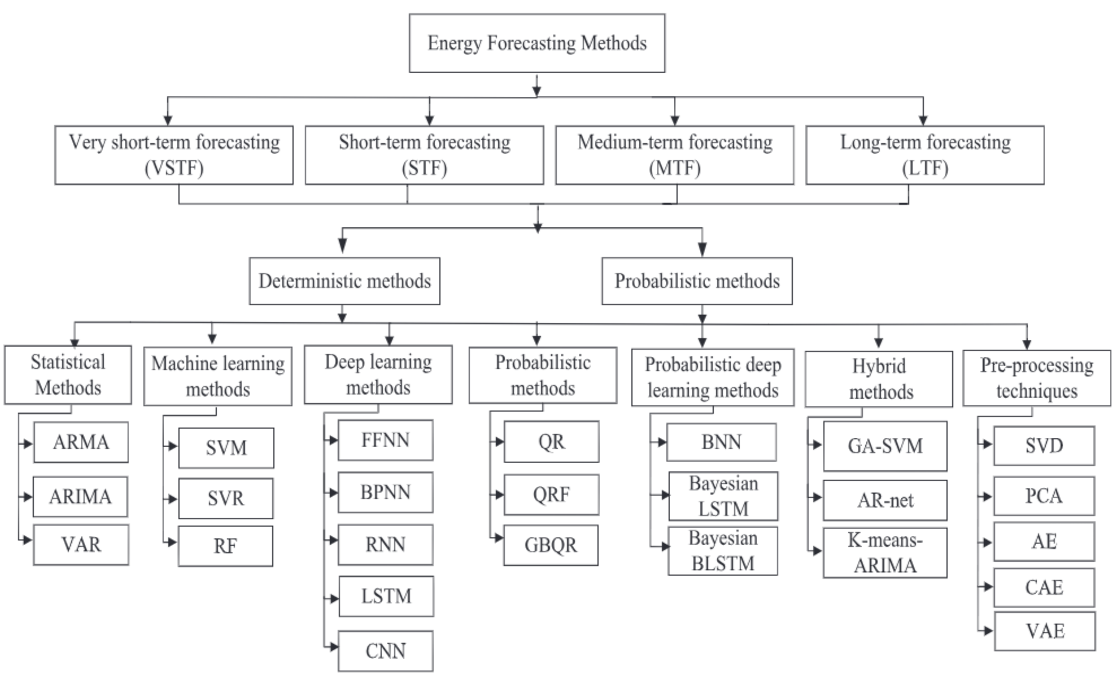


FIGURE 3 Taxonomy of energy forecasting methods for smart grid systems

Deterministic techniques generate forecasts in terms of fixed points rather than a probability distribution. These methods are also termed as point forecasting methods.5.1 Statistical methods Statistical methods are classical in nature, widely used for time-series forecasting especially for STF using the historical SG data. These methods are based on fitting a regression model on historical data points and then validating the model by find-ing the difference between actual and predicted values. This subsection stresses about existing statistical methods for the applications of SG systems forecasting.

Autoregressive moving average model (ARMA)ARMA

is the fundamental statistical method widely employed for time-series analysis. It is the combination of autoregressive (AR) and moving average (MA) methods. Sansa et al.in utilized ARMA to predict the solar irradiation for a winter day with maximum 10% variation in the generation. However, ARMA can only be applied to a stationary time-series, for which mean, and variance do not change over time and has a uniform distribution.

Autoregressive integrated moving average (ARIMA)

To deal with the non-stationary time-series, ARMA is generalized to ARIMA. ARIMA is employed in [56] for energy demand forecasting to charge the electric vehicles (EVs) using historical load data. Similarly, it has been used to forecast energy usage in and [58] for control and optimization in residential microgrids. However, the ARIMA is more suitable for linear time-series as it gives relatively high RMSE values for nonlinear data and requires a high execution time for large data sets.

Vector autoregression (VAR)

VAR model is defined as an extension to the univariate auto regression, which captures linear dependencies among multiple time-series. The authors in [33] predicted temperature, solar irradiation, and wind speed using VAR and exhibited lower RMSE values as compared to other statistical methods such as simple persistence (SP). VAR is also used in [34] for short-term wind power forecasting which is integrated with lasso estimator to recursively update the VAR parameters

Autoregressive/autoregressive integrated moving average with exogenous inputs(ARMAX/ARIMAX) ARMAX/ARIMAX

models take into account external variables that influence the forecasting accuracy of a desired parameter. It combines the AR and MA processes (along with the integrated component for ARIMA) with the external time-series parameters [60]. The authors in [60] utilized temperature, humidity, irradiance duration parameters as exogenous inputs to improve the accuracy of the solar generation forecasting. On the other hand, the authors in [61] have employed satellite images as exogenous variables integrated with recursive least squares filter. The authors in [62] have developed a seasonal ARIMA framework with weather variables and seasonality as exogenous inputs for forecasting load demand. Though the aforementioned methods achieve good forecasting accuracy, the methods that are discussed in the following subsection are found to be more efficient and scalable to deal with nonlinear traits of SG datasets with multiple time-series.

PROBABILISTIC TECHNIQUES

With the integration of RES in the modern power grid, forecasting trends are shifting from point to probabilistic in regards to the future demand and generation at disaggregated levels [93].Hong et al. presented a review for probabilistic methods and emphasized their importance over point forecasting with everchanging needs of power industry

Parametric versus non-parametric approaches

This subsection identifies the existing literature for probabilistic energy forecasting, which is mainly categorized under parametric and non-parametric approaches. The authors in provided a brief review of these two approaches for wind generation forecasting. Parametric approaches assume a certain probability density function for the parameter distribution, such as normal distribution. Dowell et al. proposed a parametric probabilistic scheme based on Bayesian probability and sparse VAR to forecast very short-term wind power generation in Southeastern Australian wind farms with a 5 min interval [96]. The authors confirmed that their method achieves least RMSE in comparison to the standard AR and VAR methods. They further utilized the similar parametric approach in [51] to forecast long-term probabilistic horizons for load consumption. A hybrid probabilistic deterministic approach for wind generation forecasting has been developed, where temporally local Gaussian processes are used to investigate forecasting errors. Furthermore, for the application of price forecasting, probabilistic methods have been utilized by various authors and the relevant contributions are outlined in [98–100]. Though this approach simplifies the analysis and reduces computation cost, sometimes the parameter distribution may not accurately fit with a known function. On the other hand, non-parametric approaches do not assume a fixed function for the probability density of the output parameters. In this approach, the predictive probability densities of the parameter are represented by a range of quantile forecasts. For wind power forecasting in the presence uncertainties, the authors in implemented a kernel density estimation model and represented the wind power with a number of kernel functions. An ensemble-based probabilistic forecasting model was developed in [102] which transforms meteorological data to wind power output and generates predictive distributions in a non-parametric manner. On the other hand, a non-parametric approach is considered in [103] to obtain prediction intervals as outputs of neural network models for generating wind power forecasts.

International Journal of Electrical and Electronics Engineering Studies, 10 (1), 1-18, 2024
ISSN 2056-581X (Print),
ISSN 2056-5828(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

State-of-the-art methods

Quantile regression

Quantile regression is an extended version of linear regression and aims to obtain predictive quantiles [9]. Taieb et al. in [104] emphasized the need to predict the future demand in probabilistic format, which can further support the grid operations related to the energy generation and distribution. They employed additive quantile regression (QR) to forecast the individual energy consumption for 3696 smart houses in Ireland with 30 min interval. To support probabilistic load forecasting, Liu et al. pro-posed quantile regression averaging (QRA) to obtain PIs for consumption with 90% percentiles [10]. The authors claimed effective forecasts using probabilistic evaluation metrics such as Pinball and Winkler score. However, the existing probabilistic approaches in energy forecasting often suffer high computational complexities and thus, more efficient methods need to be developed and reviewed considering the current energy market scenarios [105,106].· Quantile regression averaging (QRA): It is a combination approach for forecasting to compute prediction intervals. It involves applying quantile regression to the smaller forecast models in a hybrid form. The authors in [107] performed price forecasting using QRA to improve the profitability of electricity trading. In terms of load forecasting, the authors in [108] performed QRA on point forecasts to develop a computationally inexpensive yet accurate estimate of the load variability.· Quantile regression Factoring (QRF): This approach utilizes a fac-tor model to account for unobserved factors that impacts the distribution of the observable parameters. Quantile fac-tor analysis is performed to estimate the number of factors for each quantile [109]. The authors in [110] developed a multi-factor QR model to demonstrate how electricity price forecasts are impacted by volatility of market prices, demand forecasts as well as oil, coal and gas prices. In [99] authors extended QRA integrated with factor models and principal component analysis (PCA) to provide interval price forecasts. After applying QRA, PCA is used to select from individual forecast models available for averaging.

Robabilistic deep learning (PDL)

Bayesian probability incorporated with DL methods can be used to provide forecasting results in the form of PIs, con-trary to traditional deep neural networks that are deterministic in nature and generate point forecasts. PDL expresses model parameters as a function of probability distributions such as nor-mal distribution. To be precise, PDL models can predict future with different percentiles that can explain the certain or uncertain factors in the energy data and hence, enables better decision making. This section outlines the significant contributions in the field of PDL to support the applications of demand, generation, and price forecasting in modern power systems. Bayesian neural networks (BNN): The concept of Bayesian probability incorporated with ANN is defined as BNN. Yanget al. in [12] proposed BNN technique to forecast individual energy demands at household level after quantifying the shared uncertainties between different customer groups.

In addition, a clustering-based data pooling system is presented to tackle the issue of over-fitting by increasing the data volume and diversity. The authors demonstrated lower Winkler and Pinball scores for probabilistic methods. Furthermore, a comparative study is presented with the benchmark point forecasting and probabilistic techniques such as QRA and quantile regression factoring (QRF).· Bayesian LSTM: Sun et al. proposed a BNN integrated LSTM approach to curb the challenges posed by weather uncertainty in distributed PV generators and thereafter generated the net-load forecasts in the form of PIs with greater accuracies [77]. In addition, they improved the forecast-ing performance by clustering individual sub profiles based on similar energy consumption patterns prior to applying Bayesian approach. They implemented their method on areal SG dataset acquired from Ausgrid involving rooftop PV generation measurements for every half an hour for a period of 3 years.· Bayesian BLSTM: In a similar manner, authors in [39] pro-posed a PDL technique to deal with the problem of uncertainty in energy markets. The authors integrated LSTM with bidirectional RNN by enabling the propagation of training sequence forwards and backwards and proposed their method as bidirectional-LSTM (BLSTM). The pro-posed network is then trained to generate a Gaussian non-parametric predictive distribution of the dependent variables present in the energy data such as PV generation. Furthermore, Copula-based sampling is employed over predicted distributions to generate predictive scenarios However, probabilistic methods are computationally expensive due to large sample space. In this regard, [111] proposed dropout as one of the potential solutions which works as an approximator to make the Bayesian inference process less complex, computationally. However, the computational complexity of the Bayesian approach remains one of the main concerns and therefore, more generalized and effective solutions need to be explored in the future research.· Deep quantile regression: These methods combine QR methods with deep learning models to predict the quantiles involved with the uncertainty. The authors in [112] used QR neural network to forecast load probability density in a short-term horizon. On the other hand, regional wind power fore-casts using QR neural networks have been investigated in[113], where the authors used ramp functions to avoid the crossing of multiple quantiles. Similarly, LSTM integrated with penalized QR has been implemented in [114] to obtain probabilistic load forecasts in terms of PI.· Deep Gaussian processes: Deep Gaussian processes involve stacking of Gaussian processes as in the layers of neural net-works. Each Gaussian process layer works as a single layer neural network. These processes can learn the model uncertainty similar to Bayesian models. Deep Gaussian processes have been used for load forecasting in [115] and [116]. The authors in [115] found that this method becomes computationally intensive for larger data sets in comparison to other benchmark probabilistic deep learning methods. However, the benefit lies in the requirement for smaller number of training samples, which is critical for STF.

CASE STUDY

Comparative analysis of benchmark methods

This section presents a case study for comparative analysis of various forecasting methods discussed in above sections using two different datasets. To be specific, RNN and LSTM are chosen as benchmark DL models due to their superior performance reported in existing papers. For similar reasons, ARIMA, SVR, Bayesian ANN, and BLSTM are chosen as representatives of commonly used statistical, ML, and PDL methods, respectively.

Data description

In order to evaluate the performance of forecasting methods, two energy datasets from the AEP [147] and Victorian energy consumption benchmark [148] are taken. AEP dataset con-sists of hourly energy consumption values from October 2004to August 2018. Victorian dataset includes half hourly energy consumption values from April 2012 to March 2014 for 25 Victorian houses. To implement point forecasting methods, 5 time lags (equivalent to 5 h for AEP dataset and 2.5 h for Victorian dataset) are considered. For PDL methods, 24 time lags (equivalent to 24 h for AEP dataset and 12 h for Victorian dataset)are considered. Results are implemented in python environment using Keras and Tensor flow libraries.

Evaluation metrics

RMSE and MAE are the most commonly used evaluation metrics to evaluate the point forecasting methods [39],[77]. The RMSE represented by ϵ and MAE represented by ρ are define during following equations:

$$\epsilon = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_{pred}(t) - Y_{act}(t))^2}, \qquad (1)$$

$$\rho = \frac{1}{n}\sum_{t=1}^{n} | Y_{pre(t)} - Y_{act(t)} |, \qquad (2)$$

where $Y_{pred}(t)$ and $Y_{act}(t)$ are the predicted and actual values at time stamp t and n is the total number of samples, respectively. Thus, the main objective for forecasting model is to minimize the error in predicted values as:

$$\min_{\beta} \epsilon, \rho, \qquad (3)$$

where θ represents the model parameters for each forecasting method such as weights and biases or lag coefficients. Furthermore, average Pinball loss is an important evaluation metric for probabilistic methods [12, 77]. So, to evaluate Bayesian ANN and Bayesian BLSTM, average Pinball loss is computed focusing on the sharpness and consistency of the approximated distribution.

In this regard, least value of Pinball loss is more desirable. For actual data values and predictions at time-stamp (Yt ,q), Pinball loss over percentile q ∈ [0, 1] is formulated as:

$$Pinball(Y_{act}, \hat{Y}_{t,q}, q) = \begin{cases} (Y_{act} - \hat{Y}_{t,q})q & Y_t \geq \hat{Y}_{t,q} \\ (\hat{Y}_{t,q} - Y_{act})(1-q) & Y_{act} \leq \hat{Y}_{t,q} \end{cases}. \quad (4)$$

Furthermore, for confidence $(1-\gamma) \times 100$, Winkler score is computed using:

$$Winkler = \begin{cases} \delta & lb_t \leq y_t \geq ub_t \\ \delta + 2(lb_t - y_t)/\gamma & lb_t \geq y_t \\ \delta + 2(y_t - ub_t)/\gamma & ub_t \leq y_t \end{cases}, \quad (5)$$

where lbt and ubt represent the lower and upper bounds of prob-abilistic forecasts at interval t , respectively. And, δ = ubt − lbt isthe PI width at t .Furthermore, skill score in the terms of Brier Score (BS)[149] is also computed to evaluate the error in probabilistic predictions defined as:

$$BS = \frac{1}{n}\sum_{t=1}^{t=n}(f_t - y_t)^2, \quad (6)$$

where f t stands for estimated forecasting values at interval tusing approximated posterior distribution and n is the number of samples form testing set.

RESULT DISCUSSION

As shown in Table 3, both datasets are further divided into two cases each to evaluate the impact of different sample sizes. For the Victorian energy, 2 and 1 year of consumption sample size are considered. While for AEP, 10 and 1 year of consumption data are considered. Then, ARIMA, SVR, RNN, LSTM,CNN, Bayesian ANN, and Bayesian BLSTM are trained with70% of the total samples for each case and tested for remaining30%. For DL methods, 40 neural units per layer,

International Journal of Electrical and Electronics Engineering Studies, 10 (1), 1-18, 2024
ISSN 2056-581X (Print),
ISSN 2056-5828(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

100 epochs, and batch size of 1000 are chosen. Adam is used as an optimizer to minimize the loss function of mean squared error(MSE). Additionally, RNN and LSTM methods are considered with two different variations in the form of activation functions namely, hyperbolic tangent function (tanh), and rectified linear unit (Relu). Activation functions play an important role in activating the hidden layer in neural networks. Also, SVR is modeled with radial basis function (rbf) and linear kernel. Then, RMSE and MAE values are computed on the testing dataset using (1)and (2), respectively. As demonstrated in the table, Bayesian BLSTM performs similar to LSTM (with tanh) for larger training datasets and stationary time-series, that is, with 10 years of data for AEP.

TABLE 3 Comparative analysis on evaluation metrics for Victorian and AEP case study

		Victorian energy consumption dataset				AEP dataset			
		2 years		1 year		10 years		1 year	
Sr. no.	Method	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
1	ARIMA	–	–	0.0753	0.0566	–	–	0.0429	0.0370
2	SVR-rbf	0.0813	0.0795	0.0792	0.0605	0.0363	0.0298	0.0415	0.0331
3	SVR-linear	0.0676	0.0651	0.0749	0.0562	0.0310	0.0242	0.0362	0.0285
4	RNN-tanh	0.0736	0.0559	0.0774	0.0579	0.0193	0.0141	**0.0371**	**0.0285**
5	RNN-relu	**0.0631**	**0.0463**	**0.0701**	**0.0519**	0.0290	0.0228	0.0527	**0.0424**
6	LSTM-tanh	0.0805	0.0608	0.0851	0.0637	**0.0189**	**0.0136**	0.0472	0.0368
7	LSTM-relu	0.0689	0.0516	0.0804	0.0599	0.0239	0.0183	0.0551	0.0430
8	CNN	0.0665	0.0495	0.0796	0.0597	0.0197	0.0142	0.0445	0.0342
9	Bayesian ANN	0.0481	0.0417	0.0438	0.0405	0.0198	0.0154	0.0564	0.0445
10	Bayesian BLSTM	**0.0251**	**0.0281**	**0.0305**	**0.0260**	**0.0168**	**0.0126**	**0.0167**	**0.0121**

Bold values represent the best performing methods in the form of least errors for each case of dataset.

TABLE 4 Pinball score by PDL forecasting methods

Sr. no.	Method	Victorian (2 years)	Victorian (1 year)	AEP (10 years)	AEP (1 year)
1	Bayesian ANN	0.0019	0.0210	0.0050	0.0150
2	Bayesian BLSTM	0.0016	0.0180	0.0030	0.0070

TABLE 5 Winkler score by PDL forecasting methods

Sr. no.	Method	Victorian (2 years)	Victorian (1 year)	AEP (10 years)	AEP (1 year)
1	Bayesian ANN	0.7264	0.7834	0.1273	0.2341
2	Bayesian BLSTM	0.5432	0.6865	0.0703	0.0902

However, for non-stationary and smaller training samples, Bayesian BLSTM provides the best results by achieving the least error value, as reflected in Table 3. It should be noted that from point forecasting methods, RNN with Relu gives second best results in terms of lower error values. In addition, Bayesians provides PIs for future predictions and forecasts with more accuracy in relation to the ground truth, as reflected inFigure 6. RNN and LSTM have less error as compared to the statistical ARIMA and traditional ML methods such as SVR.For AEP dataset, LSTM is performing better with larger samples. For smaller sample size, RNN demonstrates less error and computational cost. Models with tanh function perform better with uniform data, such as the AEP dataset. However, with higher variability and seasonal trends, that is, for the Victorian dataset, Relu performs better as it sparsely activates neurons rather than activating all neurons at the same

International Journal of Electrical and Electronics Engineering Studies, 10 (1), 1-18, 2024

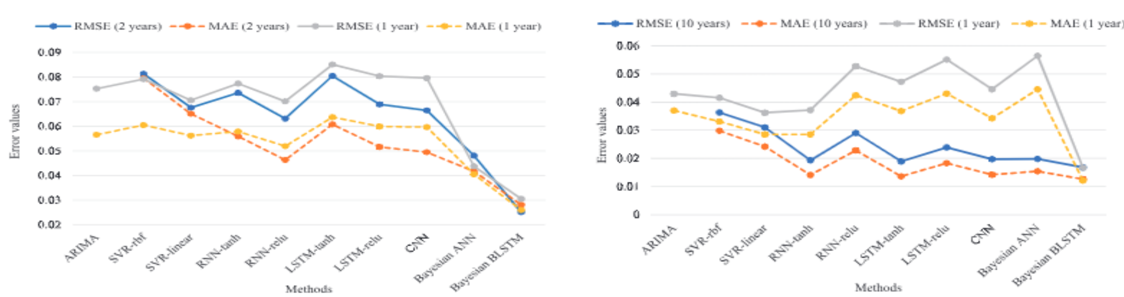
ISSN 2056-581X (Print),

ISSN 2056-5828(Online)

Website: https://www.eajournals.org/

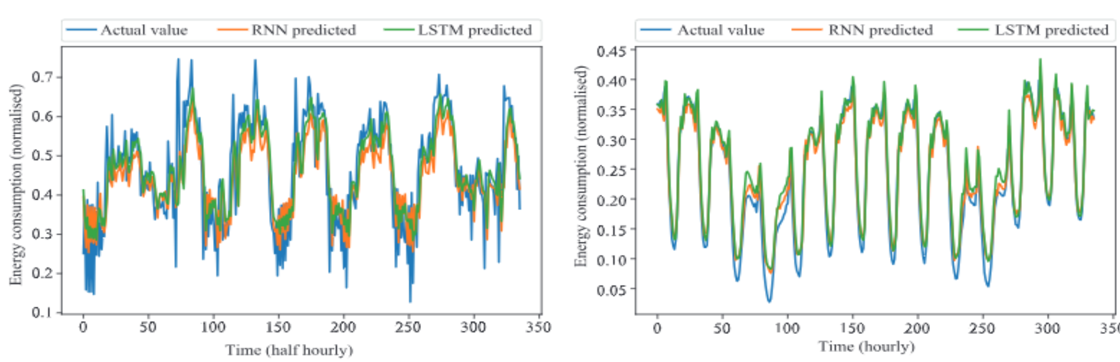
Publication of the European Centre for Research Training and Development -UK

time. Furthermore, datasets with more samples tend to train better and yield greater accuracy. In addition, ARIMA takes a longer time to train for larger datasets and thus, fails to generate the output. SVR with linear kernel has less error compared to SVR with rbf kernel. When dataset has linearly separable features, linear kernel will be sufficient to achieve superior performance. Tables 4 and 5 display the Pinball and Winkler scores, respectively, given by Bayesian ANN and Bayesian BLSTM for different data sizes. It can be inferred from the values that in case of more data and hence larger training set, scores are lower indicating the improvement in accuracy. Also, Bayesian BLSTM outperforms Bayesian ANN with the help of extensive training capability of bidirectional layer. Similarly, the BS has been computed for the two probabilistic methods across the different dataset sizes in Table 6. The performance comparison shows that Bayesian BLSTM achieves a lower BS for all considered scenarios. Figure 4 reflects the comparison graphs of errors observed by forecasting methods considered in Table 3 with respect to each sample size. It can be inferred from the figures that for point forecasting, standard DL methods provide least error irrespective of the data size and variability. However for PDL forecasting, Bayesian BLSTM outperforms all the other comparative methods and generate PIs. Furthermore, Figure 5 shows the actual and predicted energy consumption values using RNN and LSTM on each dataset. The results are plotted for activation functions which pro-duce superior performance for a given algorithm. Normalized energy consumption values are plotted over half hourly and per hour intervals for Victorian and AEP dataset, respectively.



(a) Victorian dataset for different forecasting methods (b) AEP dataset for different forecasting methods

FIGURE 4 Comparative analysis of forecasting methods based on evaluation metrics for different datasets with different sample sizes



(a) Victorian energy consumption dataset (b) AEP dataset

FIGURE 5 Predicted versus actual values for RNN and LSTM with one week and two weeks of test data for the Victorian and AEP dataset, respectively

From Figure 5a, it is inferred that the Victorian dataset has higher variability leading to more uncertainty in the energy consumption and hence, the performance of the models is slightly poorer. To deal with its non-stationary nature, logarithmic trans-formation is adapted. However, in Figure 5b, AEP dataset is uniformly distributed for which predicted values show greater similarity with actual values and thus, less errors. Further-more, Figure 6 represents a comparison graph of actual energy consumption values (ground truth) with predictive mean over90% and 50% prediction PIs generated by Bayesian BLSTM on both the datasets. PIs reflect the intervals for future probabilities on different percentiles which enable the PDL methods to quantify for uncertainties. Figure 6a clearly shows wider PIs as compared to Figure 6b as there is more variability in Victorian consumption values than AEP. In addition, a comparative plot for the execution time taken by each method is presented in Figure 7. Note that, the reported execution time precisely belongs to the training process of each forecasting method. Time taken for error computation and parameter tuning is not included. As shown in the figure, PDL methods take more time compared to point forecasting meth-ods. There is tradeoff involved between the time complexity and forecasting accuracy for the PDL methods, which needs to be maintained.

CONCLUSION AND FUTURE DIRECTIONS

In this paper, a comprehensive review of classical and advanced forecasting methods is presented for modern energy systems. A number of statistical, AI-based, probabilistic, and hybrid meth-ods are discussed in detail with respect to their applications in energy systems. In addition, impact of data pre-preprocessing techniques on forecasting performance is also highlighted. After conducting a comparative case study on two different data sets, it is inferred that DL methods with appropriate activation function and hyper-parameter tuning yield higher forecasting accuracy than the traditional statistical and ML methods. How-ever, uncertainty in the energy data from exogenous factors is a major challenge that can be tackled more efficiently with probabilistic methods. To support this, we implemented Bayesian ANN and Bayesian BLSTM as PDL forecasting techniques and the latter outperformed all the other comparative methods by demonstrating least error values. However, it exhibit high computational cost in the terms of time complexity and processing units. So, probabilistic forecasting techniques need to be explored more in the future work and potential solutions to reduce computational cost need to be proposed in the energy domain. Moreover, forecasting multi-dimensional demand and generation while ensuring high accuracy is another issue that requires substantial attention

REFERENCES

1. Gungor, V.C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., et al.:Smart grid technologies: Communication technologies and standards.IEEE Trans. Ind. Inf. 7(4), 529–539 (2011)

2. Çimen, H., Çetinkaya, N., Vasquez, J.C., Guerrero, J.M.: A microgridenergy management system based on non-intrusive load monitoring viamultitask learning. IEEE Trans. Smart Grid (2020)

3. Islam, S.N., Baig, Z., Zeadally, S.: Physical layer security for the smartgrid: Vulnerabilities, threats, and countermeasures. IEEE Trans. Ind. Inf.15(12), 6522–6530 (2019)

4. Agoua, X.G., Girard, R., Kariniotakis, G.: Short-term spatio-temporalforecasting of photovoltaic power production. IEEE Trans. SustainableEnergy 9(2), 538–546 (2017)

5. Jahangir, H., Tayarani, H., Gougheri, S.S., Golkar, M.A., Ahmadian, A.,Elkamel, A.: Deep learning-based forecasting approach in smart gridswith micro-clustering and bi-directional LSTM network. IEEE Trans.Ind. Electron. 68(9), 8298–8309 (2021)

6. Wang, Y., Chen, Q., Hong, T., Kang, C.: Review of smart meter data ana-lytics: Applications, methodologies, and challenges. IEEE Trans. SmartGrid 10(3), 3125–3148 (2018)

7. Kaur, D., Kumar, R., Kumar, N., Guizani, M.: Smart grid energy manage-ment using rnn-lstm: A deep learning-based approach. In: 2019 IEEEGlobal Communications Conference (GLOBECOM), pp. 1–6. IEEE,Piscataway (2019)

8. Kaur, D., Islam, S.N., Mahmud, M., et al.: A bayesian deep learning tech-nique for multi-step ahead solar generation forecasting. arXiv preprintarXiv:220311379 (2022)

9. Zhang, Y., Wang, J., Wang, X.: Review on probabilistic forecasting of windpower generation. Renew. Sustain. Energy Rev. 32, 255–270 (2014)

10. Liu, B., Nowotarski, J., Hong, T., Weron, R.: Probabilistic load forecastingvia quantile regression averaging on sister forecasts. IEEE Trans. SmartGrid 8(2), 730–737 (2015)

11. Ali, M., Prakash, K., Pota, H.: A bayesian approach based on acquisitionfunction for optimal selection of deep learning hyperparameters: A casestudy with energy management data. Sci. Proc. Ser. 2(1), 22–27 (2020)

12. Yang, Y., Li, W., Gulliver, T.A., Li, S.: Bayesian deep learning-based prob-abilistic load forecasting in smart grids. IEEE Trans. Ind. Inf. 16(7),4703–4713 (2019)

13. Torabi, M., Hashemi, S., Saybani, M.R., Shamshirband, S., Mosavi, A.: Ahybrid clustering and classification technique for forecasting short-termenergy consumption. Environ. Prog. Sustain. Energy 38(1), 66–76 (2018)

14. Kaur, D., Aujla, G.S., Kumar, N., Zomaya, A.Y., Perera, C., Ranjan, R.:Tensor-based big data management scheme for dimensionality reductionproblem in smart grid systems: Sdn perspective. IEEE Trans. Knowl.Data Eng. 30(10), 1985–1998 (2018)

15. Kaur, D., Islam, S.N., Mahmud, M., et al.: A vae-based bayesianbidirectional lstm for renewable energy forecasting. arXiv preprintarXiv:210312969 (2021)

16. Manjili, Y.S., Vega, R., Jamshidi, M.M.: Data-analytic-based adaptive solarenergy forecasting framework. IEEE Syst. J. 12(1), 285–296 (2017)

17. Huang, Y.C., Huang, C.M., Chen, S.J., Yang, S.P.: Optimization of mod-ule parameters for pv power estimation using a hybrid algorithm. IEEETrans. Sustain. Energy 11(4), 2210–2219 (2020)

18. de Souza, J.C.S., Assis, T.M.L., Pal, B.C.: Data compression in smart dis-tribution systems via singular value decomposition. IEEE Trans. SmartGrid 8(1), 275–284 (2015)

19. Ryu, S., Choi, H., Lee, H., Kim, H.: Convolutional autoencoder basedfeature extraction and clustering for customer load analysis. IEEE Trans.Power Syst. 35(2), 1048–1060 (2019)

20. Voyant, C., Notton, G., Kalogirou, S., Nivet, M.L., Paoli, C., Motte, F.,et al.: Machine learning methods for solar radiation forecasting: A review.Renewable Energy 105, 569–582 (2017)

21. Akhter, M.N., Mekhilef, S., Mokhlis, H., Shah, N.M.: Review on fore-casting of photovoltaic power generation based on machine learningand metaheuristic techniques. IET Renewable Power Gener. 13(7),1009–1023 (2019)

22. Kumar, D.S., Yagli, G.M., Kashyap, M., Srinivasan, D.: Solar irradianceresource and forecasting: a comprehensive review. IET Renewable PowerGener. 14(10), 1641–1656 (2020)

23. Wang, H., Lei, Z., Zhang, X., Zhou, B., Peng, J.: A review of deep learningfor renewable energy forecasting. Energy Convers. Manage. 198, 111799(2019)

24. Hernandez, L., Baladron, C., Aguiar, J.M., Carro, B., Sanchez.Esguevillas,A.J., Lloret, J., et al.: A survey on electric power demand forecasting:future trends in smart grids, microgrids and smart buildings. IEEECommun. Surv. Tutorials 16(3), 1460–1495 (2014)

25. Rossi, B., Chren, S.: Smart grids data analysis: A systematic mappingstudy. IEEE Trans. Ind. Inf. 16(6), 3619–3639 (2019)