Energy Reports 8 (2022) 13333–13352

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Research paper

Hyperparameter optimization of support vector machine using adaptive differential evolution for electricity load forecasting

M. Zulfiqar [a], M. Kamran [b], M.B. Rasheed [c,*], T. Alquthami [d], A.H. Milyani [d]

[a] Department of Telecommunication Systems, Bahauddin Zakariya University, Multan, Pakistan
[b] Department of Electrical Engineering, University of Engineering and Technology Lahore, Pakistan
[c] Universidad de Alcalá, Escuela Politécnica Superior, ISG, Alcalá de Henares, Spain
[d] Electrical and Computer Engineering Department, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Article history:
Received 10 May 2022
Received in revised form 22 August 2022
Accepted 28 September 2022
Available online 18 October 2022

Keywords:
Peak load forecasting
Support vector machine
Adaptive differential evolution
Multivariate empirical mode decomposition
Convergence rate

ABSTRACT

Peak load forecasting plays an integral part in the planning and operating of energy plants for the utility companies and policymakers to devise reliable and stable power infrastructure. However, the electricity load profile is considered a complex signal due to the non-linear and stochastic behavior of the consumer. Therefore, a rigid forecasting model with assertive stochastic and non-linear behavior capturing abilities is required to estimate the demand capacity accurately. To handle these uncertainties, this paper proposed a hybrid model that integrates the multivariate empirical modal decomposition (MEMD) and adaptive differential evolution (ADE) algorithm with a support vector machine (SVM). MEMD allows the decomposition of multivariate data to deteriorate over time to effectively extract the unique information from exogenous variables over different time frequencies to ensure high computational efficiency. The ADE algorithm obtains and tunes the SVM model's appropriate parameters to effectively avoid trapping into local optimum and returns accurate forecasting results. Consequently, the proposed MEMD-ADE-SVM forecasting model simultaneously achieves good accuracy (93.145%), stability, and convergence rate, respectively. A historical load dataset from the independent system operator (ISO) New England (ISO-NE) energy sector is analyzed to verify the MEMD-ADE-SVM hybrid model. The results show that the developed MEMD-ADE-SVM model outperforms the benchmark frameworks such as; SVR-based model by hybridizing variational mode decomposition, the chaotic mapping mechanism, and the grey wolf optimizer (VMD-SVR-CGWO), SVM based on data preprocessing and whale optimization algorithm (DCP-SVM-WO), intelligent optimized SVR model based on variational mode decomposition and Fast Fourier transform (VMD-FFT-IOSVR), SVR model based on multivariate empirical mode decomposition and particle swarm optimization (EMD-SVR-PSO), and MEMD-ADE-LSTM for day-ahead and week-ahead electricity peak load forecasting in terms of accuracy, stability, and convergence rate.

1. Introduction

A Smart grid (SG) has emerged as an intelligent energy system that has exclusively gained popularity due to its significance in peak load forecasting (PLF) (Yu et al., 2019). Numerous new studies are being conducted on PLF. However, we still need a more precise and resilient electric load forecasting (ELF) model. Accurate estimation of future electrical load fluctuations is essential for both utilities and consumers regarding decision-making and grid operations (Xiao et al., 2016; Jacob et al., 2020). However, the main obstacles to future PLF are various factors such as occupancy patterns, temperature, climate change, calendar indicators, humidity, and social practices. Appropriate mapping of these facets and load variability is significantly rebellious owing to the non-linear, arbitrary, and stochastic behavior of the load (Bashir et al., 2022). The versatility of communication technology, SG's sensor system, and advanced metering infrastructure (AMI) allow us to apprehend, examine and monitor the impact of these influential facets on PLF (Haq and Ni, 2019; Haq et al., 2020; Massaoudi et al., 2021). PLF provides essential information to policymakers to ensure efficient planning and optimal scheduling using unit commitment and economic dispatch (Talaat et al., 2020). PLF is a prediction of electric load using load history and weather information (Wood, 2022). Therefore, utilities strive to design accurate, resilient, fast, and simple STLF models. In addition, accurate predictions can help to identify potential disruptions and ensure reliable grid operations.

* Corresponding author.
E-mail address: Muhammad.rasheed@uah.es (M.B. Rasheed).

https://doi.org/10.1016/j.egyr.2022.09.188

Nomenclature

η	Grey rational coefficient
γ_i	Corresponding to r_i
$\mathcal{G}_{\vert}$	Grey rational grade
$\mathcal{R}$	Final residue
$\mathcal{Y}_{\mathfrak{q}}^{ml}$	Day ahead mean load
$\mathfrak{I}_{\mathfrak{mf}}^{\mathfrak{j}}$	Intrinsic mode function
$\mathfrak{K}$	Kernel function
$\mathfrak{P}_k$	Principal component
$\mathfrak{r}$	Initial residue
$\mathfrak{V}$	Covariance of matrix $\mathbb{R}$
$\mathfrak{Z}_j(k)$	Original sequence
b^*	randomly selected item
D	Difference between $\mathfrak{r}_1$ and $\mathfrak{r}_2$
ev	Eign value features
j	Number of trails
$\mathfrak{Z}\mathfrak{max}_j\ (k)$	Largest value of $\mathfrak{Z}_j^*(k)$
$\mathfrak{Z}\mathfrak{min}_j\ (k)$	Smallest value of $\mathfrak{Z}_j^*(k)$
α	Random forest evaluator
β	Relief-F evaluator
κ	Classification variable
λ	Eign value
$\mathcal{D}$	Dimensional space
$\mathcal{Y}$	Time series
$\mathcal{Y}_{\mathfrak{q}}^{\tau}$	Temperature load
$\mathcal{Y}_{\mathfrak{q}}^{pl}$	Day ahead peak load
$\mathcal{Y}_{\mathfrak{q}}^{\upsilon}$	Day ahead valley load
$\mathfrak{B}$	Trail vector
$\mathfrak{ff}\ ()$	Fitness function
$\mathfrak{f}^*$	Feature space
$\mathfrak{L}_j^{\mathsf{a}}$	Targeted load consumption pattern
$\mathfrak{M}_{en}$	Mean envelopes
$\mathfrak{r}_{\mathfrak{f}}$	Residue function
$\mathfrak{R}_f$	Regularized risk function
$\mathfrak{Z}_0^*\ (k)$	Reference sequence
$\mathfrak{Z}_j^*\ (k)$	Comparability sequence
$\mathfrak{Z}_j^*(k)$	Sequence after preprocessing
$\Sigma_{0j}\ (k)$	Grey rational coefficient
$\mathsf{W}_{\mathsf{g}(j)}$	White Gaussian noise series
ν	Distinguishing coefficient
ϕ	Map feature space & input data
σ	Feature selecting threshold
$\varkappa$	Insensitive loss function
$\wp_0$	Ratio of data versus noise
a	explanation for the abbreviation
C	Class
D	Number of directive
L	Time series length
$r\ [n]$	n number of decision trees
υ	Number of extreme
υ_j	Forecasting parameter
$y(t)$	Actual data-set
y^{ϕ_j}	Directive vector
$y_o(t)$	Pristine data-set
$\mathfrak{F}_{\mathrm{i}}{}^{f}[\tau_k]$	Feature importance by α
$\mathfrak{F}_{\mathrm{i}}{}^{r}[\tau_k]$	Feature importance by β
$M_t\ (j,\ k)$	Modified mutant vector
$P_t\ (j,\ k)$	Modified parent vector
$p_t\ (j,\ k)$	Parent vector

In current decades, the forecasting time frame is categorized into four formats: long-term (Guan et al., 2021; Kazemzadeh et al., 2020), medium-term (Niu et al., 2021; Ahmad and Chen, 2019), short-term (Massaoudi et al., 2021; Zhang et al., 2021), and ultra-short-term LF (Liu and Gao, 2020). STLF is generally accomplished one day to numerous weeks ahead of energy load (Chen et al., 2020). STLF needs accuracy for successful grid distribution decisions as compared to medium and long-term forecasts (Massaoudi et al., 2021). As a result, STLF prediction accuracy is becoming more and more important (Hernandez et al., 2014). Accurate STLF is important for the operation of power systems, such as energy planning, scheduling, and financial processes to assemble the system safer and more stable (Talaat et al., 2020). Therefore, this research concentrates on the day ahead and week ahead of PLF for short-term network optimization and unit commitment to faulty distribution systems.

Various researchers have proposed different PLF frameworks to enhance predictive performance. It is mainly categorized into univariate and multivariate substreams. The univariate method uses only time series, while the multivariate method evaluates various aspects such as gross domestic product (GDP), population growth (PG), and PLF meteorological data (Kazemzadeh et al., 2020). The behavior of load time series flows is generally complex. PLF complexity is caused by the non-linear and unsteady forms of load time series (Heydari et al., 2020). Hence, due to the complexity of the load signal, it is advantageous to develop a reliable prediction framework for PLF with high accuracy. Numerous frameworks have been developed for PLF. For instance, classical statistics approaches like grey forecasting (GF) (Liu and Tian, 2013), fuzzy logic (FL) (Coelho et al., 2016), linear regression (LR) (Nalcaci et al., 2019; Jacob et al., 2020), autoregressive integrated moving average (ARIMA) (Al-Musaylh et al., 2018a), and exponential smoothing (ES) (Deng et al., 2021), are considered for linear PLF trends. However, these methods cannot accurately forecast nonlinear and time-series signals because they do not capture large fluctuations in power demand, and they can only address the physical characteristics of electricity substances (Memarzadeh and Keynia, 2021). In addition, definitive predictive frameworks give inadequate results on days of unusual or special events, such as weekends and holidays (Chen et al., 2020). However, state-of-the-art computational models are good at predicting nonlinear time series. As a result, scientists are moving from traditional theoretical analytical approaches to artificial intelligence (AI) based on computational techniques using numerous techniques such as artificial neural networks (ANNs) (Massaoudi et al., 2021; Nalcaci et al., 2019) and support vector machines (SVMs) (Li et al., 2020; Zhang et al., 2021; Al-Musaylh et al., 2018b; Hafeez et al., 2019) to perform their tasks more accurately. The ML hybrid approaches are effective strategies for addressing non-linear, transient, and non-stationary features of peak loads evaluated by researchers after considering the inherent limitations of individual models and incomplete predictions of univariate time series.

A revolutionary optimization technique such as differential evolution (DE) has been portrayed as a distinctive evolutionary algorithm method, transcending prior approaches such as PSO and GA (Vesterstrom and Thomsen, 2004). DE has the benefit of being simple to execute and requiring fewer parameters. According to the survey, the DE algorithm is more stable, resilient, and robust. It produces more satisfactory results than different algorithms. The DE or modified DE or enhanced DE has been often used in the economic-environment dispatch problems of energy systems (Santos Coelho and Mariani, 2006; Yuan et al., 2008, 2009). It has yielded a more satisfactory solution than other evolutionary algorithms like the GA or PSO. Due to SVM's extensive theoretical foundations and inference capabilities (Li et al., 2020; Zhang et al., 2021; Al-Musaylh et al., 2018b; Hafeez et al., 2019), this study uses SVM as a modeler. Furthermore, we employed adaptive differential evolution (ADE) to optimize SVM's hyperparameters to improve forecast accuracy. However, it is worth stating that the priority of this study is to tackle the uncertainties in the load data and the effectiveness of the designed hybrid framework, not an approximation of the variant implementation using various optimization methods such as a genetic algorithm (GA) (Moazzami et al., 2013), fruit fly optimization (FFO) (Zhang et al., 2018), comprehensive learning particle swarm optimization (CLPSO) (Hu et al., 2014), and modified artificial bee colony (MABC) (Li et al., 2015) for fluctuations in energy demand.

This study defines the devised novel MEMD-based hybrid model as MEMD-ADE-SVM. In a few words, the multivariate channels are first inputted into the devised MEMD-ADE-SVM framework to be deteriorated simultaneously by employing the MEMD technique. After that, three hyper-parameters of SVM are optimized by ADE. Finally, ADE-based SVM is utilized to launch a model and anticipate each component extracted using the MEMD approach. The actual load dataset from the ISO New England (ISO-NE) energy sector is considered to justify the performance of the presented framework and other comparative frameworks. The prediction of each element is incorporated to acquire the final forecast. This analysis shows that the developed hybrid model accurately estimates the day ahead and week ahead PLF.

1.1. Real contributions

A robust hybrid day-ahead modeling framework, MEMD-ADE-SVM for PLF, has been devised to effectively assemble stochastic scheduling with multi-dimensional input variables to minimize the loss of underestimated or overestimated energy systems. Although several types of research have been performed to analyze the forecasting reliability and significance, no studies, evidence from Table 1, communed to PLF using MEMD for peak load data decomposition and SVM have been documented in the literature. The real contributions of this study are outlined as follows.

1. **A transition towards hybridization:** A new revitalizing MEMD-ADE-SVM framework is being developed that integrates MEMD and ADE algorithms with SVMs. MEMD allows multivariate data decomposition to efficiently capture unique information between related variables of different time frequencies during multivariate deterioration over time. The ADE system proactively selects and adjusts SVM hyperparameters to enhance prediction accuracy and stability while boosting the rate of convergence. The hybridization of the MEMD and ADE algorithms makes it possible to effectively implement SVM technology. The proposed MEMD system decomposes historical loads & meteorological variables simultaneously, unlike other decomposition methods. This dynamically manages the nonlinearity and non-stationary peak loads, efficiently retrieving various features from different levels of time frequencies related to accurate forecasting of the day ahead and week ahead peak load.
2. **Appropriate hyperparameters modification and adaption through ADE technique:** The SVM framework has three hyperparameters: the intense loss function (ε), the kernel function parameters, and the parameters that represent the trade-off between training errors and function flatness. These variables positively affect the stable performance of SVMs in PLF. Selecting and modifying these variables for accurate and consistent execution is challenging. The presented ADE algorithm is merged with the SVM model to address the problem of hyperparameters that are difficult to tune in the SVM model. By combining the ADE algorithm with the SVM model, optimal hyperparameter selection and adjustment are achieved.
3. **Innovative performance evaluation criteria:** Four typical statistical metrics (mean absolute percentage error (MAPE), directional accuracy (DA), root mean square error (RMSE), and R-squared (R^2). The four-evaluation criteria can be used as a baseline for energy system decision-making. The need for the evaluation criterion is to test the framework's effectiveness and confirm its applicability. Furthermore, two statistical test approaches the Analysis of Variance (ANOVA) (Xiong et al., 2014) and the Diebold–Mariano (DM) (Li et al., 2020) validate the MEMD-ADE-SVM hybrid model to other competing models in forecasting reliability and accuracy while identifying a significant difference in the testing datasets.

1.2. Design goals

The designed MEMD-ADE-SVM framework aims to signify day ahead and week ahead PLF efficiently and accurately. We must process the raw data, identify the appropriate features, and carefully tune the classifier to perform this. As a result, the metrics listed below are indispensable for the processing performance of our presented system.

- Accuracy and convergence rate: These are the core goals of our devised framework.
- Dimensional reduction rate: In this devised framework, the performance of MEMD influences the accuracy of classification directly.
- Time-efficiency: Applied in PLF, the framework should run fast.

1.3. Paper organization

The rest of the paper is organized as follows: The literature survey is discussed in Section 2. Section 3 explains the devised MEMD-ADE-SVM model and methods used in it. Section 4 explains the research formulation by describing the dataset description, performance metrics, and experimental implementation, while Sections 5 and 6 show the simulation results and discussion, respectively. Finally, Section 7 concludes this work by outlining limitations and potential future directions.

2. Literature survey

STLF usually covers the hourly forecasting period and is essential for grid decision-making. Statistical and ML models are generally used in the STLF literature. These models are split into two sub-streams to better understand the current STLF models: models with univariate time series load data and a model with multivariate time series load data.

Researchers note that hybrid models that use time–frequency analysis are considered promising because of the advantages of a data pre-filtering approach in extracting the unique characteristics of time-series data. Time–frequency studies determine the affinity between the most effective physical quantities (time and frequency domain). Hybrid systems extend the current paradigm of modeling decomposed ensembles by relying on the actual properties of time series data. This is a step beyond the above systems. Moazzami, Khodabakhshian (Moazzami et al., 2013) use

Table 1
Recent and relevant literature survey's brief summary considering Frameworks, Objectives, Limitations, Advantages, Performance metrics (Accuracy, Convergence rate, and Computational complexity).

STLF frameworks	Objectives	Time scale	Limitations	Advantages	Performance metrics		
					Accuracy	Convergence rate	Computational complexity
SVM-ANN with K-Medoids clustering (Haq et al., 2020)	Peak load forecasting	Short-term (30 mints)	Improved accuracy with complex framework	The framework has large complexity	High	Low	High
Weather information based ELF of a bulk power system (Kazemzadeh et al., 2020)	Accuracy improvement for effective performance of bulk power system	Daily	This framework is suitable and quite effective only for bulk power system	Performance increased by integrating exogenous parameters	High	Low	High
ANN-based forecasting framework (Heydari et al., 2020)	Accuracy improved by reducing RMSE	Daily	Accuracy achieved at the expense of convergence	Convergence is decreased due to sigmoid function and model complexity	High	Low	High
A big data approach for ELF (Hu et al., 2014)	Forecast accuracy improvement for scalable models	Daily	The framework has complex structure and slow convergence rate	Forecast accuracy is improved at the expense of high complexity.	High	Low	High
Intelligent model for ELF based on SVM and FFI algorithm Khalid and Javaid (2020)	Distribution energy generation forecasting and structure analysis	Daily	The framework is designed for short horizon of prediction	High accuracy and better generalization is achieved at the cost of framework complexity	High	Low	High
LSTM-RNN based LF (Amjady et al., 2010)	Accuracy improvement to facilitate the residential consumers	Hourly	Proposed framework improved forecasting accuracy	Accuracy is improved while convergence is compromised	High	Low	High
ANN based DE-PSO (Sakurai et al., 2019)	Day ahead forecasting considering outliers.	Daily	Accuracy increased with compromising convergence rate	Improved accuracy in comparison traditional PSO	High	Low	High

ANN's wavelet decomposition (WLD) and GA optimization using PLF univariate time series meteorological data to decompose the load time series, taking into account the low and high-frequency dimensions of ANN. By determining the detection target, the prediction accuracy has been improved for grasping complex features at different frequencies. However, pre-selecting basis function wavelets is also very complicated. In recent years, empirical modal decomposition (EMD) has gained increasing attention to overcome wavelet defects (Huang et al., 1998). For example, AlMusaylh, and Deo (Al-Musaylh et al., 2018b) used an improved empirical adaptive noise mode decomposition combined with SVR optimized by a two-phased PSO approach for PLF using univariate time series load data. Recently, EMD studies have developed theoretical treatments for bivariate (Rilling et al., 2007), trivariate (ur Rehman and Mandic, 2009), and multivariate (Rehman and Mandic, 2010) patterns, respectively. Multivariate EMD (MEMD) is a multivariate, ingenious, adaptable, multiscale decomposition approach stemming from EMD (Huang et al., 1998; Rehman and Mandic, 2010). EMD and MEMD reinforce non-linear, inconsistent, uncertain, unstable, unsteady networks that can decompose the original data into intrinsic mode functions (IMFs). However, as mentioned above, various aspects such as environment variables, days of the week, holidays, and consumer social interactions affect the energy load (Hu et al., 2015b,a). The complexity of the energy system suggests that it is not enough to consider the past load of the univariate PLF. This study deals with nonlinear and transient load sequences and their influencing factors, considering the usefulness of MEMD. However, MEMD was proposed by Rehman and Mandic to decompose multivariate load series simultaneously for more accurate predictions (Rehman and Mandic, 2010). The results show that MEMD has improved performance compared to EMD. Nevertheless, it is difficult to estimate all the aspects that influence load. Temperature is a meteorological factor, based on comprehensive literature research (Hu et al., 2015a; Sobhani et al., 2020; Selakov et al., 2014; Jang et al., 2020). In this study, historical data consisting of load and temperature are considered as input variables to reduce the computational cost of the hybrid models. The MEMD algorithm decomposed multidimensional input variables to extract IMFs with similar frequencies and multidimensional residuals. PLF is the output of the predictive target. Researchers consider ML hybrid approaches are potent techniques that deal with non-stationary, non-linear, and transient characteristics of peak loads due to the inadequate predictive control of univariate time series and the inherent limitations of individual models. Table 1 shows various short-term PLF models.

The emerging hybrid and integrated predictive models are intelligent solutions that completely utilize the preferred features of single models to provide excellent efficiency (Raza and Khosravi, 2015; Yu et al., 2019). For example, a hybrid framework is devised based on a wavelet neural network (WNN) and improved differential evolution (IDE) for PLF (Liao, 2014). The applicability of the presented framework is confirmed by a practical comparing it with other frameworks like ANN-GA, ANN evolutionary programming (ANN-EP), and ANN-PSO. The authors developed a converged model of repulsion PSO (RPSO) and SVM algorithm for PLF (Dai and Zhao, 2020). The presented composite framework is validated using Singapore's historical data compared to traditional training frameworks in assessing accuracy. A nonlinear AR, GA, and an extraneous NN hybrid framework are devised for STLF (Jawad et al., 2018). The proposed framework is being optimized by using statistics and pattern recognition-based schemes. GA is used for the weight and bias of the NN training selection. The framework is validated by comparing existing mean and regression tree models with external inputs. The author presented a resilient STLF model with a computerized data preprocessing and prefiltering strategy for ELF of distribution feeders (Huyghues-Beaufond et al., 2020). The previous day's building level LF model was proposed based on DL (Cai et al., 2019). The devised DL model is validated by hovering the accuracy of definitive models. The navel hybrid model VMD-LSTM-BO has been developed (He et al., 2019). This model is considered superior to current models in stability and accuracy. In Wu et al. (2019), authors proposed GNRR and the multi-purpose cuckoo search algorithm (CSA) based improved hybrid framework. The developed model used real-time load data from an Australian energy market operator (AEMO) to evaluate forecast accuracy against the benchmark frameworks. The author developed a Neural Elman (NE) network-based forecast engine to signify the future load of SG. The proposed intelligent optimization algorithm optimally acclimates the biases and weights of this network to obtain precise forecastings (Vrablecová et al., 2018). The authors provided an STLF model based on SSVR (Li et al., 2018). The foremost objective is to enhance the efficacy and accuracy of comparative forecasts. The output of the forecasting module is passed through the optimization engine, and the accuracy and efficiency of relative predictions are enhanced by fine-tuning the parameters. However, it improves forecast accuracy at the outlay of computational complexity. The developed framework attains more increased accuracy than other frameworks. Besides, an integrated SVR algorithm and chaotic krill herd (CKH) framework is proposed for load forecasting time series (Zhang et al., 2020). However, the results acquired are not stable and have no accuracy. The hybrid framework is designed based on SVR and DE to improve prediction accuracy by modifying the SVR's hyper-parameters (Zhang et al., 2016). The proposed model performs better than typical regression models, SVR, ANN, and back-propagation (BP)- ANN models. A model combining fruit fly (Ff) and SVR was developed in Cao and Wu (2016) to solve the problem of parameter selection and improve the accuracy of PLF. Also, a new method is developed, hybridizing SVR with the firefly optimization (FFO) in Kavousi-Fard et al. (2014), Xiao et al. (2016) to ensure accurate PLF by optimal tuning hyper-parameters.

The aforementioned hybrid frameworks can be deemed promising, optimistic, and practical in enhancing forecast accuracy by suitably modifying super-parameters. However, the authors of these articles concentrate on optimizing bias initialization and random weight or appropriately altering and picking hyper-parameters. Also, none of these models considered accuracy, rate of convergence, and stability simultaneously. From plenty of analysis and investigation, we inferred that only one factor (bias initialization and random weight optimization or appropriate hyperparameters setting and selection) and only one criterion (convergence rate or accuracy or stability) are insufficient. Therefore, a robust hybrid model is needed to overwhelm the problems of current models while improving predictive accuracy and stability with fast convergence rates.

From the recent rational work above, we can draw three conclusions: (i) There is no perfect versatile predictive model in all respects, but some frameworks are appropriate for some objectives and conditions. (ii) There are problems with overfitting. Overfitting means that the model is above average in training and worst in prediction. (iii) There is a trade-off between the prediction accuracy and the convergence rate, and increasing prediction accuracy affects the convergence rate. The reverse is also true.

Several studies have compared forecasting models and identified the top-performing models for electricity PLF. However, the prevalence of these has targeted distinctive areas. No analysis involving definitive or recently developed hybrid models has exhaustively concentrated on the PLF. Hence, this research can present complete and long-term trends in procedure and policy than previous research via the relative analysis of diverse high-level forecasting frameworks. It can be confidently inferred from the literature that vast improvement has been made in ELF for energy management. However, the existent techniques are counterproductive in dealing with big data. It is formidable to tune control parameters, resulting in high computational complexity and the ineptitude to converge fast because redundancy, irrelevancy, and dimensional reduction are not intercepted. Moreover, the literature mentioned above does not simultaneously cater to forecast accuracy and convergence rate. A fast and accurate framework is the need of the day to unravel such concerns. Hence, in Shiri et al. (2015), the integration of gradient descent (GD) algorithm with an SVM-based framework is devised. This framework has much computational complexity and is untrained to converge. Some authors focus on feature selection algorithms such as traditional classifier decision trees (DTs) and artificial neural networks (ANN) (Jiang et al., 2016). However, DTs meet the issue of overfitting, which means that a DT portrays excellently in training but incorrectly in prediction. The ANN has restricted generalization capacity and has a dilemma of regulating its convergence. In Wang et al. (2017), the authors framed a hybrid feature selection, extraction, and classification-based model for ELF. However, this method has high system complexity and is incompetent to converge. Therefore, a new hybrid forecasting model has been devised. The developed framework strives to establish high-quality daily to weekly load forecasts over a period with a relatively high convergence rate for SG decision-making.

3. Proposed MEMD-ADE-SVM framework

This work proposes a novel hybrid framework based on MEMD, SVM, and ADE algorithms for day ahead and week ahead PLF. The devised model has three main parts: (i) MEMD based data decomposition module, (ii) forecaster based on SVM, and (iii) optimizer based on ADE algorithm.

The whole workflow of the devised MEMD-ADE-SVM framework is presented in Fig. 1. In this study, there are three key goals for PLF:

- The first is simultaneous preprocessing of multidimensional time series using a time frequency process. This can significantly improve future peak loads by adopting a univariate approach.
- Second, it avoids overfitting and training the developed model with excellent generalization capabilities and diminishes modeling and predictive computational complexity.
- Finally, we need to consider the optimal modeling parameters to ensure the optimal solution for our PLF.

The following are the relevant analyses of the proposed ADE-SVM based framework, as illustrated in Fig. 1:

Step 1: Using the MEMD, the multivariate channel $\mathcal{Y} = \{$ temperature, valley, mean, peak $\}$ is initially split into m multivariate $\mathfrak{r}_f$ and $\Im_{mf}$ extracted. For accuracy excellency, time domain studies can take into account the possible characteristics of factors that utilize energy load patterns. In contrast, frequency-domain analysis allows us to search more precisely for features specific to our dataset. The Fig. 3 shows a time-domain study applied to a real-time dataset to capture trends in power load. When implementing MEMD and the time–frequency method to deteriorate time-domain signals to finite waves of different frequencies, these spontaneous signals become stable and predictable signals of different frequencies instead of related variables. Effectively extract information and make predictions.

Step 2: ADE is used to optimize the hyper-parameters of SVM. MAPE carefully tunes three SVM parameters, notably (C, ϵ, γ), and, in training sets to ensure SVM predicting accuracy. Before analyzing and predicting each constituent with SVM, the characteristics of $\Im_{mf}$ and $\mathfrak{r}_f$ derived from the original signal $\mathcal{Y}$ are

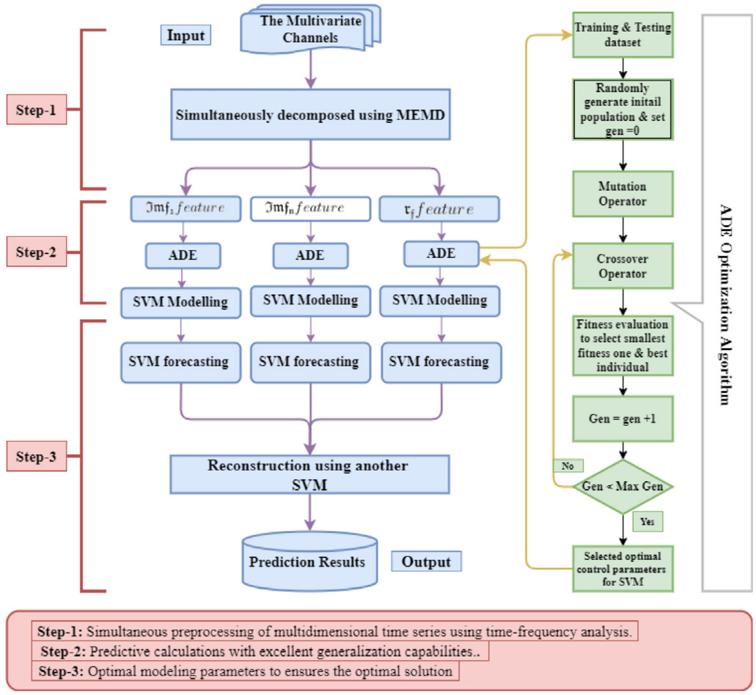


Fig. 1. Proposed system model MEMD-ADE-SVM.

investigated to find the best SVM settings. The developed ADE is a random search heuristic method based on group differences. It is an innovative and efficient technique. GA was used to develop the ADE. The function of ADE algorithm is to optimize the SVM parameters, that is, to identify the optimal parameter combination (C, ϵ, γ), so that the SVM model performs the best in classification. The SVM prediction model uses the SVR parameters that correspond to the best global solution.

Step 3: SVM is used in this study to build the system and forecast each element retrieved using the MEMD approach. ADE explored the optimum SVM hyper-parameters. However, the appropriate enhanced prediction of each removed element in *step 1* may be developed, which must then be revamped to obtain the final energy PL estimate.

In a nutshell, multivariate channels are input to the initially proposed MEMD-ADE-SVM model and simultaneously deteriorate using MEMD technology. Next, ADE optimizes the three hyperparameters of the SVM. Finally, use ADE-based SVM to build the framework and anticipate each feature. Each feature was extracted using the MEMD method. Combine the predictions from each feature to get the final prediction.

3.1. MEMD based decomposition module

The data lacks consistency, errors, preliminary numbers, and no evident behaviors. Before employing the data, we must depart the data through the decomposition stage to prepare the data and convert it into an understandable layout.

MEMD (Rehman and Mandic, 2010) is a generic extension of EMD (Huang et al., 1998). It is widely employed in a modular form to evaluate multi-channel time series data. MEMD translates historical peak load and influencing variables data into $\Im_{mf}$ in an efficient manner. The various layers of the $\Im_{mf}$ have a frequency range. The first decomposed segment obtains the maximum wavelength. As the number of decomposition rises, the frequency reduces considerably. The residue function $(\mathfrak{r}_f)$ is the final component of this computation. Neighborhood maximum and minimum are not well defined (Rehman and Mandic, 2010). While the conventional EMD approach can reconstruct a convoluted univariate load to a finite set of $\Im_{mf}$ and a $\mathfrak{r}_f$. The $\Im_{mf}$ components extracted by EMD from various. Different load TS may extract a different number of $\Im_{mf}$ elements, and TS may not always correlate to a certain frequency. In computing cost, matching the $\Im_{mf}$ differences obtained from different TS is tough. To overwhelm the inherent weaknesses of EMD, the MEMD method is applied in this study to improve forecasting accuracy while significantly lowering the computational cost. MEMD makes a significant contribution by estimating the local mean of n-dimensional signals.

In this work, the four input variates for day ahead and week ahead PLF are considered as input variables: the mean peak load, valley load, temperature, and peak load. The proposed MEMD can simultaneously decomposed the $\mathfrak{p}$-variate inputs $\mathcal{Y}$ into k-multivariate IMF $(\Im_{mf}(\mathfrak{q}))$ and a multivariate residue $(r_k(\mathfrak{q}))$. The $\mathcal{Y}$ is presented in Eq. (1):

$$\mathcal{Y} = \left\{\mathcal{Y}_1(\mathfrak{q}), \mathcal{Y}_2(\mathfrak{q}), \ldots, \mathcal{Y}_{\mathfrak{p}}(\mathfrak{q})\right\} \tag{1}$$

where each component $\Im_{\mathrm{mf}_j}(\mathfrak{q})$ presents $\left\{\Im_{\mathrm{mf}_j^1}(\mathfrak{q}), \Im_{\mathrm{mf}_j^2}(\mathfrak{q}), \ldots, \Im_{\mathrm{mf}_j^m}(\mathfrak{q})\right\}$ $(j = 1, \ldots, k)$, while the residue $r_k(\mathfrak{q})$ represents $\left\{{r_k}^1(\mathfrak{q}), {r_k}^2(\mathfrak{q}), \ldots, {r_k}^{\mathfrak{p}}(\mathfrak{q})\right\}$ of L length times series.

Fig. 2 shows the decomposition process using the proposed MEMD and its briefly description is as under:

1. Firstly, The Hammersley function is used to generate the appropriate point set for the given multivariate input signals described in Eq. (1).
2. The corresponding angles with the normalized Hammersley sequences are normalized in the range of 0 and 2π. The directive vector set (y^{ϕ_j}) presented in Eq. (2) and corresponding to angles (ϕ^d) presented in Eq. (3) are established in order to compute the projection $\left\{\mathrm{P}^{\phi_d}(\mathfrak{q})\right\}_{d=1}^{D}$ of $\mathcal{Y}$ along dth directions.

$$y^{\phi_j} = \left\{y_1^d, y_2^d, \ldots, y_m^d\right\}, d = 1, \ldots, D \tag{2}$$

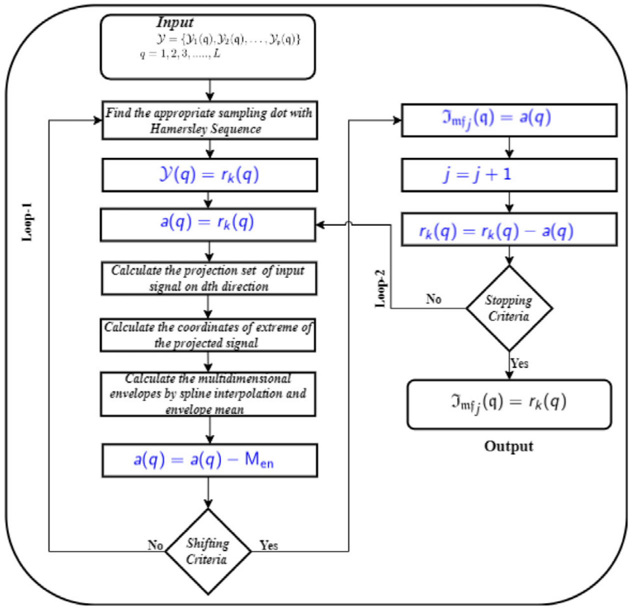


Fig. 2. MEMD algorithm flow chart. Loop-1 is used to test through shifting criterion while loop-2 is used to determine through stopping criterion.

where D is the number of directions.

$$\phi^d = \left\{\phi_1^d, \phi_2^d, \ldots, \phi_{m-1}^d\right\} \tag{3}$$

3. The extreme of $\left\{\mathrm{P}^{\phi_d}(\mathfrak{q})\right\}_{d=1}^{D}$ is evaluated on the instantaneous time $\left\{\mathfrak{q}_\ell^{\phi_d}\right\}_{d=1}^{D}$, where $\ell \in [1, L]$ denotes the position of the extreme. Furthermore, the coordinates $\left(\mathfrak{q}_\ell^{\phi_d}, \mathcal{Y}\left(\mathfrak{q}_\ell^{\phi_d}\right)\right)$ of the extreme dots are computed.
4. To obtain multi-directional wrapped curves on $\left(\mathfrak{q}_\ell^{\phi_d}, \mathcal{Y}\left(\mathfrak{q}_\ell^{\phi_d}\right)\right)$ by performing spline interpolation strategy. The mean of wrapped curves is calculated in Eq. (4):

$$\mathsf{M}_{en}(\mathfrak{q}) = \frac{1}{D-\upsilon} \sum_{d=1}^{D-\upsilon} exp^{\phi_d}(\mathfrak{q}) \tag{4}$$

where υ denotes the number of extrema of the projected signal less than 3. The Eq. (4) shows that a mode would be excluded when the projected signal has inadequate extrema.
5. The components $\mathfrak{I}_{mf}(q)$ are extracted in sequence from high to low frequency. As shown in Fig. 2. Loop 1 is used to test whether transfer function $a(q)$ is obtained through the sifting criterion or again calculation of projection of input signals along dth directions, and Loop 2 is used to determine whether $r(q)$ is the residue or is used to obtain next multivariate $\mathfrak{I}_{mf}1(q)$ through the stopping criterion.

Through the above procedures of decomposition, the multidirectional signal can be expressed as. The multidimensional signal $\{\mathcal{Y}(\mathfrak{q})\}_{\mathfrak{q}=1}^{L}$ can be expressed in Eq. (5) using the decomposition mechanisms described above:

$$\mathcal{Y}(\mathfrak{q}) = \sum_{j=1}^{m} \mathfrak{I}_{mf_j}(\mathfrak{q}) + r_k(\mathfrak{q}) \tag{5}$$

where m is the total number of $\mathfrak{I}_{mf}$ obtained $j = 1, \ldots$, m and the residue is r_k.

The day ahead multivariate historical data comprising electricity load and temperature are taken from ISO-NE for PLF work investigation to highlight the MEMD results for $\{\mathcal{Y}(\mathfrak{q})\}$ as depicted in Table 2. Consider $\mathcal{Y}_\mathfrak{q}^{pl}$, $\mathcal{Y}_\mathfrak{q}^{ml}$, $\mathcal{Y}_\mathfrak{q}^{\tau}$ and $\mathcal{Y}_\mathfrak{q}^{V}$, are day ahead peak load, mean load, temperature, and valley load respectively. The MEMD is capable of decomposing the four-variable inputs $\mathcal{Y}(\mathfrak{q})$ concurrently into m multivariate time series of $\mathfrak{I}_{mf}$ and $\mathfrak{r}_f(\mathfrak{q})$ is represented in Eq. (6):

Table 2
Historical data comprising of temperature and energy load taken from ISO-NE.

Date	Mean	Valley	Temperature	Peak
2020/2/1	8234.12	5893.21	25.2	8698.41
2020/2/2	8123.34	6010.27	23.9	9467.81
2020/2/3	7869.27	6123.11	22.9	8321.50
2020/2/4	7812.65	5867.01	21.8	8723.01
2020/2/5	8134.11	6254.21	23.7	9871.01
…	…	…	…	…
…	…	…	…	…

$$\mathcal{Y}(\mathfrak{q}) = \left\{\mathcal{Y}^{pl}(\mathfrak{q}), \mathcal{Y}^{V}(\mathfrak{q}), \mathcal{Y}^{ml}(\mathfrak{q}), \mathcal{Y}^{\tau}(\mathfrak{q})\right\}, (\mathfrak{q} = 1, \ldots, L) \tag{6}$$

where each component $\mathfrak{I}_{mf_j}(\mathfrak{q})$ of the L represents $\left\{\mathfrak{I}_{mf_j^{pl}}(\mathfrak{q}), \mathfrak{I}_{mf_j^{ml}}(\mathfrak{q}), \mathfrak{I}_{mf_j^{ml}}(\mathfrak{q}), \mathfrak{I}_{mf_j^{\tau}}(\mathfrak{q})\right\}$ $(j = 1, \ldots, m)$ while the residue r_k of L is represents $\left\{{r_k}^{pl}(\mathfrak{q}), {r_k}^{V}(\mathfrak{q}), {r_k}^{ml}(\mathfrak{q}), {r_k}^{\tau}(\mathfrak{q})\right\}$. The result of decomposition is shown in Fig. 3. Fig. 3 shows the decomposition of ISO-NE load data using the MEMD algorithm resulting in ten multivariate $\mathfrak{I}_{mf}$ elements with higher to the lowest frequency and one $\mathfrak{r}_f$ element.

3.2. SVM based forecaster

SVMs are generally used to demonstrate the problems of nonlinear and unsteady prediction. Due to its robust fictitious framework and potential for effective generalization, SVM is a modeler that confirms the efficacy of the MEMD-based time–frequency strategy presented to the day ahead and the week ahead PLF. In this paper, We model the classification problem mathematically in Eq. (7):

$$\mathsf{f}(x, y) = \sum_{k=1}^{\mathcal{D}} y_k \mathfrak{z}_k(x) + \gamma \tag{7}$$

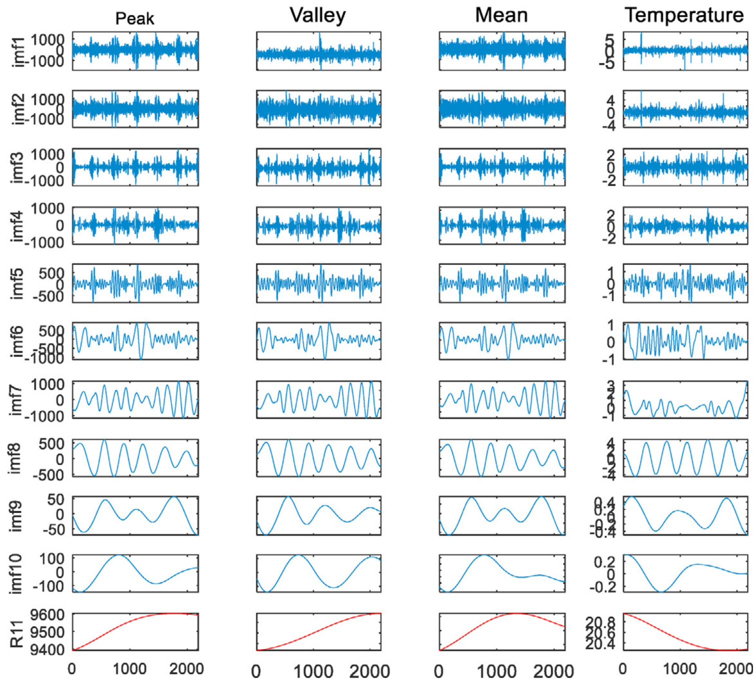


Fig. 3. Decomposition of ISO-NE load data using the MEMD algorithm resulting in ten multivariate $\mathfrak{I}_{mf}$ elements with higher to the lowest frequency and one $\mathfrak{r}_f$ element.

where $y_k^{\infty}(k = 1, 2, 3, \ldots)$ are the forecaster parameters to be determined, the dimensional space is represented by $\mathcal{D}$, and γ depends upon the distribution of data and classification variables. The objective of SVM is to define a hyperplane in $\mathcal{D}$-dimensional feature space that differentiates the data points. In this study, the hyperplane is defined by Eq. (7). The regularized risk function $\mathcal{R}_F$ is then defined in Eq. (8):

$$\mathcal{R}_F\,(y) = \frac{\sum_{k=1}^{\mathcal{D}} \left|\mathfrak{L}_k^{a} - \mathsf{f}\,(x, y)\right|_{\varepsilon} + \sigma y^2}{\mathcal{D}}, \tag{8}$$

where σ is the feature selection regulating threshold, ε is the insensitive loss function parameter, and $\mathfrak{L}_k^{a}$ is the targeted load consumption pattern. The parameter y must be obtained through minimization of this $\mathcal{R}_F$. The robust error function x is calculated in Eq. (9):

$$x = \begin{cases} 0 & \mathit{if}\ \left|\mathfrak{L}_j^{a} - \mathsf{f}\,(x, y)\right| < \varepsilon \\ \left|\mathfrak{L}_j^{a} - \mathsf{f}\,(x, y)\right| & \mathit{otherwise}. \end{cases} \tag{9}$$

Eq. (9) employs a function to minimize Eq. (8) and can be modeled in Eq. (10):

$$\mathsf{f}\left(x, \pi, \pi^*\right) = \sum_{k=1}^{N} \left(\pi^* - \pi\right)\mathcal{K}^*\,(x, x_k) + \gamma, \tag{10}$$

where $\pi^* \geq 0$ for all values of k. $\mathcal{K}^*\,(x, w)$ is the SVM kernel function that shows the multiplication of radial basis KPCA in the feature space f^* as in Eq. (11):

$$\mathcal{K}^*\,(x, w) = \sum_{k=1}^{\mathcal{D}} \mathfrak{Z}_k\,(x)\mathfrak{Z}_k\,(w) \tag{11}$$

In an infinite feature space, the $\mathcal{K}^*$ eliminates the requirement for $\mathfrak{Z}_k$ feature will be calculated. By maximizing the quadratic form, the π and π^* can be obtained in Eq. (12):

$$\mathcal{R}\left(\pi^*, \pi\right) = -\varepsilon \sum_{k=1}^{N} \left(\pi_k^* + \pi_k\right) + \sum_{k=1}^{N} \mathfrak{L}_k^{a}\left(\pi_k^* - \pi_j\right) - \frac{1}{2}\sum_{k,i=1}^{N} \left(\pi_k^* + \pi_k\right)\left(\pi_k^* - \pi_k\right)\mathcal{K}^*\left(x_k, w_k\right). \tag{12}$$

The ELF pattern is fed into the optimizer unit, which improves accuracy by dropping other errors.

3.3. ADE

This subsystem intends to improve predictive performance even further by dropping the $\mathcal{R}_F$. Since the SVM-based forecaster's returned value of the $\mathcal{R}_F$ is the smallest within its limits, The optimizer unit is combined with the forecasting module based on SVM to minimize further $\mathcal{R}_F$. However, as an objective function, the optimization module applies $\mathcal{R}_F$ minimization. Yet, this feature is guided to hyperparameters including the insensitive loss function $\mathfrak{L}_f$, cost penalty $\mathfrak{C}_{\mathfrak{P}}$, and kernel $\mathcal{K}$. Optimizing these hyperparameters is robust for efficient, accurate, and effective LF. Scholars have used a variety of methods to improve hyperparameters, including cross-validation, back-propagation (BP), and gradient descent (GD) (Kumar et al., 2016). On the other hand, these strategies have high dimensionality and are untrained to converge. However, DE is favored over-optimization method for two rationales: (i) premature convergence avoidance, and (ii) it provides superior quest ability. The authors used an efficient adaptation of DE (EDE) that was presented in Storn and Price (1997). In terms of trial vector generation reliability and convergence rate, the study in Amjady et al. (2010) is enhanced. As a result, proposed adaptive DE is used with the SVM model to optimize control parameters. The following is a brief discussion: The trial vector ($\mathfrak{B}$) for the jth individual in the t iteration is described in Amjady et al. (2010) and presented in Eq. (13):

$$\mathfrak{B}_t\,(j,\ k) = \begin{cases} m_t\,(j,\ k) & \mathit{if}\ r_m(j) \leq \mathfrak{ff}\,(m_t(j)) \\ p_t\,(j,\ k) & \mathit{if}\ r_m(k) > \mathfrak{ff}\,(m_t(j)) \end{cases} \tag{13}$$

where $m_t\,(j,\ k)$ represents the mutant vector and $p_t\,(j,\ k)$ represents the parent vector. In Eq. (13), $\mathfrak{ff}\,()$ represents the fitness

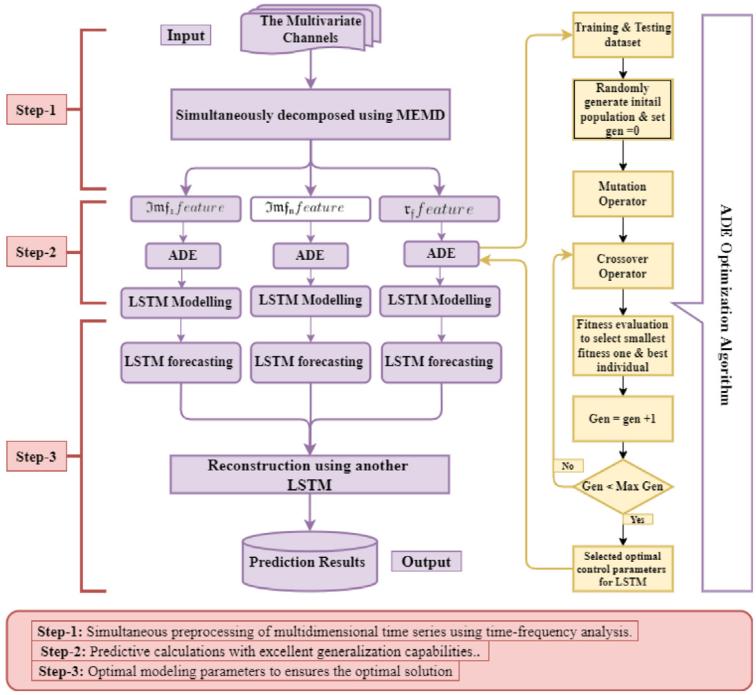


Fig. 4. Proposed benchmark model MEMD-ADE-LSTM for comparison purpose.

function with a range of 0 to 1, and r_m represents a random number with a range of 0 to 1. The next generation $\mathfrak{B}_{t+1}(j)$ offspring is generated based on $p_t(j)$ and $m_t(j)$ presented in Eq. (14):

$$\mathfrak{B}_{t+1}(j,\ k) = \begin{cases} \mathfrak{B}_t(j,\ k)\ \ if\ \Re_f(\mathfrak{B}_t(j)) \le \mathfrak{ff}(p_t(j)) \\ p_t(j,\ k)\ \ otherwise. \end{cases} \tag{14}$$

It is clear from Eqs. (13) and (14) that the former $\mathfrak{B}$ influences the selection of the following reproduction $t+1$ offspring, which is based on the r_m and $\mathfrak{ff}()$ functions. The EDE method in Amjady et al. (2010) compares r_m with $\mathfrak{ff}()$ to update load values. This erratic upgrading of the load is a significant dilemma. As a result, This issue is addressed by eliminating the offspring selection's reliance on the genuinely random quantity. The process for changing load values is established by comparing the future load value of $\mathfrak{ff}$ to the previous load value. As a result, the new load values will become suitable, increasing forecast accuracy. The designed adjustments in Eq. (13) are as follows:

$$\mathfrak{B}_t(j,\ k) = \begin{cases} m_t(j,\ k) & if\ \dfrac{P_t(j)}{P_t(j_{\max})} \le \mathfrak{ff}(M_t(j)) \\ p_t(j,\ k) & if\ \dfrac{P_t(j)}{P_t(j_{\max})} > \mathfrak{ff}(M_t(j)) \end{cases} \tag{15}$$

The $\mathfrak{ff}$ of parent and mutant vectors, according to this perspective, is defined in Amjady et al. (2010) and presented in Eq. (16) and (17):

$$\mathfrak{ff}(M_t(j)) = \frac{\frac{1}{\Re_f(M_t(j))}}{\frac{1}{\Re_f(M_t(j))} + \frac{1}{\Re_f(P_t(j))}} \tag{16}$$

$$\mathfrak{ff}(P_t(j)) = \frac{\frac{1}{\Re_f(P_t(j))}}{\frac{1}{\Re_f(P_t(j))} + \frac{1}{\Re_f(M_t(j))}} \tag{17}$$

It is assumed in the $\mathfrak{ff}$ of Eq. (16) and (17) that every arithmetic function, such as divide and adding, requires 1 unit of time to complete. Because Eq. (16) and (17) will require 5 units of time to execute each iteration, the total number of iterations for the EDE method is 100, according to Amjady et al. (2010). At each iteration, EDE calculates 1 $\mathfrak{ff}$ in 500 time units and 2 in 1000 time units. As a consequence, the $\mathfrak{ff}$s in Eq. (16) and (17) are modified to minimize processing time and boost generalization ability as follows:

$$\mathfrak{ff}(M_t(j)) = \frac{\Re_f(P_t(j))}{\Re_f(M_t(j)) + \Re_f(P_t(j))} \tag{18}$$

$$\mathfrak{ff}(P_t(j)) = \frac{\Re_f(M_t(j))}{\Re_f(P_t(j)) + \Re_f(M_t(j))} \tag{19}$$

Using Eqs. (14) and (15), the approach calculates 2 $\mathfrak{ff}$s in 100 iterations in 400 units. As a response, the convergence performance of the EDE method used in Amjady et al. (2010) is improved.

3.4. Proposed benchmark model MEMD-LSTM-ADE for comparison

The comprehensive flow of the devised benchmark MEMD-ADE-LSTM framework compared to the proposed model is elaborated in detail in Fig. 4. The explicit studies of the suggested LSTM-based benchmark model are as follows: LSTM is used to verify the framework and forecast each element extracted by employing the MEMD approach. ADE analyzed the optimal control parameters of LSTM. Hence, the affiliated improved forecast of each part removed from *Step 1* can be acquired, which requires reconstruction to fetch the final electricity PLF. In a few words, the multivariate channels are first inputted into the proposed MEMD-ADE-LSTM model to be decomposed simultaneously using the MEMD technique. Then hyper-parameters of LSTM are optimized by ADE. Finally, ADE-based LSTM establishes a model and forecasts each element extracted using the MEMD technique. The estimates of each component are incorporated to fetch the final prediction.

4. Research formation

4.1. Datasets description

A real-world load dataset from ISO-NE (ISO, 2022) is used to validate the excellency of the devised MEMD-ADE-SVM hybrid framework. The main reason for using the samples is that ISO-NE is an independent, not-for-profit corporation responsible for keeping electricity flowing across the six New England states. The first $2by3$ of daily historical data is used as training samples. At the same time, the rest is used to validate and test a model to estimate the PL. Table 3 shows the experimental distribution of the data sets. For the multi-dimensional actual series, consider valley load, peak load, mean load and temperature represented as y^{Vl}, y^{Pl}, y^{Ml}, and y^{T} respectively. Therefore, a multi-dimensional vector (y_q)is represented as in Eq. (20):

$$y_q = \left\{y_q^{pl}, y_q^{Vl}, y_q^{Ml}, y_q^{T}\right\}, (q = 1, \ldots, \tau) \tag{20}$$

It is composed of affecting parameters represent the input on the qth day and y_{q+1}^{pl} that reflects the result on the next day. Dataset of the proposed model is represented in Eq. (21):

$$\mathfrak{D} = \left\{\left(Y_j, Z_j\right) \in \left(\mathcal{R}^{4u} \times \mathcal{R}\right)\right\}_{j=u}^{\tau} \tag{21}$$

where

$$Y_j = \left[y_j^{Pl}, y_q^{Vl}, y_q^{Ml}, y_q^{T}, \ldots, y_{q-u+1}^{Pl}, y_{q-u+1}^{Vl}, y_{q-u+1}^{Ml}, y_{q-u+1}^{T}\right] \tag{22}$$

The input to specify the delay value of the historical data is depicted in Eq. (22), and $Z_j = y_{q+1}^{Pl}$ is set as the output of Y_j. Through Y_j and Z_j, I/O pair $\tau - u + 1$ have been developed for modeling and prediction. Where u the embedded dimension is picked by trial and error, Section 4.3 contains more information about it.

4.2. Performance metrics

Four statistical errors are chosen to calculate the prediction accuracy of the developed MEMD-ADE-SVM hybrid framework: mean absolute percentage error (MAPE), root mean square error (RMSE), R-squared ($\mathfrak{R}^2$), and directional accuracy (DA). The four evaluation criteria can be used as a baseline for energy system decision-making. $\mathfrak{R}^2$ is characterized by Eq. (23), which is widely used to assess the appropriate levels of various benchmarks on the same test dataset. The greater the value of R^2, the superior the predicting performance for the framework is Zhang et al. (2021), He et al. (2019). Eq. (24) defines RMSE, that is a term commonly used to quantify the relative squared error between the actual and forecasted loads. Closer the forecasted value to the true value, the RMSE value would be lessened (He et al., 2020; Dewangan et al., 2020; Kumar et al., 2016). MAPE is defined as Eq. (25), and It is frequently used to compute the average absolute inaccuracy between actual and predicted loads. The smaller the MAPE value, the superior the prediction performance of the model (Memarzadeh and Keynia, 2021; He et al., 2020; Dewangan et al., 2020). DA is expressed as Eq. (26), a method for measuring the accuracy of forecasting direction and giving investors with the current trend (Wang et al., 2019).

$$R^2 = 1 - \frac{\sum_{q-1}^{M}\left(X_q - \hat{X}_q\right)^2}{\sum_{q=1}^{M}\left(X_t - X\right)^2} \tag{23}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{q-1}^{M}\left(X_q - \hat{X}_q\right)^2}{M}} \tag{24}$$

$$\text{MAPE} = \frac{1}{M}\sum_{q=1}^{M}\left|\frac{X_q - \hat{X}_q}{X_q}\right| \times 100 \tag{25}$$

$$\mathfrak{DA} = \frac{\sum_{q-2}^{M} u_q}{M-1} \times 100, q = 2, \ldots, N \tag{26}$$

$$s.t. \quad u_q = \begin{cases} 1, \text{ if } \left(x_q - x_{q-1}\right)\left(\hat{x}_q - \hat{x}_{q-1}\right) \geq 0 \\ 0, \text{ otherwise} \end{cases} \tag{27}$$

Table 3
The duration of time and the volume of the experimental data sets.

Data origin	Time span	Size of samples	Training dataset	Testing dataset
ISO-NE	January 1, 2017 to December 31, 2020	2308	1615	692

where M denotes the size of the testing dataset, X_q and $\hat{X}_q$ are actual and forecast PLs at the qth point in time, respectively. u_q signifies the forecast direction and the observation, and $\bar{X}$ is the mean of the actual PL. In addition, a quantitative testing strategy is used to evaluate the expected validity of the developed MEMD-ADE-SVM hybrid framework against other competing models. A powerful preliminary test called ANOVA tests the null hypothesis of equivalence to determine if there is a significant performance gap between all the models compared (Xiong et al., 2014). Secondly, Diebold–Mariano (DM) test shows a significant difference in predictive performance between the developed hybrid framework and other equivalent frameworks at certain probability values as demonstrated in Eq. (28) (Li et al., 2020). However, DM tests are used to successfully remove the constraints of stochastic variance of instances and to determine and minimize framework prediction errors compared to other frameworks, which can provide stability throughout the analysis.

$$\mathcal{DM} = \frac{l_{\text{mean}}}{l_{\text{std}}}$$
$$s.t \begin{cases} F_c = \left[c^1, c^2, \ldots, c^{\tau}\right] \\ F_d = \left[d^1, d^2, \ldots, d^{\tau}\right] \\ l^i = c^j - d^j \\ l_{\text{mean}} = \frac{\sum_{j-1}^{\tau} l^2}{\tau} \\ l_{\text{std}} = \sqrt{\frac{\sum_{j=1}^{\tau}\left(l^2 - l_{\text{mean}}\right)^2}{\tau - 1}} \\ j = 1, \ldots, \tau \end{cases} \tag{28}$$

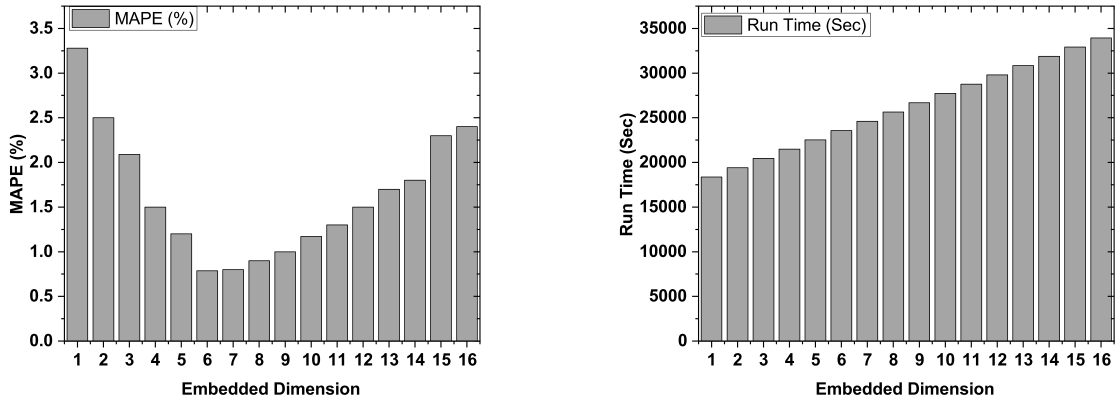
4.3. Experimental implementation

The devised MEMD-ADE-SVM hybrid framework is run in the MATLAB R2020b environment, where MEMD plays a vital role in enhancing forecasting accuracy. The ADE is used to find the best SVM parameters to enhance prediction performance. ADE settings are chosen by trial and error. LIBSVM (Version 3.24), an SVM library offered by Chang and Lin (2011), is used to implement SVM. To enhance forecasting accuracy, three SVM hyperparameters, C, ϵ, γ, are fine-tuned in the training phase using ADE-based hyperparameter optimization. The search space of parameters is defined: $C \in [0.1, 1000]$ $\gamma \in [0.001, 1000]$, and $\varepsilon \in [0.001, 0.1]$, respectively. The fitness function of ADE is used as the average of MAPE to produce and assess the optimal parameters in SVM. The lower the MAPE value, the better the particle's modeling and prediction. The possible size of the embedded dimension is specified from 1 to 16 throughout the training process. To find the appropriate embedded dimension, we must trade-off prediction accuracy and computing time in a real-world sample. As a result, the best one $u = 6$ is chosen, as shown in Table 4 and Fig. 5.

5. Experimental results

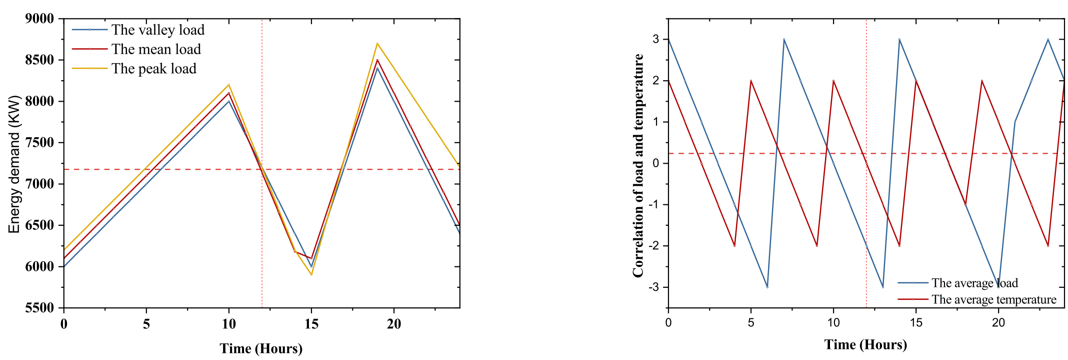
5.1. Simulation for data analysis

Hourly based historical load data gathered from the ISO New England Control Area from 2017 to 2020, having more than 5000 records that is freely available to the public (ISO, 2022). The



(a) u according to MAPE **(b)** u according to computational cost

Fig. 5. Achieved trade off between forecasting accuracy and computational cost for optimal u in real world samples. The best one $u = 6$.



(a) The historical load wave form for the peak load, the valley load, and the mean of load. **(b)** Correlation of the mean historical load with the daily mean temperature.

Fig. 6. Significant variation of historical load waveforms for peak loads, valley loads and average loads, and the correlation between daily mean historical loads and mean daily temperatures.

Table 4
Embedded dimensions related to performance by trial and error.

Embedded dimension	MAPE% ISO-NE	Processing time (sec) ISO-NE
1	2.9125	12156.6561
2	2.6091	13156.7698
3	1.9960	15491.7601
4	1.4912	18171.1378
5	1.3012	19876.9881
6	0.7866	17279.1995
7	0.9245	2010.1871
8	1.0988	21642.7023
9	0.9987	22161.9813
10	1.2133	23121.1451
11	1.2994	23891.1341
12	1.6117	24181.0341
13	1.8134	27191.7810
14	1.9145	3001.1656
15	2.4127	3391.1903
16	2.5678	3651.1676

y^{Vl}, y^{Pl}, y^{Ml} along with y^{T} are eliminated and accumulated to generate experimental datasets. The presence of non-linearity, non-stationarity, and transients as seasonally fluctuation features and the considerable variation of the current waveform of the y^{Vl}, y^{Pl}, y^{Ml}, and y^{T}, respectively is depicted in Fig. 6. It is easily observed that the y^{Pl} curve has highly stochastic fluctuations. Meteorological, seasonality, temporal segmentation, vacations, load lags, load distribution, and cultural, geopolitical, and economic considerations are crucial in daily electricity usage. In this article, temperature is considered an independent factor to avoid exponentially increasing errors and reduce the computational complexity of optimization techniques. Fig. 6(a) demonstrate how the daily mean of y^{pl} is chosen to examine the relationship with the daily y^{T}. The Fig. 6(b) shows that whether the temperature is at its lowest or highest point, the y^{PL} approaches the entire bag. Temperature's importance in day-ahead PLF has been established by the association between y^{pl} and y^{T}. As a result, selecting historical y^{T} as input variables is advantageous.

5.2. Deterioration analysis

The proposed MEMD algorithm decomposed the multivariate input variables in the dataset simultaneously to accurately predict the day peak load and determine load demand instability patterns specific to the result. As seen in Fig. 3, decomposition yields ten $\Im_{mf}$ multivariate constituents ranging in frequency from higher to the lowest, and one $\mathfrak{r}_f$ element. However, it must be noticed that the number of multi-variable objects accessed $\Im_{mf}$ is dependent by the random sample. The number of features retrieved varies if the sample size obtained from data varies significantly or is an integral multiple of 2. Furthermore, the developed time–frequency spectrum indicates that all $\Im_{mf}$ constituents swing vertically and horizontally at point zero, but the waveform of $\Im_{mf}$ remains asymmetric. We recreate the elements of each channel to verify the underlying evolution trend of variance in load

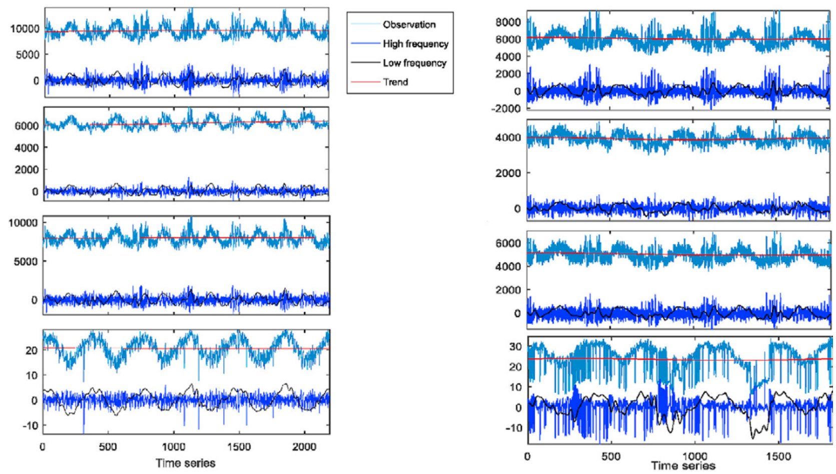


Fig. 7. Regeneration of $\mathfrak{I}_{mf}$ & $\mathfrak{r}_{f}$ of each channel to investigate the patterns of fluctuations in energy demand inherent in evolution.

Table 5
The devised and benchmark frameworks parameters for simulation (Sideratos et al., 2020; Zhang et al., 2017).

Parameters	Values
Population	24
Max-upper bound population	0.9
Min-lower bound population	0.1
Decision variables	2
Max iterations	100
Rate of learning	0.002
Decaying weight	0.0002
Initial value of weights	0.1
Initial bias	0
No of objectives	2
Momentum	0.5
Feature selection threshold	0.5

demand. As illustrated in Fig. 7, the observation of the actual signal is mostly based on the $\mathfrak{r}_{f}$, which presents the key pattern of each channel and is non-influenced by the parameters of linked $\mathfrak{I}_{mf}$ particles. The foreseeing of $\mathfrak{r}_{f}$ is critical for decision-makers of LTLF, whereas high-frequency parts $\left(\mathfrak{I}_{mf_1} \sim \mathfrak{I}_{mf_5}\right)$ fluctuated with small amplitude meditate everyday instability objects in PLF, such as workday, weekend, and resident load demand. Furthermore, low frequency $\left(\mathfrak{I}_{mf_6} \sim \mathfrak{I}_{mf_{10}}\right)$ elements took advantage of important characteristics such as social activity and seasonality. However, some unavoidable spontaneous factors simulate predictive accuracy and complexity.

5.3. Statistical analysis

The 30% of load data is kept for testing and validation reasons, with the remaining 70% used for training. The devised framework is verified against benchmark frameworks such as DCP-SVM-WO (Sideratos et al., 2020), VMD-FFT-IOSVR (Fan et al., 2017), MEMD-PSO-SVR (Huang et al., 2017), VMD-SVR-CGWO (Zhang and Hong, 2021), and MEMD-ADE-LSTM considering convergence rate and accuracy. These frameworks were chosen for their architectural similarity to the proposed framework; this is necessary for a valid comparison. Table 5 demonstrates the model variables for the proposed and benchmark methods. The following is a comprehensive narrative of the simulation consequences:

5.3.1. Day ahead and week ahead PLF

The day ahead PLF profile with an hourly time horizon of the devised model and benchmark models like MEMD-PSO-SVR, DCP-SVM-WO, VMD-SVR-CGWO, MEMD-ADE-LSTM, and VMD-FFT-IOSVR are presented in Fig. 8(a). From the graph, we can see that all predictive models, including the ones we have developed, can grasp the behavior of nonlinear loads from historical data and predict future electrical loads based on the captured behavior. It is also clear that models such as MEMD-PSO-SVR, DCP-SVM-WO, and VMD-FFT-IOSVR use the sigmoid activation function, multivariate AR, and Levenberg Marquardt algorithms for network training. In contrast, the proposed SVM network has a short run time and is trained using kernel functions. Fig. 8(b) confirms that devised framework is closer to the targeted curve compared to the benchmark models. Table 6 shows numerical observations of exposures in MAPE for the day ahead ISO-NE load data with hourly resolution. The proposed SVM-based model's MAPE error is 0.57%; the EMD-PSO-SVR model is 2.63%, and the DCP-SVM-WO model is 1.92%, the VMD-FFT-IOSVR model is 1.77%, the VMD-SVR-CGWO model is 1.39%, and the MEMD-ADE-LSTM model is 1.22%. The MAPE of the developed model is lower than the comparison models, and the lower the MAPE, the higher the accuracy. However, this improved accuracy comes at the expense of longer execution times. The devised SVM-based model is superior to EMD-PSO-SVR, DCP-SVM-WO, VMD-FFT-IOSVR, VMD-SVR-CGWO, and MEMD-ADE-LSTM due to integrating the ADE-based optimization engine.

Fig. 8(a) depicts a comparison of targeted EL vs. predicted EL for the devised and benchmark frameworks during the week of 04/16/2020 to 04/22/2020. The simulation results in Table 7 demonstrate the proposed framework and benchmark models. Evaluation of the proposed framework and the benchmark frameworks like EMD-PSO-SVR, DCP-SVM-WO, VMD-FFT-IOSVR, VMD-SVR-CGWO, and MEMD-ADE-LSTM in terms of MAPE (%) for the week with a day time horizon. The proposed and benchmark models such as EMD-PSO-SVR, DCP-SVM-WO, VMD-FFT-IOSVR, VMD-SVR-CGWO, and MEMD-ADE-LSTM have an average MAPE of 0.57, 2.63, 1.92, 1.77, 1.39 and 1.22 respectively.

5.4. Accuracy considering load data with temperature

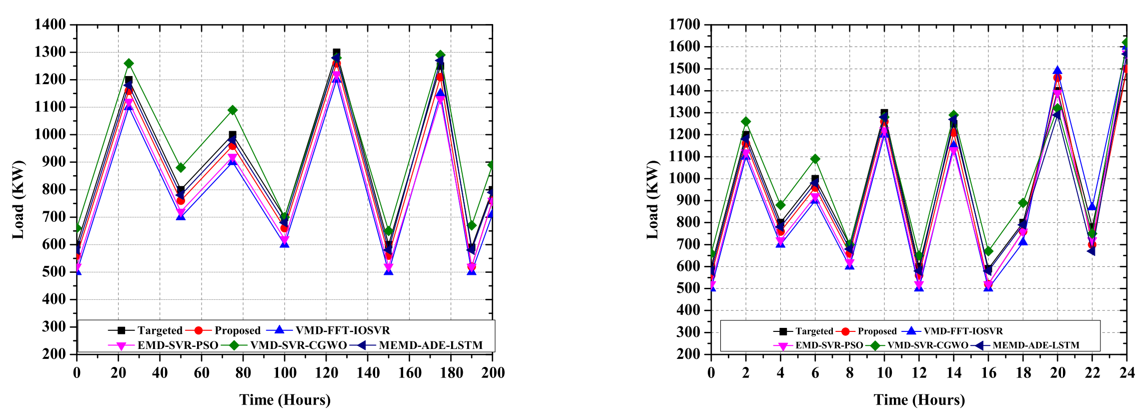
Table 8 compares the mean difference between DA, RMSE, R^2 and MAPE for all models across the IS-ONE dataset. A single models (SVM and LSTM) have the lowest DA and MAPE performance of all models. The reason for this could be an essential limitation of the PLF's single SVM and LSTM models, such as overfitting the training set and local optimization. We compared the traditional EMD-SVR-PSO with the MEMD-ADE-SVM model; the EMD-SVR-PSO model obtains the lowest accuracy

Table 6
Evaluation of the devised and the benchmark frameworks for the 15 April 2020 With hour resolution in terms of MAPE (%).

Proposed and benchmark forecasting models						
Hours	VMD-FFT-IOSVR MAPE (%)	DCP-SVM-WO MAPE (%)	EMD-SVR-PSO MAPE (%)	VMD-SVR-CGWO MAPE (%)	MEMD-ADE-LSTM MAPE (%)	Proposed MAPE (%)
0	2.4	2.1	1.9	1.4	1.1	0.9
1	2.3	1.9	1.8	1.5	1.2	0.7
2	2.2	2	1.7	1.6	1.3	0.8
3	2.1	1.9	1.9	1.3	1.1	0.5
4	2.1	1.8	1.6	1.4	1.4	0.9
5	2.	1.8	1.8	1.3	1.3	0.7
6	1.9	1.75	1.7	1.4	1.4	0.7
7	2.7	1.7	1.5	1.3	1.5	0.6
8	3.1	1.65	1.6	1.2	1.3	0.6
9	2.5	1.6	1.5	1.3	1.3	0.8
10	2.4	1.55	1.4	1.3	1.3	0.6
11	2.6	1.5	1.9	1.2	1.2	0.5
12	2.6	2.3	1.9	1.2	1.2	0.4
13	2.7	2.4	2	1.3	1.3	0.5
14	2.7	2.5	2.1	1.7	1.1	0.6
15	2.8	2.1	2.2	1.1	1.1	0.7
16	2.8	2	1.8	1.2	1.2	0.4
17	2.9	1.8	1.7	1.7	1.2	0.3
18	2.9	1.9	1.6	1.3	1.3	0.6
19	3	1.8	1.5	1.4	1.4	0.5
20	3.1	1.8	1.8	1.5	1.1	0.4
21	3.2	1.7	1.7	1.7	1.1	0.5
22	2.7	1.6	1.6	1.6	1	0.3
23	2.8	2.2	2.2	1.3	0.9	0.2
Average	**2.63**	**1.92**	**1.77**	**1.39**	**1.22**	**0.57**

Table 7
Evaluation of the devised and the benchmark frameworks for the week time horizon of 04/16/2020 to 04/22/2020 in terms of MAPE (%).

Proposed and benchmark forecasting models						
Days	VMD-FFT-IOSVR MAPE (%)	DCP-SVM-WO MAPE (%)	EMD-SVR-PSO MAPE (%)	VMD-SVR-CGWO MAPE (%)	MEMD-ADE-LSTM MAPE (%)	Proposed MAPE (%)
Monday	2.6	2	1.7	1.3	1.3	0.5
Tuesday	2.5	1.9	1.9	1.6	1.4	0.8
Wednesday	2.4	1.8	1.8	1.7	1.2	0.7
Thursday	2.7	1.5	2.1	1.7	1.3	0.5
Friday	2.8	2.2	1.5	1.2	1.4	0.4
Saturday	2.7	2	1.7	1.1	1.2	0.6
Sunday	2.7	2.1	1.8	1.1	1.1	0.5
Average	**2.63**	**1.93**	**1.78**	**1.38**	**1.27**	**0.57**

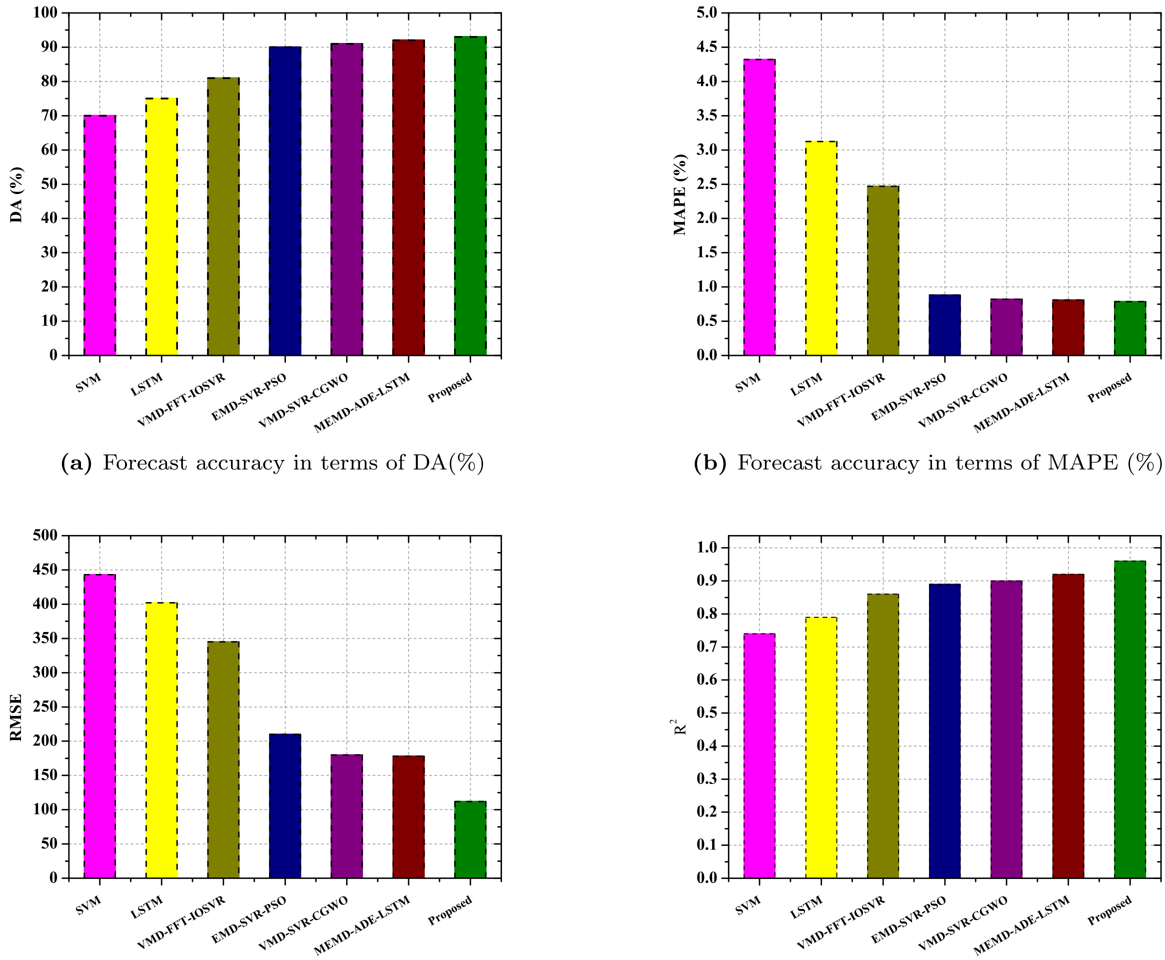


(a) Week ahead forecasting (16-04-2020 to 22-04-2020.) **(b)** Day ahead forecasting.

Fig. 8. Evaluation of devised and benchmark frameworks of ISO-NE energy sector hourly load dataset. (a) Day ahead forecasting; (b) Week ahead forecasting (16-04-2020 to 22-04-2020.).

Table 8
The predictive accuracy of proposed and other frameworks in the real-world testing sets considering historical load along-with temperature.

Models	Input variables		DA (%)	MAPE (%)	RMSE	R^2
	Load	Temperature				
SVM	✓	✓	70.3165	4.231	443.132	0.743
LSTM	✓	✓	75.12	3.124	402.1	0.798
VMD-FFT-IOSVR	✓	✓	81.431	2.471	345.871	0.867
DCP-SVM-WO	✓	✓	84.231	2.141	184.241	0.8771
EMD-SVR-PSO	✓	✓	90.231	0.881	210.918	0.8923
VMD-SVR-CGWO	✓	✓	91.227	0.823	180.918	0.9013
MEMD-ADE-LSTM	✓	✓	92.227	0.819	178.218	0.9231
Proposed	✓	✓	93.145	0.786	112.147	0.9612



(a) Forecast accuracy in terms of DA(%)

(b) Forecast accuracy in terms of MAPE (%)

(c) Forecast accuracy in terms of R^2

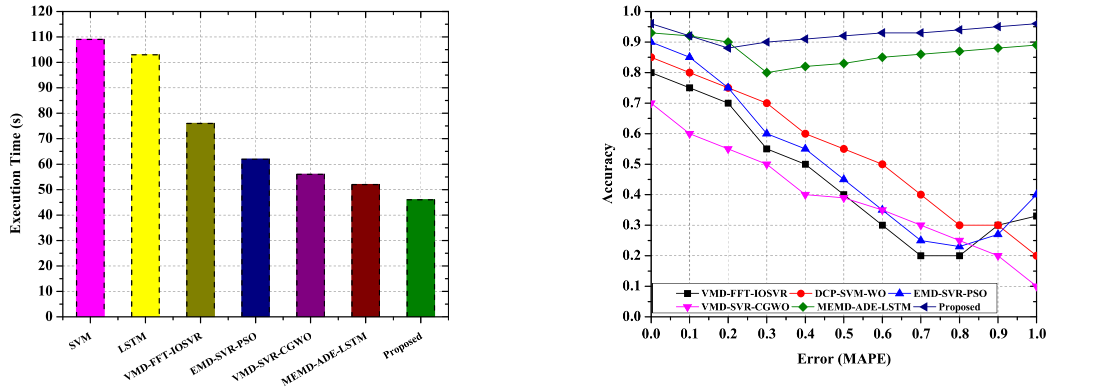
(d) Forecast accuracy in terms of RMSE

Fig. 9. Forecast accuracy of the devised and other benchmark frameworks in real world testing data.

compared to the MEMD-ADE-SVM model. However, these hybrid frameworks perform better than a single model due to the time–frequency resolution algorithm. In addition, MEMD-ADE-LSTM offers excellent performance in DA and R^2 as compared to the EMD-SVR-PSO. This is due to the ADE optimization technique. The devised MEMD-ADE-SVM framework is more accurate by optimizing ADE compared to the relevant benchmark models. The hybrid model presented better performance than the analyzed error metric analysis models. Similarly, the VMD-based framework VMD-SVR-CGWO proposed in Zhang and Hong (2021). We approximated the procedures of univariate STLF-based SVR optimized by using PSO presented a model in this paper. This report accentuates the significance of irrelevant variables and the efficacy of multivariate decomposition predictions. Table 8 and Fig. 9 show the highest performance error protection.

5.5. Evaluation of convergence rate

Comparative analysis of the devised SVM based model and other benchmark models like DCP-SVM-WO, VMD-FFT-IOSVR, EMD-PSO-SVR, VMD-SVR-CGWO, and MEMD-ADE-LSTM consider the system's convergence rate shown in Fig. 10(a). There is a trade-off between rate of convergence and prediction accuracy. The accuracy of the VMD-FFT-IOSVR strategy has been enhanced compared to other models, including the proposed model. Since the optimization engine is integrated into the VMD-FFT-IOSVR strategy, this improved accuracy is achieved at the expense of



(a) Comparative analysis of the proposed MEMD model and other reference models taking into account the convergence rate on the ISO-NE load data .

(b) Robustness analysis of our devised and other benchmark frameworks adding noise to each feature and observing accuracy.

Fig. 10. Performance evaluation in terms of convergence rate and robustness.

longer execution times. Fig. 10(a) show that the execution time has increased from 15 s to 109 s as the optimization module is integrated into the forecasting module. The developed model has reduced execution time for the following reasons: (i) Provides mostly abstract attributes as input to the training and prediction engine, reducing network training time. (ii) Use kernel functions. (iii) Use ADE instead of the PSO, CGWO, and FFT algorithms because the convergence speed is relatively fast. The proposed STLF model reduced the running time from 109 s to 46 s due to adaptations in current models. In contrast, the SVM model does not have a built-in optimizer, so SVM outperforms the other models in evaluating the rate of convergence. This behavior is clearly illustrated in Fig. 10(a).

5.6. Evaluation of robustness

Stochastic noises (Harmonic noise, white noise, asymmetric dichotomous noise, and Lévy noise) have a tremendously adverse influence on the forecasting accuracy of PLF. Pre-filtering in real-time can effectively improve measurement accuracy. Pretreating and inspecting the electric power load data statistically is essential to illustrate the stochastic noise of electric load. The proposed MEMD is used to denoise load data. MEMD significantly reduced the stochastic noise amplitude of load data. Therefore, the proposed time series and MEMD method can effectively suppress the stochastic noise of load data and enhance the forecast accuracy of load to sustain the robustness and resilience of the proposed model. Stochastic noise of load is presented with different correlation time and power spectral density, and total variance method is effective to evaluate five kinds of stochastic noise of load data before and after feature engineering, including load random walk (L), bias instability (B), rate ramp walk (K), rate ramp (R) and quantization noise (Q). Table 9 is each stochastic noise coefficient before and after FE. Fig. 10(b) illustrates the resilience analysis of the devised and benchmark models. The assessment is performed by including error (noise) to each feature and observing the accuracy of each Strategy. The devised framework is more robust and resilient than the benchmarks frameworks as illustrated in Fig. 10(b), because the noise within features has low effect on accuracy, and therefore, less important and irrelevant features are dropped during the feature engineering process. Thus, the proposed framework is also robust against noise in the features.

5.7. Validation of forecasting performance

Two statistical methods (DM (Li et al., 2020) and ANOVA (Xiong et al., 2014) tests) used to further verify the predictive performance that there was a statistically significant difference between the evaluated frames of the PLF. Firstly, the ANOVA test is employed to find a statistically significant performance difference when comparing the devised MEMD-ADE-SVM with others frameworks. However, the null hypothesis of the ANOVA is denied when not finding the pairwise differences of the frameworks. Then the DM test is performed to detect the pairwise discrepancies simultaneously with a 5% significant level. DM tests have been found to efficiently extract the limits of arbitrary sample operation and provide a complete assessment of accuracy and stability with fast convergence rates in a short period of time. The DM test was created to assess the predictive inaccuracies of different fused models. In all conditions, the DM test results continued to exceed the 5% significant rating absolute limit. Table 10 shows p values correlated with statistical results, explaining that the proposed hybrid forecasting model (MEMD-ADE-SVM) at the 1% significance level is excellent compared to other models . From the concrete measurements of the DM test, we can infer that the developed model structure is more accurate in prediction than other models, there is a significant difference at a certain level, and the performance consistency of the model proposed by PLF is confirmed.

Table 9
Stochastic noise coefficients before and after MEMD.

Noise coefficient	Before MEMD	After MEMD
Q	1.49e−3	1.81e−4
L	1.02e−5	1.32e−6
B	4.21e−4	4.87e−5
K	6.35e−3	9.12e−4
R	5.21e−2	4.20e−3

Table 10
The DM test values between the devised and other seven models in the ISO-NE real-world testing set.

Frameworks (Individual and Hybrids)	ISO-NE	
	DM	P-value
SVM	7.0125	0.0000
LSTM	9.3244	0.0000
VMD-FFT-IOSVR	5.6834	0.0000
EMD-SVR-PSO	4.9951	0.0000
DCP-SVM-WO	4.5726	0.0000
VMD-SVR-CGWO	4.3213	0.0000
MEMD-ADE-LSTM	4.2732	0.0000
Proposed	4.1031	0.0000

6. Discussion

Our observations could be used to build a framework for day-ahead PLF using the multivariate time–frequency method. The proposed model provided two advantages: (1) This has reduced the computational overhead required to analyze the day ahead PL properly. (2) it provides a flexible PLF framework for the financial functioning of the energy system. The following are our primary performance strategies: meteorological factors are considered. As described at the beginning, the day ahead PLF is essential for the management and transmission of energy systems. However, the univariate historical load implementation appears deficient. Considering the factors that affect all PLs, it should be emphasized that inaccuracies and the complexity of the integration framework's calculations. Temperature is an important part of the PLF meteorological parameter, but it is included as an input variable in this study.

Secondly, combining different strategies into a hybrid framework appeared viable for improving forecast accuracy. Despite the fact that individual frameworks can account for additional load-related factors, predictive accuracy remains inadequate. As a result, we used the time–frequency analysis method in our research. MEMD takes advantage of the non-linearity of multi-dimensional time series to restore unique information from the raw time series produced by the wavelet decomposition.

The time–frequency deterioration method has the potential to improve PLF accuracy significantly as depicted in Table 8. In addition, we employed the EMD approach exhaustively characterize the MEMD approach due to not considering the temperature data and its integration into the devised framework. We employed univariate time series modeling with SVM (either with an optimization algorithm or without) to apprehend the differences, because EMD cannot deal with multivariate time series data. The EMD-SVM model exemplifies a single SVM model for predictive accuracy, even if temperature is included in the SVM model. In addition, the MEMD-SVM approach with temperature variables is superior to all error measurements compared to models without temperature variables. Therefore, we emphasize the advantages of the MEMD-ADE-SVM model proposed for the day ahead PLF. First, using an adaptive time–frequency analysis approach, the MEMD technique can thoroughly comprehend multi-dimensional data's non-linear and non-stationary properties. Second, SVM enhances three parameters using ADE, this allows us to overcome the multi-parameter limitations of other benchmark models. Finally, the proposed MEMD-ADE-SVM framework outperforms the developed and other models in DA, and MAPE, while the DM test results and P-values further demonstrate the proposed model's superiority depicted in Table 10.

6.1. Superiority analysis of proposed MEMD-ADE-SVM

The performance and superiority investigation of MEMD-ADE-SVM and benchmark frameworks in computational time (τ) is depicted in Table 11. The individual/single models such as LSTM, and SVM, without integrating decomposition techniques (MEMD/EMD/VMD) and optimization algorithms for tuning hyperparameters, have low computational time and worst error performance. However, when both FE and optimizing modules are merged with these individual models, the computational time is increased. The error is reduced due to the trade-off between accuracy and convergence rate. The statistical evaluations of the proposed framework and benchmark frameworks in terms of forecast errors and computational time are listed in Table 11. The average values of computational speed and forecast error for both individual (LSTM, and SVM) models and hybrid (DCP-SVM-WO, VMD-FFT-IOSVR, EMD-PSO-SVR, VMD-SVR-CGWO, and MEMD-ADE-LSTM) models for daily, and weekly time horizon are listed in Table 11. The computational speed of individual models (LSTM, and SVM) without the integration of MEMD/VMD/EMD decomposition technique and optimization modules and hybrid (DCP-SVM-WO, VMD-FFT-IOSVR, EMD-PSO-SVR, VMD-SVR-CGWO, and MEMD-ADE-LSTM) models with the integration of decomposition techniques (MEMD/VMD/EMD) and optimization algorithms is computed on a Core(TM)-i3-3110M, CPU@2.40 GHz with 8 GB RAM system. The individual models LSTM, and SVM have the lowest computational time for day ahead 170s, and 185, respectively, while the hybrid models have increased computational time for day ahead 242s, 228 s, 224 s, 221 s, and 218 s for DCP-SVM-WO, VMD-FFT-IOSVR, EMD-PSO-SVR, MEMD-ADE-LSTM, VMD-SVR-CGWO, and devised frameworks, respectively. The increase in computational time is encountered as the optimizer module or the MEMD/EMD/VMD module, or both modules are integrated with the individual forecaster models. Moreover, this increment in time is due to the trade-off between convergence speed and accuracy, thus achieving more accuracy at the cost of surplus computational time. The proposed framework reduces the computational time by devising modifications in the DE optimization algorithm to produce ADE (see Eqs. (18) and (19)).

6.2. Transcends of shallow learning network over the deep learning network

Deep learning (DL) has become a robust ML approach that has been widely prosperous in multiple applications. DL is one of the best techniques for yanking knowledge from large sets of raw data in a (nearly) self-organized manner. The technical design of DL depends on the feed-forward information flow principle of ANNs with multiple layers of hidden neurons, which form deep neural networks (DNNs). DNNs have various architectures and parameters and are often developed for specific applications. However, the training process of DNNs can be extended based on the application and training set size (Gong et al., 2015). Moreover, finding the most precise and efficient architecture of a DL system at a suitable time is a possible problem associated with this strategy. Darwish, Hassanien, and Das (Darwish et al., 2020) propose a survey on DL-based forecasting models. In this study, the authors state that hyperparameter optimization is critical and time-consuming, so optimization approaches should be used. This paper reviews only those forecasting algorithms whose hyperparameters are optimized using swarm intelligence and evolutionary algorithms. The result of these optimization algorithms on the forecast accuracy of DL-based forecasting strategies is investigated for big data applications Besides, generally used DL methods are also debated, along with their flaws and resilience. Xiao et al. explore the potential applications of data mining techniques in PLF using big data in Xiao et al. (2017). DL-based approaches are employed to analyze data and forecast load consumption. There are two types of DL-based techniques: supervised and unsupervised. Former is used as forecasting models and later for feature extraction.

Shallow learning (SVM) is an ML method generated from the structural risk minimization principle and statistical learning theory. On the one hand, SVM uses the structural risk minimization principle instead of the empirical risk minimization principle to underestimate the training error to evade falling into local optimum like the ANN. On the other hand, SVM is a feature model that can non-linearly map the training data in a low-dimensional plane to a high-dimensional space. Based on these characteristics, SVM can effectively overwhelm the shortcomings of DL models (e.g., weak classification ability and over-fitting) and thus is widely used in ELF, data mining, and other fields.

We prefer the shallow framework over the DL framework because some problems and challenges regarding DL techniques and their architectures still need to be addressed (Bäck et al., 2013; Bae et al., 2016)

Table 11
Evaluation of actual and forecasted PL in terms of MAPE and computational time of Individual frameworks (LSTM and SVM) and hybrid frameworks (DCP-SVM-WO, VMD-FFT-IOSVR, EMD-PSO-SVR, MEMD-ADE-LSTM, VMD-SVR-CGWO) for day-ahead and week-ahead time horizon.

Frameworks without and with MEMD/ EMD/ VMD and optimization modules														
	LSTM		SVM		EMD-PSO-SVR		VMD-FFT-IOSVR		MEMD-ADE-LSTM		VMD-SVR-CGWO		MEMD-ADE-SVM	
	τ(s)	MAPE (%)	τ(s)	MAPE (%)	τ(s)	MAPE (%)	τ(s)	MAPE (%)	τ(s)	MAPE (%)	τ(s)	MAPE (%)	τ(s)	MAPE (%)
Day ahead	170	2.25	185	2.2	242	2.98	228	1.65	224	1.25	221	0.89	218	0.786
Week ahead	185	2.25	198	2.2	310	2.98	289	1.65	276	1.25	267	0.89	256	0.786

- The selection of DL architectures: Till now, different DL architectures have been used and applied in the literature to solve complex problems. However, there is no motivation or documentation on why these architectures have been established (Bera et al., 2014)

- Lack of benchmarking results: There are few benchmark results in the literature, such as studies in Reddy et al. (2016), Badem et al. (2017), Chen et al. (2018). In these studies, authors have involved different deep architectures and compared the results with DTs and BPs to produce the best training results. In addition, the loss of information in any system under analysis can influence the stability of the whole system. This issue must be benchmarked to determine what the best performance is being fetched.

- The cost of implementing the architecture: Features extraction can be done beforehand, and then the suitable algorithm can be enforced as in Gadekallu et al. (2020), Junbo et al. (2015), Liu et al. (2016). This procedure strives to lessen the mandated training time and computational power.

- Reasonable run-time: The high dimensionality of some datasets with many parameters in some DL architectures, such as the DNN model, represents a challenge for DL to acquire accurate DNN in a reasonable run time.

- Overfitting in DNNs: In complex applications, many parameters are related to the unseen dataset. This can cause a difference in the training dataset's error and the error faced in the new unseen dataset. However, the efficiency of the DL model-based ANN can be evaluated by the ability to perform unseen datasets.

- The optimization of hyper-parameters: The hyper-parameters are those whose value is defined before the learning process. Any modification in these parameters simulates the performance of the DL model.

- High hardware performance is required: High processing power is needed to vend with a real-world application using DL solutions. Therefore, engineers and specialists are trying to develop multi-core, high-performing GPUs and similar processing units like the recently upcoming Tensor Processing Units (TPUs).

- Lack of flexibility: DL models can yield accurate and efficient solutions to a disseminated problem. On the other hand, the shallow network architectures are highly specialized to specific application domains.

7. Conclusions

PLF plays a vital role in balancing power distribution, economics, and safe and reliable energy system operations. Accurate PLF reduces energy grid failure, ameliorates costs and risks, improves energy grid security, and helps policymakers in optimal planning and decision making, making the energy grid cost-effective and environment friendly. Moreover, it provides a strategic understanding of the current energy situation to evaluate the energy imported or exported. On this note, most of the existing literature focuses on accuracy improvement. However, considering only the accuracy index is insufficient, convergence rate and stability indices are indispensable. These indices are equally crucial in forecasting. Thus, a novel hybrid framework MEMD-ADE-SVM is devised in this research. The proposed hybrid framework has a novel ADE algorithm for appropriate parameter selection and tuning of SVM. Meanwhile, the MEMD approach simultaneously decomposes historical loads and meteorological variables adaptively to handle the non-linearity and non-stationarity of the day and week ahead PL. It effectively extracts various features at different levels of time frequencies associated with predicting the next day's peak load more accurately. The purpose is to simultaneously acquire high accuracy, excellent stability, and fast convergence. The proposed framework evaluated in terms of accuracy, stability, and convergence rate by comparing it with benchmark frameworks such as DCP-SVM-WO, VMD-FFT-IOSVR, EMD-PSO-SVR, MEMD-ADE-LSTM, and VMD-SVR-CGWO. From the experimental results, the achieved DAs of the benchmark frameworks and the proposed framework are 81.43%, 84.23%, 90.231%, 91.227%, 92.22% and 93.145%, respectively. Therefore, the proposed framework would be the most appropriate option for policymakers and decision-makers to use for load forecasting to ensure power systems' reliable and safe operation.

7.1. Limitations and future work

Energy grids are expected to be more convoluted with the evolution of materializing renewable energy (RE) technologies. The uncertainties of SG systems are boosting as many aspects may affect electricity demand. This paper concentrates not on the future load demand from a long-term perspective but the short-term load fluctuation. The forecasting framework devised in this research does not evaluate other related factors but is only based on the detailed historical short-term load. Many key impacts may be missing, and there are also substantial research gaps there. From the standpoint of life cycle assessment (LCA), research on the whole system from "cradle to grave" is introduced, which can be used in forecasting models. Moreover, scientific scenarios can be inducted to combine long-term and short-term forecasting, and more work must be done in related fields. Follow-up studies could be performed in future work, including but not confined to:

- The forecasting model can evaluate additional characteristics or parameters to enhance the PLF's efficacy.

- Research on energy systems, notably the use of RE, must be studied so that the distribution and structures of future RE can be well known, which is a crucial aspect for PLF.

- Paying close concentration to develop data cleaning technologies to deal with irregular and unstable short-term load data so that the adverse impacts of noise can be effectively handled.

- A dynamic model selection strategy could be evaluated when selecting the weights of hybrid or combined models.
- LCA-based modeling and design analysis can be oriented into forecasting models.
- More case studies in different SG systems could be done to demonstrate the scalability of the proposed forecasting model.

CRediT authorship contribution statement

M. Zulfiqar: Conceptualization, Methodology/Study design, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review and editing, Visualization. **M. Kamran:** Conceptualization, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review and editing, Visualization, Supervision, Project administration. **M.B. Rasheed:** Conceptualization, Methodology/Study design, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review and editing, Visualization, Supervision, Project administration, Funding acquisition. **T. Alquthami:** Conceptualization, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review and editing, Visualization, Supervision, Project administration, Funding acquisition. **A.H. Milyani:** Conceptualization, Validation, Investigation, Resources, Writing – review and editing, Project administration, 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.

Data availability

No data was used for the research described in the article.

Acknowledgment

This project has received funding from the European Union Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 754382, GOT ENERGY TALENT. Furthermore, The Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia has also funded this project, under grant no. (RG-34-135-42). Therefore, authors greatly acknowledge technical and financial support from the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

References

Ahmad, T., Chen, H., 2019. Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems. Sustainable Cities Soc. 45, 460–473.

Al-Musaylh, M.S., Deo, R.C., Adamowski, J.F., Li, Y., 2018a. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. Adv. Eng. Inform. 35, 1–16.

Al-Musaylh, M.S., Deo, R.C., Li, Y., Adamowski, J.F., 2018b. Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting. Appl. Energy 217, 422–439.

Amjady, N., Keynia, F., Zareipour, H., 2010. Short-term load forecast of microgrids by a new bilevel prediction strategy. IEEE Trans. Smart Grid 1 (3), 286–294.

Bäck, T., Foussette, C., Krause, P., 2013. Contemporary Evolution Strategies, Vol. 86. Springer.

Badem, H., Basturk, A., Caliskan, A., Yuksel, M.E., 2017. A new efficient training strategy for deep neural networks by hybridization of artificial bee colony and limited–memory BFGS optimization algorithms. Neurocomputing 266, 506–526.

Bae, C., Kang, K., Liu, G., Chung, Y.Y., 2016. A novel real time video tracking framework using adaptive discrete swarm optimization. Expert Syst. Appl. 64, 385–399.

Bashir, T., Haoyong, C., Tahir, M.F., Liqiang, Z., 2022. Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN. Energy Rep. 8, 1678–1686.

Bera, S., Misra, S., Rodrigues, J.J., 2014. Cloud computing applications for smart grid: A survey. IEEE Trans. Parallel Distrib. Syst. 26 (5), 1477–1494.

Cai, M., Pipattanasomporn, M., Rahman, S., 2019. Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques. Appl. Energy 236, 1078–1088.

Cao, G., Wu, L., 2016. Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting. Energy 115, 734–745.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2 (3), 1–27.

Chen, J., Zeng, G.-Q., Zhou, W., Du, W., Lu, K.-D., 2018. Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. Energy Convers. Manage. 165, 681–695.

Chen, Y., Zhang, F., Berardi, U., 2020. Day-ahead prediction of hourly subentry energy consumption in the building sector using pattern recognition algorithms. Energy 211, 118530.

Coelho, V.N., Coelho, I.M., Coelho, B.N., Reis, A.J., Enayatifar, R., Souza, M.J., Guimarães, F.G., 2016. A self-adaptive evolutionary fuzzy model for load forecasting problems on smart grid environment. Appl. Energy 169, 567–584.

Dai, Y., Zhao, P., 2020. A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization. Appl. Energy 279, 115332.

Darwish, A., Hassanien, A.E., Das, S., 2020. A survey of swarm and evolutionary computing approaches for deep learning. Artif. Intell. Rev. 53 (3), 1767–1812.

Deng, C., Zhang, X., Huang, Y., Bao, Y., 2021. Equipping seasonal exponential smoothing models with particle swarm optimization algorithm for electricity consumption forecasting. Energies 14 (13), 4036.

Dewangan, C.L., Singh, S., Chakrabarti, S., 2020. Combining forecasts of day-ahead solar power. Energy 202, 117743.

Fan, G.-F., Peng, L.-L., Zhao, X., Hong, W.-C., 2017. Applications of hybrid EMD with PSO and GA for an SVR-based load forecasting model. Energies 10 (11), 1713.

Gadekallu, T.R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P.K.R., Srivastava, G., 2020. Deep neural networks to predict diabetic retinopathy. J. Ambient Intell. Humaniz. Comput. 1–14.

Gong, M., Liu, J., Li, H., Cai, Q., Su, L., 2015. A multiobjective sparse feature learning model for deep neural networks. IEEE Trans. Neural Netw. Learn. Syst. 26 (12), 3263–3277.

Guan, Y., Li, D., Xue, S., Xi, Y., 2021. Feature-fusion-kernel-based Gaussian process model for probabilistic long-term load forecasting. Neurocomputing 426, 174–184.

Hafeez, G., Javaid, N., Riaz, M., Ali, A., Umar, K., Iqbal, Z., 2019. Day ahead electric load forecasting by an intelligent hybrid model based on deep learning for smart grid. In: Conference on Complex, Intelligent, and Software Intensive Systems. Springer, pp. 36–49.

Haq, E.U., Lyu, X., Jia, Y., Hua, M., Ahmad, F., 2020. Forecasting household electric appliances consumption and peak demand based on hybrid machine learning approach. Energy Rep. 6, 1099–1105.

Haq, M.R., Ni, Z., 2019. A new hybrid model for short-term electricity load forecasting. IEEE Access 7, 125413–125423.

He, F., Zhou, J., Feng, Z.-k., Liu, G., Yang, Y., 2019. A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm. Appl. Energy 237, 103–116.

He, F., Zhou, J., Mo, L., Feng, K., Liu, G., He, Z., 2020. Day-ahead short-term load probability density forecasting method with a decomposition-based quantile regression forest. Appl. Energy 262, 114396.

Hernandez, L., Baladron, C., Aguiar, J.M., Carro, B., Sanchez-Esguevillas, A.J., Lloret, J., Massana, J., 2014. A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings. IEEE Commun. Surv. Tutor. 16 (3), 1460–1495.

Heydari, A., Nezhad, M.M., Pirshayan, E., Garcia, D.A., Keynia, F., De Santoli, L., 2020. Short-term electricity price and load forecasting in isolated power grids based on composite neural network and gravitational search optimization algorithm. Appl. Energy 277, 115503.

Hu, Z., Bao, Y., Chiong, R., Xiong, T., 2015a. Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection. Energy 84, 419–431.
Hu, Z., Bao, Y., Xiong, T., 2014. Comprehensive learning particle swarm optimization based memetic algorithm for model selection in short-term load forecasting using support vector regression. Appl. Soft Comput. 25, 15–25.
Hu, Z., Bao, Y., Xiong, T., Chiong, R., 2015b. Hybrid filter–wrapper feature selection for short-term load forecasting. Eng. Appl. Artif. Intell. 40, 17–27.
Huang, X., Hong, S.H., Li, Y., 2017. Hour-ahead price based energy management scheme for industrial facilities. IEEE Trans. Ind. Inform. 13 (6), 2886–2898.
Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 454 (1971), 903–995.
Huyghues-Beaufond, N., Tindemans, S., Falugi, P., Sun, M., Strbac, G., 2020. Robust and automatic data cleansing method for short-term load forecasting of distribution feeders. Appl. Energy 261, 114405.
ISO, 2022. ISO: ISO New England Data, URL http://www.iso-ne.com/.
Jacob, M., Neves, C., Vukadinović Greetham, D., 2020. Forecasting and Assessing Risk of Individual Electricity Peaks. Springer Nature.
Jang, Y., Byon, E., Jahani, E., Cetin, K., 2020. On the long-term density prediction of peak electricity load with demand side management in buildings. Energy Build. 228, 110450.
Jawad, M., Ali, S.M., Khan, B., Mehmood, C.A., Farid, U., Ullah, Z., Usman, S., Fayyaz, A., Jadoon, J., Tareen, N., et al., 2018. Genetic algorithm-based non-linear auto-regressive with exogenous inputs neural network short-term and medium-term uncertainty modelling and prediction for electrical load and wind speed. J. Eng. 2018 (8), 721–729.
Jiang, H., Wang, K., Wang, Y., Gao, M., Zhang, Y., 2016. Energy big data: A survey. IEEE Access 4, 3844–3861.
Junbo, T., Weining, L., Juneng, A., Xueqian, W., 2015. Fault diagnosis method study in roller bearing based on wavelet transform and stacked auto-encoder. In: The 27th Chinese Control and Decision Conference. 2015 CCDC, IEEE, pp. 4608–4613.
Kavousi-Fard, A., Samet, H., Marzbani, F., 2014. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. Expert Syst. Appl. 41 (13), 6047–6056.
Kazemzadeh, M.-R., Amjadian, A., Amraee, T., 2020. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. Energy 204, 117948.
Khalid, R., Javaid, N., 2020. A survey on hyperparameters optimization algorithms of forecasting models in smart grid. Sustainable Cities Soc. 61, 102275.
Kumar, N., et al., 2016. Market clearing price prediction using ANN in indian electricity markets. In: 2016 International Conference on Energy Efficient Technologies for Sustainability. ICEETS, IEEE, pp. 454–458.
Li, Y., Che, J., Yang, Y., 2018. Subsampled support vector regression ensemble for short term electric load forecasting. Energy 164, 160–170.
Li, K., Ma, Z., Robinson, D., Lin, W., Li, Z., 2020. A data-driven strategy to forecast next-day electricity usage and peak electricity demand of a building portfolio using cluster analysis, Cubist regression models and particle swarm optimization. J. Clean. Prod. 273, 123115.
Li, S., Wang, P., Goel, L., 2015. Short-term load forecasting by wavelet transform and evolutionary extreme learning machine. Electr. Power Syst. Res. 122, 96–103.
Liao, G.-C., 2014. Hybrid improved differential evolution and wavelet neural network with load forecasting problem of air conditioning. Int. J. Electr. Power Energy Syst. 61, 673–682.
Liu, Y., Gao, F., 2020. Ultra-short-term forecast of power load based on load characteristics and embedded system. Microprocess. Microsyst. 103460.
Liu, S., Tian, L.-X., 2013. The study of long-term electricity load forecasting based on improved grey prediction model. In: 2013 International Conference on Machine Learning and Cybernetics, Vol. 2. IEEE, pp. 653–656.
Liu, D., Zeng, L., Li, C., Ma, K., Chen, Y., Cao, Y., 2016. A distributed short-term load forecasting method based on local weather information. IEEE Syst. J. 12 (1), 208–215.
Massaoudi, M., Refaat, S.S., Chihi, I., Trabelsi, M., Oueslati, F.S., Abu-Rub, H., 2021. A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting. Energy 214, 118874.
Memarzadeh, G., Keynia, F., 2021. Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm. Electr. Power Syst. Res. 192, 106995.
Moazzami, M., Khodabakhshian, A., Hooshmand, R., 2013. A new hybrid day-ahead peak load forecasting method for Iran's National Grid. Appl. Energy 101, 489–501.
Nalcaci, G., Özmen, A., Weber, G.W., 2019. Long-term load forecasting: Models based on MARS, ANN and LR methods. CEJOR Cent. Eur. J. Oper. Res. 27 (4), 1033–1049.
Niu, D., Ji, Z., Li, W., Xu, X., Liu, D., 2021. Research and application of a hybrid model for mid-term power demand forecasting based on secondary decomposition and interval optimization. Energy 234, 121145.
Raza, M.Q., Khosravi, A., 2015. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew. Sustain. Energy Rev. 50, 1352–1372.
Reddy, K.K., Sarkar, S., Venugopalan, V., Giering, M., 2016. Anomaly detection and fault disambiguation in large flight data: A multi-modal deep auto-encoder approach. In: Annual Conference of the PHM Society, Vol. 8, no. 1.
Rehman, N., Mandic, D.P., 2010. Multivariate empirical mode decomposition. Proc. Royal Soc. A 466 (2117), 1291–1302.
Rilling, G., Flandrin, P., Gonçalves, P., Lilly, J.M., 2007. Bivariate empirical mode decomposition. IEEE Signal Process. Lett. 14 (12), 936–939.
Sakurai, D., Fukuyama, Y., Iizaka, T., Matsui, T., 2019. Daily peak load forecasting by artificial neural network using differential evolutionary particle swarm optimization considering outliers. IFAC-PapersOnLine 52 (4), 389–394.
Santos Coelho, L.d., Mariani, V., 2006. ERRATUM-correction to" combining of chaotic differential evolution and quadratic programming for economic dispatch optimization with valve-point effect". IEEE Trans. Power Syst. 21 (3), 1465.
Selakov, A., Cvijetinović, D., Milović, L., Mellon, S., Bekut, D., 2014. Hybrid PSO–SVM method for short-term load forecasting during periods with significant temperature variations in city of burbank. Appl. Soft Comput. 16, 80–88.
Shiri, A., Afshar, M., Rahimi-Kian, A., Maham, B., 2015. Electricity price forecasting using support vector machines by considering oil and natural gas price impacts. In: 2015 IEEE International Conference on Smart Energy Grid Engineering. SEGE, IEEE, pp. 1–5.
Sideratos, G., Ikonomopoulos, A., Hatziargyriou, N.D., 2020. A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks. Electr. Power Syst. Res. 178, 106025.
Sobhani, M., Hong, T., Martin, C., 2020. Temperature anomaly detection for electric load forecasting. Int. J. Forecast. 36 (2), 324–333.
Storn, R., Price, K., 1997. Differential evolution–A simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. 11 (4), 341–359.
Talaat, M., Farahat, M., Mansour, N., Hatata, A., 2020. Load forecasting based on grasshopper optimization and a multilayer feed-forward neural network using regressive approach. Energy 196, 117087.
ur Rehman, N., Mandic, D.P., 2009. Empirical mode decomposition for trivariate signals. IEEE Trans. Signal Process. 58 (3), 1059–1068.
Vesterstrom, J., Thomsen, R., 2004. A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems. In: Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No. 04TH8753), Vol. 2. IEEE, pp. 1980–1987.
Vrablecová, P., Ezzeddine, A.B., Rozinajová, V., Šárik, S., Sangaiah, A.K., 2018. Smart grid load forecasting using online support vector regression. Comput. Electr. Eng. 65, 102–117.
Wang, R., Wang, J., Xu, Y., 2019. A novel combined model based on hybrid optimization algorithm for electrical load forecasting. Appl. Soft Comput. 82, 105548.
Wang, K., Xu, C., Zhang, Y., Guo, S., Zomaya, A.Y., 2017. Robust big data analytics for electricity price forecasting in the smart grid. IEEE Trans. Big Data 5 (1), 34–45.
Wood, D.A., 2022. Trend decomposition aids short-term countrywide wind capacity factor forecasting with machine and deep learning methods. Energy Convers. Manage. 253, 115189.
Wu, Z., Zhao, X., Ma, Y., Zhao, X., 2019. A hybrid model based on modified multi-objective Cuckoo search algorithm for short-term load forecasting. Appl. Energy 237, 896–909.
Xiao, L., Shao, W., Wang, C., Zhang, K., Lu, H., 2016. Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting. Appl. Energy 180, 213–233.
Xiao, F., Wang, S., Fan, C., 2017. Mining big building operational data for building cooling load prediction and energy efficiency improvement. In: 2017 IEEE International Conference on Smart Computing. SMARTCOMP, IEEE, pp. 1–3.
Xiong, T., Bao, Y., Hu, Z., 2014. Interval forecasting of electricity demand: A novel bivariate EMD-based support vector regression modeling framework. Int. J. Electr. Power Energy Syst. 63, 353–362.
Yu, Z., Niu, Z., Tang, W., Wu, Q., 2019. Deep learning for daily peak load forecasting–A novel gated recurrent neural network combining dynamic time warping. Ieee Access 7, 17184–17194.
Yuan, X., Wang, L., Yuan, Y., Zhang, Y., Cao, B., Yang, B., 2008. A modified differential evolution approach for dynamic economic dispatch with valve-point effects. Energy Convers. Manage. 49 (12), 3447–3453.
Yuan, X., Wang, L., Zhang, Y., Yuan, Y., 2009. A hybrid differential evolution method for dynamic economic dispatch with valve-point effects. Expert Syst. Appl. 36 (2), 4042–4048.
Zhang, W., Chen, Q., Yan, J., Zhang, S., Xu, J., 2021. A novel asynchronous deep reinforcement learning model with adaptive early forecasting method and reward incentive mechanism for short-term load forecasting. Energy 236, 121492.

Zhang, F., Deb, C., Lee, S.E., Yang, J., Shah, K.W., 2016. Time series forecasting for building energy consumption using weighted support vector regression with differential evolution optimization technique. Energy Build. 126, 94–103.
Zhang, Z., Ding, S., Sun, Y., 2020. A support vector regression model hybridized with chaotic krill herd algorithm and empirical mode decomposition for regression task. Neurocomputing 410, 185–201.
Zhang, Z., Hong, W.-C., 2021. Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads. Knowl.-Based Syst. 228, 107297.
Zhang, X., Wang, J., Zhang, K., 2017. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm. Electr. Power Syst. Res. 146, 270–285.
Zhang, J., Wei, Y.-M., Li, D., Tan, Z., Zhou, J., 2018. Short term electricity load forecasting using a hybrid model. Energy 158, 774–781.