Received: April 20, 2023. Revised: May 20, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Hyperparameter Optimization of Regression Model for Electrical Load Forecasting During the COVID-19 Pandemic Lockdown Period

Saif Mohammed Al-azzawi[1] Mohanad A. Deif[2] Hani Attar[3]* Ayman Amer[3] Ahmed A. A. Solyman[4]

[1]Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Istanbul Gelisim University, Istanbul, Avcılar 34310, Turkey
[2]Department of Bioelectronics, Modern University of Technology and Information (MTI) University, Egypt
[3]Department of Energy Engineer, Zarqa University, Jordan
[4]Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Nişantaşı University, 34398 ˙Istanbul, Turkey
* Corresponding author's Email: hattar@zu.edu.jo

Abstract: Due to global lockdown policies implemented against COVID-19, there has been an impact on electricity consumption. Several countries have emphasized the significance of ensuring electricity supply security during the pandemic to maintain the livelihood of people. Accurate forecasting of electricity demand plays a crucial role in ensuring energy security across all nations; accordingly to achieve this objective, this study employs metaheuristics optimization algorithms to enhance the prediction model's operation, such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF), at an optimized level to minimize errors. Two metaheuristics optimization methods, Particle Swarm Optimization (PSO) and Genetic Algorithms (GA), are utilized. The suggested prediction models are trained using daily power usage data from three US urban regions. In terms of prediction accuracy, the findings show that KNN with PSO surpasses the other models. The COVID-19 pandemic reduced power usage by 20% relative to pre-pandemic levels.

Keywords: COVID-19, Metaheuristics optimization algorithms, Support vector machine, Particle swarm optimization, Genetic algorithms.

1. Introduction

Predicting electricity demand at the city level can help with power plant energy efficiency, resource planning, greenhouse gas emissions tracking, program evaluation, reserve requirements and system infrastructure analysis. Hence, knowing building energy consumption on a city-wide scale is crucial for promoting urban sustainability, reducing carbon emissions, and increasing global energy efficiency [1].

The consumption of electricity on a city-wide scale is influenced by temperature. The ambient temperature, population, and income are all significant drivers of energy usage in cities [2], which is mostly owing to the significant energy consumption necessary for heating and cooling buildings, which is heavily dependent on outdoor air temperature [3]. Since climate change causes frequent increase, intense, and long-lasting extreme weather events such as heat waves, evaluating temperature-sensitive city-scale power use has become critical. Climate change adaptation must be prioritized by academics, energy planners, and policymakers in order to develop successful climate change solutions [4], and to comprehend how to improve the preparedness of energy production and transmission infrastructure for high-demand situations in order to boost resilience and energy security in the context of climate change adaptation. Hou et al. [5] explored how rising ambient

temperatures influenced Shanghai's power usage [6], where they hypothesized that if present energy consumption patterns continue, the expected temperature increase will result in an increase in summer power demand and a decrease in winter demand [7].

Similarly, it was predicted by 2099 in California that rising air temperatures will demand up to 31% more transmission capacity and up to 38% more peak generating capacity [8]. A heat wave in California in August 2020 caused a power supply deficit owing to an increase in air conditioner consumption, resulting in rotating power outages for inhabitants. Following the heat wave disruptive event in California, policy makers highlighted the refinement of electricity demand forecasting by incorporating climate change, taking into account thrilling weather events and associated load effects, as the first recommended action in the response letter [9].

In addition to ambient climatic conditions, other circumstances, such as unanticipated public health crises, may influence the power demand of a metropolis. According to the research achievement in Brazil, the COVID-19 outbreak affected the power consumption patterns of Brazilians, resulting in a 7% to 20% decrease in energy demand, depending on the local economic structure, whereas the industrialized regions were less affected [10]. Similarly, a European study discovered that the strictness and intensity of lockdown measures has an effect on the energy consumption patterns of society. Power consumption profiles during the pandemic were comparable to pre-pandemic weekend profiles for the same period in 2019, but nations with less rigorous measures, such as Sweden, had a lower decline in electricity usage [11]. It is possible to track the real-time economic consequences of lockdowns through tracking the path of the power consumption. Switzerland's total electricity consumption decreased by 4.6%, including a 14.3% decrease in the canton of Ticino, where stricter curtailment limitations were enacted in addition to federal legislation [12].

The need for energy varies from season to season, every day, and every minute, as was previously described. The prediction of the necessary demand over the short-, medium-, and long-term prediction horizons is known as energy data forecasting. It is an educated forecast of how much energy will be required by homes, businesses, or other institutions. Data on energy the vast volumes of data about energy consumption are referred to as energy data. Direct measurements of energy use are included. (such as electricity load, PV generation, wind, gas, steam, heating load etc.). Energy is produced from various sources using various time scales and measurement units. The forecast that establishes the time series for a future time is obtained using this energy information, which is helpful and used as historical data to train the forecasting algorithms. The important thing to remember is that those accurate findings need information. The PV generation and electrical load datasets of residential usage have been considered in this research work to differentiate and assess the forecasting effectiveness over various energy data applications. Horizontal forecasting (time scale) The literature discusses various methods and models for forecasting energy data for both commercial and residential use. Although the types of forecasting horizons as shown in Table 1 used in the energy industries are not officially categorized, they can be grouped into four groups based on the length of the forecasting interval: very short-term forecasting (VSTF), short-term forecasting (STF), medium-term forecasting (MTF), and long-term forecasting (LTF) [13]. A new sort of forecasting known as VSTF primarily pertains to forecasting with a period of up to an hour in advance [14]. Demand sight's management system frequently uses STF, which takes a range of time into account up to a day in advance [15]. The daily operation and scheduling of the electricity and spot price calculation are the main applications for VSTF and STF, where the needed accuracy is much higher than a long-term forecast [16]. This type of forecasting is essential to ensure that the scarce electricity in developing nations is used more effectively [17]. MTF, which stands for "month ahead," is a term that has been used for years [18] to schedule maintenance and the development of the grid system. LTF, on the other hand, considers a time frame that ranges from months to years in the future when it comes to arranging power supply and resource planning. When it is important to predict power demand over a longer time, this form of forecasting is used to establish system design [16]. MTF and LTF are often only used to estimate peak loads since they frequently experience forecast mistakes over time [19] .historical load data, meteorological variables (humidity, temperature), the season/time of year, the day of the week/hour of the day, or even certain holidays or festivals, can all impact forecasting.

2. Related work

Table 2 provides information on relevant work conducted in different countries and the prediction models used, which demonstrates that there are two sorts of energy demand forecast methods: physical simulation models and data-driven models. Data-driven models are more widely utilized since physical

Table 1. Forecasting horizons

Forecasting horizons	Time scale
Very short-term	5min-1h
Short-term	1h- 24h
Medium-term	24h- weeks
Long-term	months-year

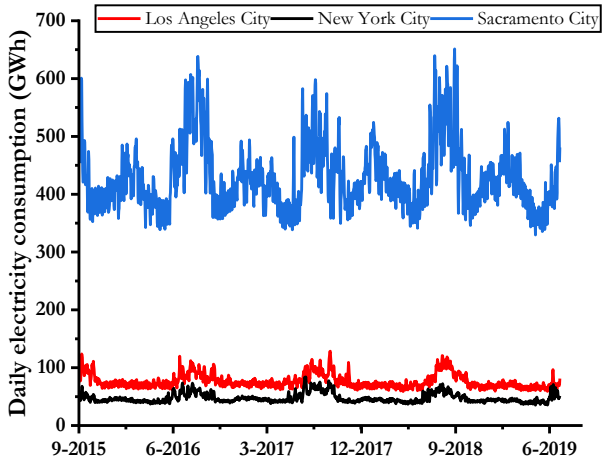


Figure. 1 Average hourly electricity usage in three metropolitan areas

simulation models need the consideration of several external parameters, which can be difficult to obtain. Some data-driven regression models, on the other hand, nevertheless need enormous data gathering of associated elements. Furthermore, the majority of data-driven models employ deep learning or machine learning techniques. Because single models have limitations, several researchers have adopted hybrid models for predictions in recent years.

The literature review reveals that the focus of many scholars was solely on the accuracy of prediction models and not on their stability. Moreover, in electricity demand prediction studies, the impact of significant global events such as COVID-19 was overlooked, and the factors considered, such as weather, may not be pertinent in the current period; consequently, these models may not be relevant during such global events. The study's noteworthy contributions are as follows:

(1) The recommendation to apply a hybrid model to forecast daily electricity consumption during the COVID-19 pandemic.
(2) The suggested model is compared to reference models in terms of accuracy and stability of prediction.
(3) Prediction is affected by the denoising process and optimizer.
(4) The usefulness of COVID-19-related parameters as prediction model inputs is addressed.

3. Materials and methods

3.1 Materials

The research paper [6] provided the dataset used in this study, which encompasses ambient temperature and city-level electricity usage data from three US metropolitan areas, namely New York (NY), Los Angeles (LA), and Sacramento (Sac), spanning from July 2015 to September 2020. The national oceanic and atmospheric administration (NOAA) provided data on ambient temperature, while the energy information administration (EIA) provided statistics on energy use at the municipal level. The raw electricity consumption in the three cities is plotted in Fig. 1.

3.2 Methods

Fig. 2 depicts the overall methodology employed to forecast city-scale daily electricity usage, comprising five phases as follows:

Phase 1: Data preparation, phase 2: Prediction phase, phase 3: Hyper-parameter optimization, phase 4: performance evaluation, Phase 5: Benchmark models and phase 6: predict energy consumption during the pandemic. Each phase is explained in the next subsections

Phase 1: Data preparation

The primary dataset was divided into power data including (date, time, electricity load, and day Type (working day / non-working day)) and weather data including (date, time, ambient temperature, and humidity), therefore, the power and weather data were merged dataset to appropriate for study phases.

The merged dataset now includes a new feature, which comprises the summation of the daily electricity load, the daily demand peaks, and the average ambient temperature, calculated for each day in the dataset. The measurement with missing values were eliminated, and then the dataset of the remaining measurements was explored to check the variables scale. Eventually, the preparation and preprocessing of the dataset assures that 80% of the dataset was randomly picked for prediction model training, while the remaining 20% was used for model validation testing.

Phase 2: Prediction phase

Four models were created to anticipate the daily

Table 2. Studies on forecasting energy usage in various countries

References	Forecasting target	Forecasting Model
[20]	Iran's demand for energy	A hybrid model combines Bayesian approach and scenario analysis
[21]	Demand for energy in Ireland	Covariance matrix adaptation evolutionary strategy
[22]	Electricity demand in India	Long Short-Term Memory network (LSTM)
[23]	Demand for electricity in New Singapore and South Wales	▪ Variational mode decomposition ▪ Support vector machine ▪ Salp Swarm Algorithm (SSA)
[24]	Natural gas demand in Germany	Functional autoregressive with convolutional neural network
[25]	Energy demand in China	autoregressive distributed lag mixed data sample
[26]	Residential natural gas demand	▪ Linear regression ▪ Kernel machine with memory ▪ Kernel machine ▪ Two-reservoir model ▪ Two-reservoir model with nonlinear memory ▪ Two-reservoir model with linear memory
[27]	Need for energy in Basilicata and Italy	Regression analysis
[28]	Load demand	▪ Harris hawk's optimization ▪ Stationary wavelet packet transform ▪ Feed-forward neural network
[29]	Energy demand	▪ Autoregressive integrated moving average model ▪ Artificial neural network ▪ Support vector machine
[30]	Building energy demand	Engineering simulation
[31]	Heating demand	▪ Artificial neural network with an online learning method
[32]	Electricity demand	▪ Artificial neural network ▪ Autoregressive integrated moving average model ▪ Multivariate adaptive regression spline
[33]	Energy load	variational mode decomposition with LSTM
[34]	Natural gas demand	▪ Autoregressive integrated moving average model ▪ Artificial neural network ▪ Extreme learning machine
[35]	Electricity demand	▪ Adaptive Fourier decomposition ▪ Support vector machine ▪ Fast Fourier transform
[36]	HVAC system energy demand	▪ Fuzzy with neural network
[37]	Fans Electricity demand	▪ Artificial neural network
[38]	Electricity demand	▪ Cyclic behavior ▪ Piecewise aggregated approximation ▪ Piecewise interpolation

power usage in New York, Los Angeles, and Sacramento. The forecasting models belong to two different modeling techniques: Time series forecasting techniques including autoregressive integrated moving average (ARIMA) model, and supervised machine learning forecasting (SMLF) techniques including support vector machine (SVM), K-nearest neighbors (KNN), and random forest (RF). The details of the six forecasting models are illustrated in the following sub-sections I and II.

I. Time series forecasting techniques

The ARIMA model is a well-known linear model

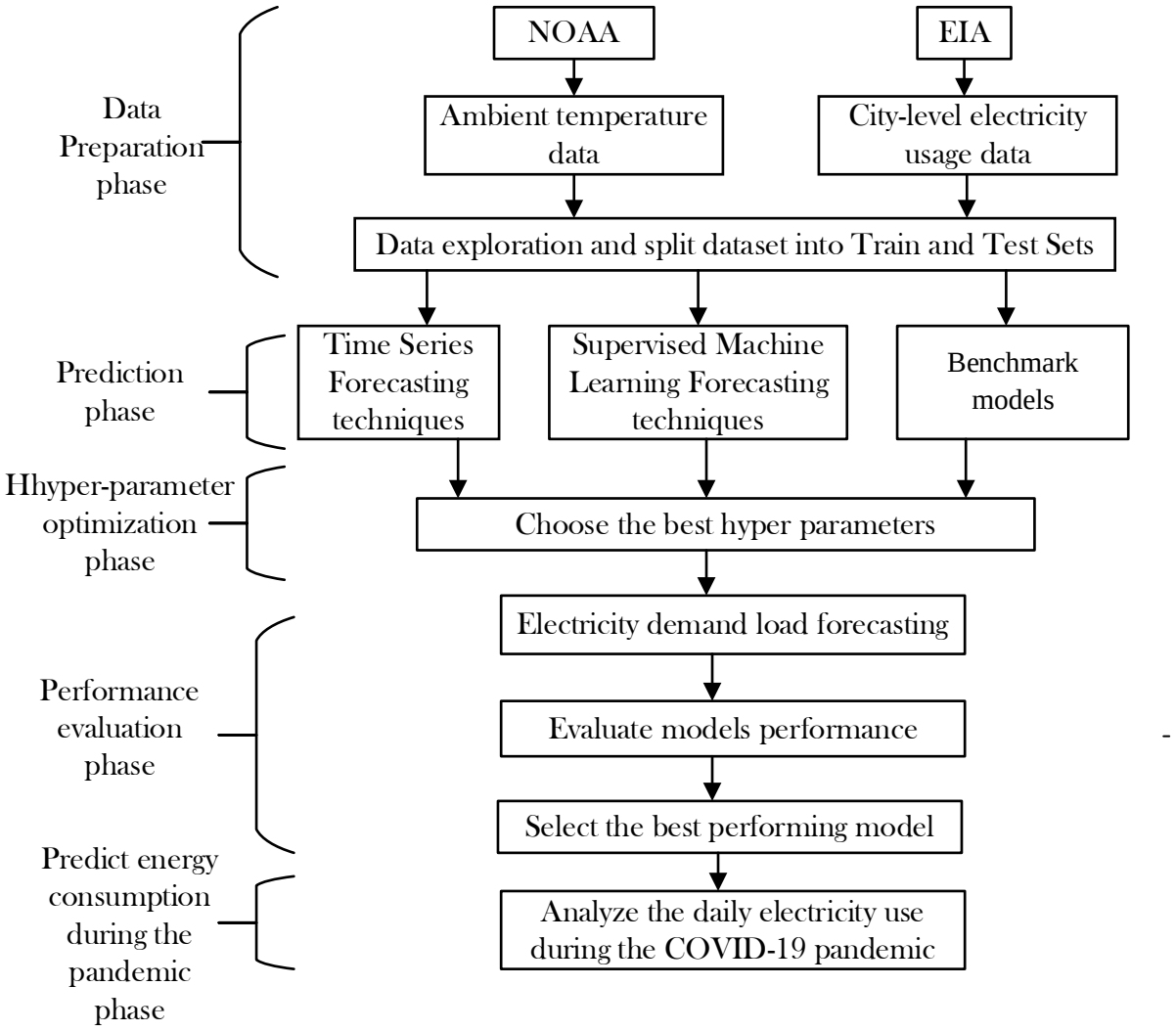


Figure. 2 Overall prediction methodology

for predicting univariate time series forecast [39], whereas the strategy is predicated on the assumption that time series may be separated into current and prior values, as well as random errors. ARIMA consists of three components: AR (p) is an additive linear function of p prior observations, MA (q) is a moving average that compensates for q random errors, and d is an integer that helps turn a series into a stationary one. The ARIMA (p, q, d) model is represented by the following equation:

$$\Delta^d y(t) = c + \sum_{j=1}^{p} \alpha_j \times y(t-j) + \epsilon(t) + \sum_{j=1}^{\bar{q}} \beta_j \times \epsilon(t-j) \quad (1)$$

The symbol Δ, which represents the operator $(1 - B)$, where B is referred to as the Backward operator, and $By(t)$ stands for the observation data $y(t-1)$ at time t, while $y(t)$ characterizes the current observation. The model consists of a constant term c, auto-regressive parameters $\alpha_1, \ldots, \alpha_p$, moving average coefficients $\beta_1, \ldots, \beta_q$, and white noise $\epsilon(t)$ at time t. To fit an ARIMA model, the following four steps need to be taken:

Step 1 involves identifying the ARIMA (p, d, q) structure, followed by parameter estimation in step 2, diagnostic checks on the estimated residuals in step 3, and finally, forecasting future values based on known data in step 4.

Box and Jenkins (1976) use the data's partial auto-correlation function (PACF) and auto-correlation function (ACF) to establish the q and p parameters of the ARIMA model.

II. Supervised machine learning forecasting techniques support vector machine (SVM)

To achieve linear separability, SVM algorithms translate data points between low-dimensional to high-dimensional spaces. If *n* data points exist (n is the predicted size of the sample,) the SVM objective function is as follows [40-42]:

$$\arg \min_{\mathbf{w}} \left\{ \frac{1}{n} \sum_{i=1}^{n} \mathrm{m} \ \{0, 1 - y_i f(x_i)\} + C\mathbf{w}^T\mathbf{w} \right\} \quad (2)$$

The error term penalty parameter is denoted as C, is a critical hyper-parameter in all SVM models, and $\mathbf{w}$ represents the normalization vector. Additionally, the choice of kernel function $f(x)$ is an essential hyper-parameter that determines the likeness measurement among, x_i and x_j data points. There are several types of kernel functions available in SVM models, including radial basis function (RBF), sigmoid kernels, polynomial, and linear. The various kernel functions can be expressed as [43]:

1 Linear kernel:

$$f(x) = x_i^T x_j \tag{3}$$

2 Polynomial kernel:

$$f(x) = \left(\gamma x_i^T x_j + r\right)^d \tag{4}$$

3 RBF kernel:

$$f(x) = \exp\left(-\gamma \|x - x'\|^2\right) \tag{5}$$

4 Sigmoid kernel:

$$f(x) = \left(\tanh\left(\gamma x_i^T x_j + r\right)\right) \tag{6}$$

After selecting a kernel type, various other hyperparameters must be adjusted, as demonstrated by the kernel function equations. When the kernel type is set to polynomial, RBF, or sigmoid, the coefficient γ, referred to as 'gamma' in sklearn, is the conditional hyperparameter of the kernel type hyperparameter. Additionally, the polynomial and sigmoid kernels have a conditional hyperparameter r, which is specified by 'coef0' in sklearn.

In addition, the polynomial kernel function contains an extra conditional hyperparameter, d, which represents the polynomial kernel function's degree. Another hyperparameter in SVR models is 'epsilon,' which specifies the distance error of the loss function [44].

Random forest (RF)

To briefly explain the RF algorithm, T decision trees are independently constructed from a training dataset with n variables and m samples (features). Each decision tree t^{th} model is built on the t^{th} bootstrap sample set from the original dataset, and at each inner node, a random selection of n' variables (where n' is much smaller than n) is used to determine the best partition. The trees are built without pruning and have the maximum depth. The prediction value y is obtained from each tree, and the last forecast value is determined by combining the results from all T trees in the forest. This is how the RF prediction is obtained:

$$\hat{y} = \frac{1}{T}\sum_{i=1}^{T} \hat{f}_i(x) \tag{7}$$

When building a tree, an additional tuning parameter called mtry, mtry applied to specify the number of variables used for node splitting at each iteration during the creation of a tree, where it is preferred that $1 <$ mtry $< n$ in order to decrease the computational time. By choosing mtry $< n$, the objective is to reduce the computational burden.

K-nearest neighbors (KNN)

By utilizing a similarity measure, specifically a distance function [45], this algorithm can make predictions for new data points. Essentially, the algorithm assigns a value to a new data sample based on its proximity to the training set points. To do this, the approach use a metric such as Euclidean, Manhattan, or Minkowski distance to calculate the distance between each sample in the test set and every sample in the training set. The K nearest samples of training data are then picked. The KNN prediction for regression is the mean of the outcomes from the K nearest neighbors.

$$y = \frac{1}{K}\sum_{i=1}^{k} y_i \tag{8}$$

The i th case in the example sample is represented by y_i, while y represents the prediction (outcome) for the query point.

Phase 3: Hyper-parameter optimization

During the model creation phase of machine learning, optimization techniques may be used to quickly search the space of hyper-parameters in order to identify the optimal hyper-parameters for the regressor models. The hyper-parameter optimization procedure typically comprises four essential components: an estimator with its objective function, a search space (also known as a configuration space), a search or optimization method to locate the optimal hyper-parameter combinations, and an evaluation function to compare the performance of various hyper-parameter configurations. According to references [46, 47], the objective of a hyper-parameter optimization work is typically to generate the best practicable results.

$$x^* = \arg\min_{x \in X} f(x) \tag{9}$$

The objective function $f(x)$, which can represent the root mean squared error (RMSE) or the error rate, is to be minimized. The hyper-parameter configuration x^* that results in the optimal value of $f(x)$ can assume any search space value X. Fig. 3 illustrates the primary steps of the hyper-parameter optimization process. Two optimization algorithms (PSO and GA) were created to optimize the hyperparameters of each machine learning model. The pseudo code of PSO and GA have shown in Algorithm 1 and Algorithm 2. Table 3 shows search space for the hyper-parameters of machine learning forecasting models.

Algorithm 1 : Pseudo code of PSO

/* Initialize all parameters for PSO */

Computation maxtime: $= Tp_{\max}, t: = 0$

Number of particle-patterns: $= m, 2 \leq m \in \boldsymbol{N}^1$;

Particle-patterns initial solution: $= \boldsymbol{P}_i^0$

Particle-patterns initial position: $= \boldsymbol{x}_{ij}^0$;

Particles initial velocity:= $\boldsymbol{v}_{ij}^0$;

PSO parameter: $= \omega, 0 < \omega \in \boldsymbol{R}^1$;

PSO parameter: $= C_1, 0 < C_1 \in \boldsymbol{R}^1$;

PSO parameter: $= C_2, 0 < C_2 \in \boldsymbol{R}^1$;

/* Start PSO */

Evaluate $(\boldsymbol{G}^0, \boldsymbol{P}^0)$;

while $t < Tp_{\max}$ do

/* Update velocities and positions */

$\boldsymbol{v}_{ij}^{t+1} = \omega \cdot \boldsymbol{v}_{ij}^t$

$+ C_1 \cdot \text{rand}() \cdot (\text{best}\left(P_{ij}^t\right) - x_{ij}^t)$

$+ C_2 \cdot \text{rand}() \cdot (\text{best}\,(G^t) - x_{ij}^t)$;

$\boldsymbol{x}_{ij}^{t+1} = \boldsymbol{x}_{ij}^t + \boldsymbol{v}_{ij}^{t+1}$;

/* if fitness value is increased, a new solution will be accepted. */

Update_Solutions $(\boldsymbol{G}^t, \boldsymbol{P}^t)$;

$t = t + 1$;

end while

Update_Solutions $(\boldsymbol{G}^t, \boldsymbol{P}^t)$;

return Best found pattern of particles as solution;

Algorithm 2 : Pseudo code of GA

/* Initialize all parameters for DGA */

Computation maxtime: $= Tg_{\max}, t: = 0$;

Number of islands: $tn, 1 \leq n \in \boldsymbol{N}^1$;

initial solution: $= \boldsymbol{P}_i^0$;

/* Start DGA //

Evaluate $(\boldsymbol{G}^0, \boldsymbol{P}^0)$;

while $t < Tg_{\max}$ do

for all islands do

Selection ();

Crossover ();

Mutation ();

end for

$t = t + 1$;

end while

Update_Solutions $(\boldsymbol{G}^t, \boldsymbol{P}^t)$;

return Best found pattern of particles as solution;

Phase 4: Performance evaluation

The RMSE indicators are used to verify the effectiveness of the forecasting model established in this paper.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - x_i)^2} \quad (10)$$

In the formula, y_i is the genuine situation value, x_i is the predicted value, and n is the predicted size of the sample.

Phase 5: Benchmark models

To highlight the advantages of the proposed model, this paper defines five classifiers models reported in research [6] as the benchmark model for

Table 3. Search space for the hyper-parameters of regressor models

MACHINE LEARNING Model	Hyper-parameter	Search Space
RF Regressor	n estimators	[10,100]
	max depth	[5,50]
	min samples split	[2,11]
	min samples leaf	[1,11]
	criterion	['mse', 'mae']
	max features	[1,13]
SVM Regressor	C	[0.1,50]
	kernel	['linear', 'poly', 'rbf', 'sigmoid']
	epsilon	[0.001,1]
KNN Regressor	n neighbors	[1,20]

Table 4. Reasons for selecting benchmark models

Model name	Reason for being selected
5-Parameter	which was initially proposed by ASHRAE in the 1990s. The major reason why model are widely used is their interoperability [48].
Degree Hour	is one of the most well-known methods used in the heating, ventilating, and air conditioning (HVAC) industry to estimate heating and cooling energy requirements. Because of its significance in this field, the heating and cooling degree day was used to determine the U.S. climate zone and has been widely used as a proxy variable to quantify the influence of climate change on electricity demand.
Decomposed model	the decomposed model can decouple the effects of different factors (weather-related temperature-dependent load and seasonal time-dependent periodical load), providing us a unique opportunity to observe the impact of an unexpected public health event on city-level demand.
lightGBM	Combining multiple trees usually outperforms a single tree in terms of model accuracy and robustness [49].
NN	It can be used to solve many different tasks [50]

comparison. The reasons for choosing these models are shown in Table 4.

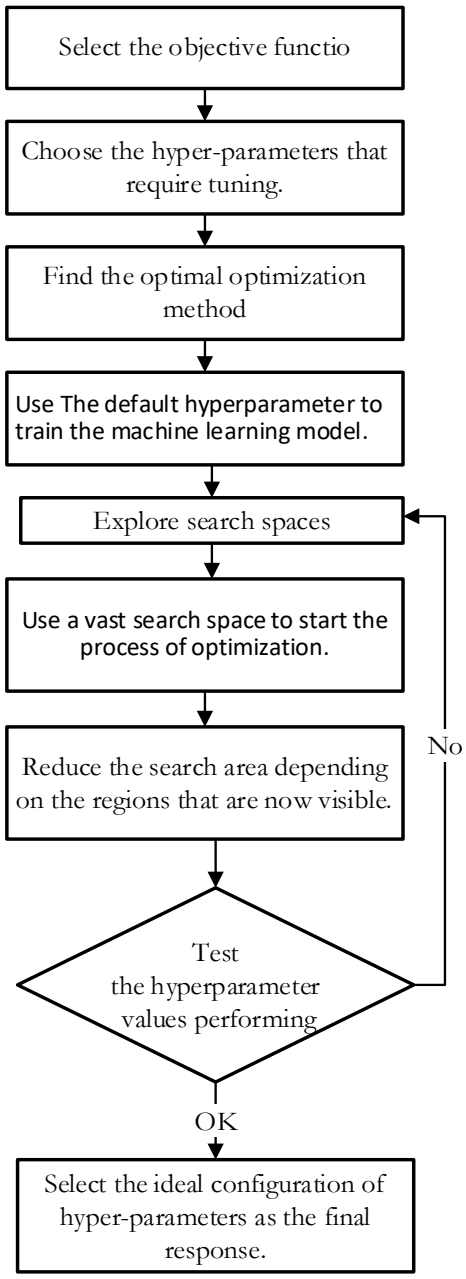


Figure. 3 The flow chart for the main process of Hyper-parameter optimization

Phase 6: Predict energy consumption during the pandemic

To address two practical issues, the most efficient models were selected. 1) how a city's daily power consumption would be affected by the ambient air temperature, particularly during a heat wave; and 2) how a city's daily power consumption would be affected by unforeseen public health concerns, in this case the COVID-19 pandemic. In this study, we anticipate the daily power consumption of a city, including the energy consumed by buildings, transportation (e.g., public electric buses/trains, electric automobiles), industries, and other public services inside the city and the surrounding rural

Table 5. The initial hyperparameters values for PSO

PSO Parameter	Parameter values
Number of particles	10
Number of generations	5
Maximum velocity of each particle	None
Coefficient of Local acceleration (c1)	1.5
Coefficient of global acceleration (c2)	1.5

Table 6. The initial hyperparameters values for GA

GA Parameter	Parameter values
population_size	10
gene_mutation	0.10
gene_crossover	0.5
generations_number	1.5

Table 7. Hyperparameter values after GA tuning

Machine learning model	**Hyper-parameter**	Los Angeles	Sacramento	New York
RF Regressor	n estimators	24	32	80
	min samples split	2		
	max depth	50		
	min samples leaf	5	2	1
	criterion	'mse'	'mae'	'mse'
	max features	10		
SVM Regressor	C	30		
	kernel	'poly'		
	epsilon	0.001		0.01
KNN Regressor	n neighbors	6	15	10

Table 8. Hyperparameter values after PSO tuning

Machine learning model	**Hyper-parameter**	Los Angeles	Sacramento	New York
RF Regressor	n estimators	22	36	82
	min samples split	6		
	max depth	40		
	min samples leaf	1	2	1
	criterion	'mse'		
	max features	8		
SVM Regressor	C	50		
	kernel	rbf		
	epsilon	0.001		0.01
KNN Regressor	n neighbors	6	5	7

regions.

4. Implementation

The experiments were implemented using Python 3.8 on an HP system with an i5 Intel Core processor working at 1.60 GHz and Windows10 operating system. Several Python open-source modules are used to analyze the machine learning models and optimization strategies involved (sklearn, optunity, hyperband, DEAP) [51-53]

5. Results

Four studies were conducted to predict the daily power use in Los Angeles, New York and Sacramento. The initial experiment evaluated the performance of supervised machine learning forecasting techniques (SVM, RF, and KNN) using default hyperparameter values. The second experiment involved utilizing optimization algorithms (PSO and GA) to optimize the hyperparameters of each machine learning model, followed by an evaluation of their performance. The best performing model from experiments 1 and 2 was compared to the time series forecasting technique (ARIMA model). Lastly, the influence of the COVID-19 pandemic on the city's daily electricity consumption was analyzed.

The initial hyperparameters values for PSO and GA algorithms used in adopted experiments are shown in Tables 5 and 6, which utilized the supplied default settings based on the Python scikit-learn library package [54].

Tables 7 and 8 show the best hyperparameters obtained with different optimization techniques (GA and PSO) for machine learning models. Our experiment employed 5-fold cross-validation to establish the optimal model parameters for Sacramento, Los Angeles, and New York City. Convergence was reached long before 100 iterations, and we stopped when the variance between iterations was less than 0.5%.

In contrast, Figs. 8 and 9 illustrate the effectiveness of each optimization approach applied to KNN, SVM, and RF regressors evaluated on the MNIST dataset after a thorough optimization phase. The four-year period covered by the study's data was from July 2015 to June 2019. We utilized the first three years to train our model, reserving the fourth year for validation. To avoid biased electricity usage behavior caused by the COVID-19 curtailment situation, the data from the year 2020 was not included.

Tables 9, 10, and 11 demonstrate that optimizing hyperparameters with optimization techniques like GA and PSO can significantly improve the regression

Table 9. Performance results of applying regression models with default hyperparameter values

	RMSE (GWh)		
	RF	SVM	KNN
Los Angeles	8.383	8.850	7.4786
Sacramento	5.343	6.4569	4.9525
New York	25.221	31.476	24.565

Table 10. Performance results of applying GA optimization algorithm to the regression models

	RMSE (GWh)		
	RF-GA	SVM-GA	KNN-GA
Los Angeles	7.7128	8.1427	6.8804
Sacramento	4.9158	5.940	4.5563
New York	23.204	28.958	22.600

Table 11. Evaluation of the performance of the PSO method when applied to regressor models

	RMSE (GWh)		
	RF-PSO	SVM-PSO	KNN-PSO
Los Angeles	6.7067	7.0806	5.9829
Sacramento	4.2746	5.1655	3.9620
New York	20.17	25.181	19.652

Table 12. Optimal hyperparameter values for ARIMA model after optimization process

City	ARIMA(p,q,d)
Los Angeles	(1,1,5)
Sacramento	(5,3,5)
New York	(1,3,5)

performance of machine learning models, which is generally observable.

In terms of performance while tuning hyperparameters of machine learning models, PSO outperformed GA, with the RMSE value for all regression machine learning models lowered by over 20% when tweaking hyperparameters values using PSO. While utilizing GA, it is lowered by over 8%.

The KNN model beat the RF and SVM models among the three regression models that were tweaked using PSO in all three metropolitan regions. KNN performs the best in Sacramento (RMSE=3.96). but performs the worst in New York (RMSE=19.65).

Additionally, the KNN regressor is easier to implement because it has lowest number of Hyper-parameter (one hyper-parameter) compared with three hyper-parameters for SVM regressor and six hyper-parameters for RF regressor. and therefore, the computational time of optimization step for KNN regression model is often much lower than another model.

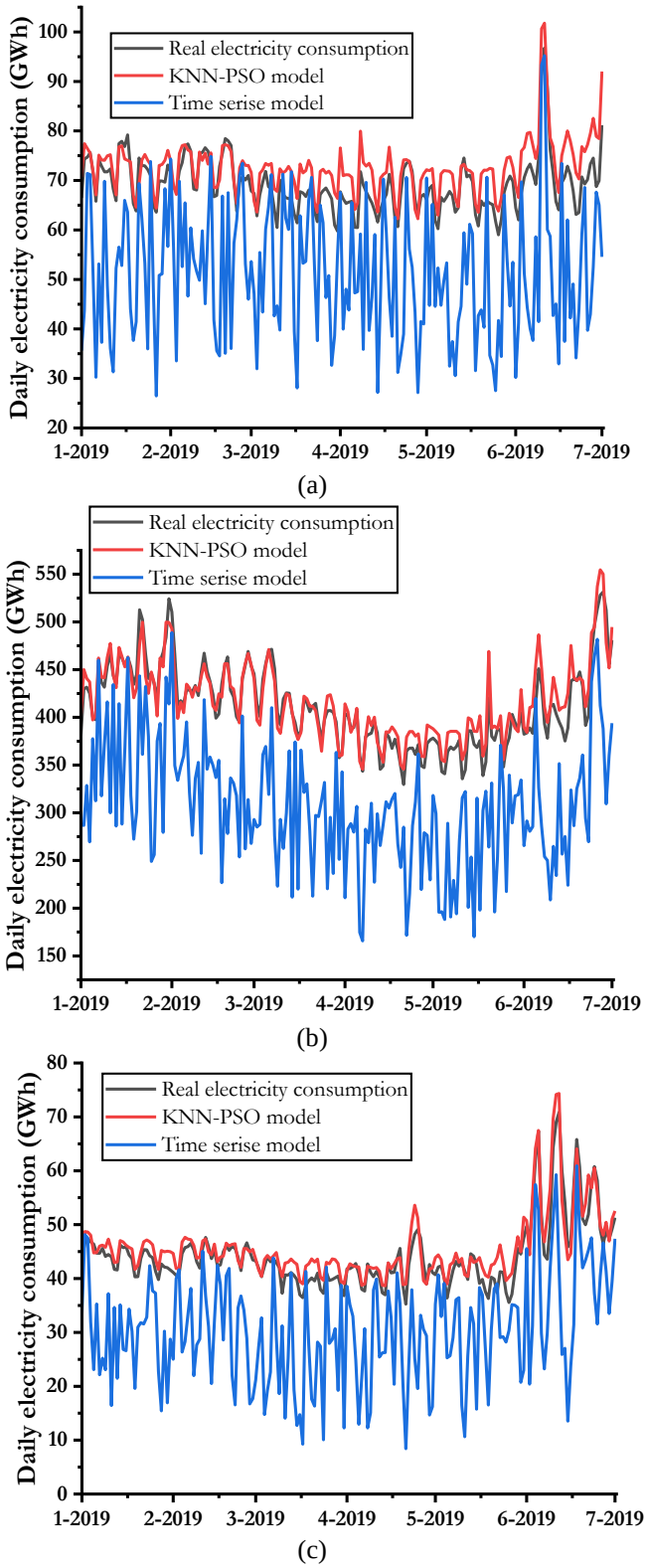


Figure. 4 Prediction results for KNN-PSO model and ARIMA model on the test dataset (Jan.2019–June 2019): (a) Los Angeles, (b) New York, and (c) Sacramento city

When the number of hyper-parameters increases, the size and complexity of the search space and the overall objective function evaluation time increase exponentially [55-57]. As a result, it is required to improve existing hyper-parameter optimization algorithms in order to lessen the impact of huge search spaces on execution time.

Based on the previous experiment, the best-performing forecasting model achieved by using KNN with the PSO method will compare with the ARIMA model that belongs to Time series forecasting techniques. The comparison results have recorded in Table 1.

The auto.arima() function [58, 59] was applied to determine the best hyper-parameters values ARIMA(p, q, d) for the ARIMA model. The optimal parameters results and RMSE value for all the three areas are shown in Table 12.

Similar to the previous experiment, the two models have trained with the first three years (from 2015 to 2018) and kept the last year (2019) for forecasting. The results for daily electricity consumption forecast in 2019 are shown in Fig. 4.

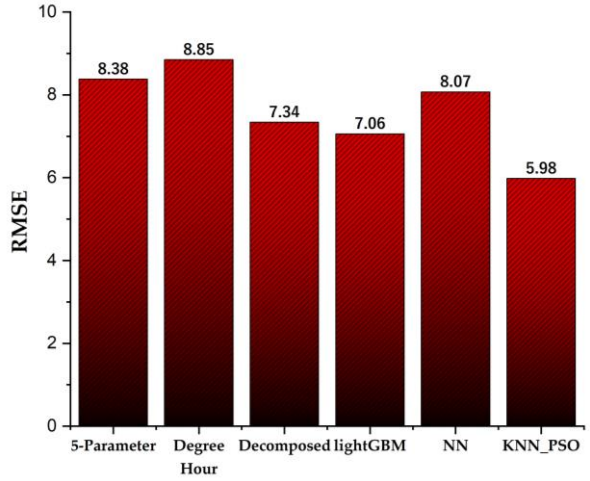
It can be observed from Fig. 4, the forecasting curve for KNN-PSO model more fitted with actual daily electricity consumption curve, while there are gap between the forecasting curve for ARIMA model and actual daily electricity consumption curve. Because the KNN-PSO model is more resilient in coping with missing data, this may be understood. Time-series modeling, on the other hand, employs the consecutive ordering of incoming data to describe temporal information. If any data is missing, data imputation is required, which augments another degree of complication to the data pretreatment process. Tabular data models, on the other hand, encapsulate temporal data using supplementary characteristics (such as holiday day and day type) and hence do not suffer if any data is absent.

In Table 13, the RMSE is used to assess the performance of the KNN-PSO and ARIMA models. The KNN-PSO surpasses the ARIMA because it has the lowest RMSE across all metropolitan regions. Among the examined models, the suggested model (KNN-PSO) has the greatest prediction accuracy.

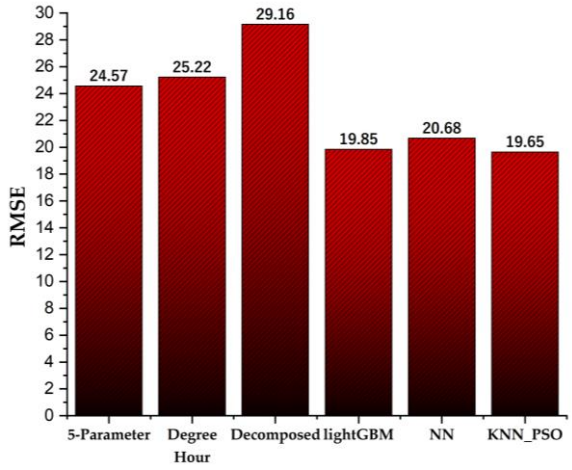
The prediction accuracy for the proposed model (KNN- PSO) with benchmark models is shown in Fig. 5. The KNN- PSO model outperforms the 5-parameter, degree hour, decomposed, and other baseline machine learning models (lightGBM and NN) in all three metropolitan areas. It can be

Table 13. Performance evaluation of KNN-PSO and ARIMA models

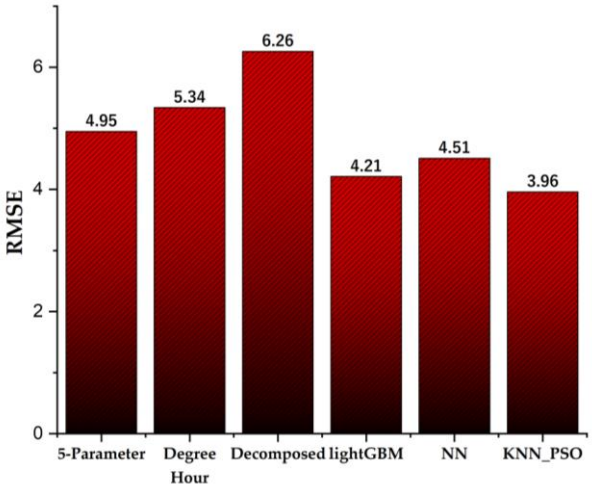
City	ARIMA	KNN-PSO
Los Angeles	25.142	5.982958
Sacramento	16.472	3.962065
New York	40.232	19.65249



(a)



(b)



(c)

Figure. 5 Prediction results for the KNN-PSO model and benchmark models: (a) Los Angeles, (b)New York, and (c) Sacramento City

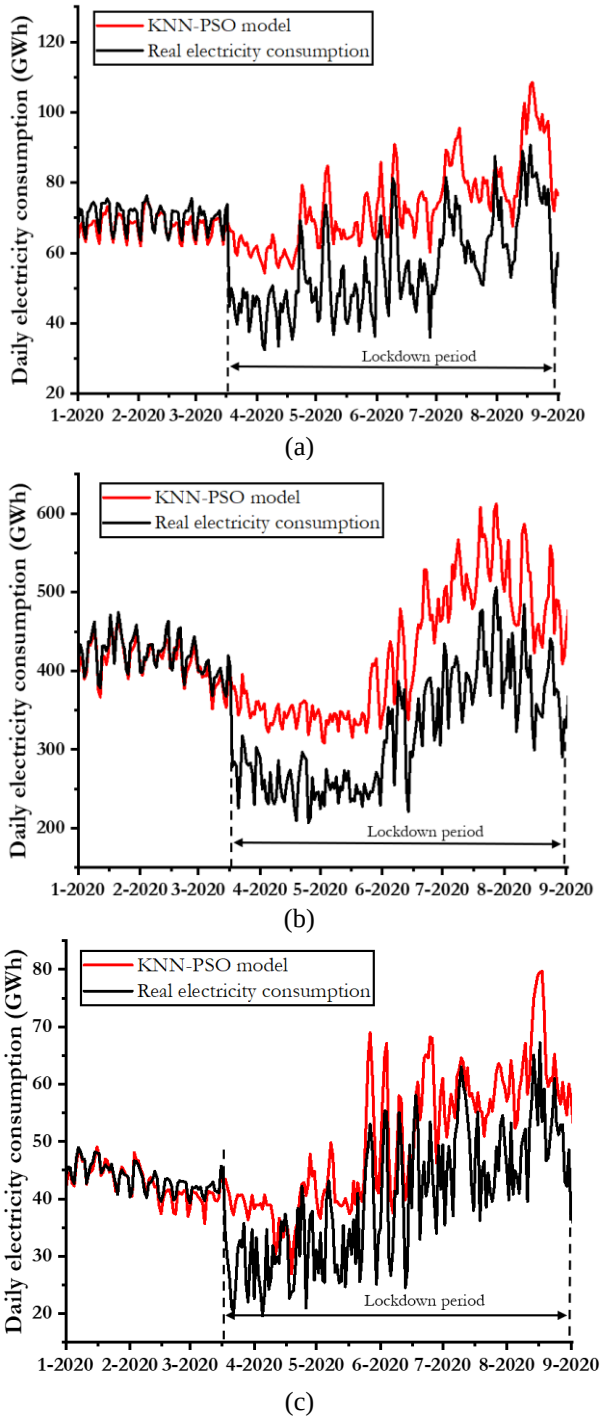


Figure. 6 KNN-PSO model predictions results before and after the COVID-19 lockdown: (a) Los Angeles, (b)New York, and (c) Sacramento city

concluded that the proposed model has the highest prediction accuracy among the evaluated models.

The third question was how government and individual measures (e.g., stay-at-home, corporate closure, reduced operation, and shelter-in-place) to mitigate the COVID-19 outbreak in U.S. cities impacted city-level power usage. City-scale electricity use can indicate shutdown severity in real time.

The KNN-PSO model was chosen to assess the effect of COVID-19 on city-scale consumption of power because of its superior accuracy when compared to other models investigated. It uses the period of the pre-pandemic from July 2015 to March 2020 to train the KNN-PSO model, and then employs the learned model to forecast power consumption after the pandemic. Our hypothesis is that if energy consumption patterns altered as a result of the COVID-19 pandemic, the model trained with data acquired before the pandemic would be incapable of reliably forecasting energy consumption after the pandemic. Fig. 6 displays the anticipated results.

Fig. 6 illustrates that the forecasted electricity consumption used after the lockdown period was over than the real electricity usage. In contrast, previous to the lockdown period, power demand predictions were correct (dotted orange line). This disparity between expected and actual electricity use after mid-March 2020 demonstrates that the COVID-19 epidemic altered electricity demand patterns. During the COVID-19 pandemic, daily power usage in the urban zones of New York, Los Angeles, and Sacramento declined by 10 to 20 percent compared to the month prior to the outbreak.

6. Conclusion

This study utilizes three distinct machine learning models (RF, SVM, and KNN) to forecast the power consumption of three metropolitan regions in the United States: New York, Los Angeles, and Sacramento. To improve the accuracy of the predictions, hyperparameter tuning was utilized to optimize the performance of each machine learning model. The proposed study found that the supervised machine learning forecasting techniques (SVM, RF, and KNN) performed better than two metaheuristics optimization approaches, PSO and GA. The performance of each technique using both default and tuned hyperparameters values was evaluated. Additionally, the obtained results in this paper were compared to a Time series forecasting technique, specifically the ARIMA model. Before and after hyperparameter tuning, the KNN regressor model had the highest degree of accuracy, with PSO optimization delivering the lowest RMSE score. Lastly, the achieved data in the proposed work found that during the COVID-19 pandemic lockdown period in 2020, power consumption in the three metropolitan regions declined from 2% to 12% compared to the same month in 2019.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization, Saif Mohammed Al-azzawi and Mohanad A. Deif; methodology, Saif Mohammed Al-azzawi ; software, Mohanad Dief and Hani Attar; validation, Ahmed Solyman, Ayman Amer, and Saif Mohammed Al-azzawi; formal analysis, Ayman Amer and Ahmed Solyman; investigation, Mohanad Dief; resources, Saif Mohammed Al-azzawi, mohanad Dief, Hani Attar, Ayman Amer, and Ahmed Solyman; data curation, Mohanad Dief; writing—original draft preparation, Saif Mohammed Al-azzawi and Mohanad Dief; writing—review and editing, Hani Attar; visualization, Mohanad Dief and Ahmed Solyman; supervision, Hani Attar and Ahmed Solyman; project administration, Hani Attar and Ahmed Solyman; funding acquisition, Hani Attar and Ayman Amer.

References

[1] C. E. Kontokosta and C. Tull, "A data-driven predictive model of city-scale energy use in buildings", *Appl. Energy*, Vol. 197, pp. 303–317, 2017.

[2] O. Deschênes and M. Greenstone, "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US", *Am. Econ. J. Appl. Econ.*, Vol. 3, No. 4, pp. 152–185, 2011.

[3] J. R. V. Canteli, S. Ulyanin, J. Kämpf, and Z. Nagy, "Fusing TensorFlow with building energy simulation for intelligent energy management in smart cities", *Sustain. Cities Soc.*, Vol. 45, pp. 243–257, 2019.

[4] S. Davoudi, E. Brooks, and A. Mehmood, "Evolutionary resilience and strategies for climate adaptation", *Plan. Pract. & Res.*, Vol. 28, No. 3, pp. 307–322, 2013.

[5] S. Sjostrand, A. Widerstrom, A. R. Ahlgren, and M. Cinthio, "Design and fabrication of a conceptual arterial ultrasound phantom capable of exhibiting longitudinal wall movement", *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, Vol. 64, No. 1, pp. 11–18, 2017, doi: 10.1109/TUFFC.2016.2597246.

[6] Z. Wang, T. Hong, H. Li, and M. A. Piette, "Predicting city-scale daily electricity consumption using data-driven models", *Adv. Appl. Energy*, Vol. 2, p. 100025, 2021.

[7] H. Y. Ling, M. H. Zhen, D. G. Tao, and S. Jun, "Influences of urban temperature on the electricity consumption of Shanghai", *Adv. Clim. Chang. Res.*, Vol. 5, No. 2, pp. 74–80, 2014.

[8] J. A. Sathaye, L. L. Dale, P. H. Larsen, G. A. Fitts, K. Koy, S. M. Lewis, and A. F. P. D. Lucena, "Estimating impacts of warming temperatures on California's electricity system", *Glob. Environ. Chang.*, Vol. 23, No. 2, pp. 499–511, 2013.

[9] H. Lu, X. Ma, and M. Ma, "A hybrid multi-objective optimizer-based model for daily electricity demand prediction considering COVID-19", *Energy*, Vol. 219, p. 119568, 2021.

[10] M. Carvalho, D. D. M. Delgado, K. M. D. Lima, M. D. C. Cancela, C. A. D. Siqueira, and D. L. B. D. Souza, "Effects of the COVID-19 pandemic on the Brazilian electricity consumption patterns", *Int. J. Energy Res.*, Vol. 45, No. 2, pp. 3358–3364, 2021.

[11] A. Bahmanyar, A. Estebsari, and D. Ernst, "The impact of different COVID-19 containment measures on electricity consumption in Europe", *Energy Res. \& Soc. Sci.*, Vol. 68, p. 101683, 2020.

[12] B. Janzen and D. Radulescu, "Electricity use as a real-time indicator of the economic burden of the COVID-19-related lockdown: Evidence from Switzerland", *CESifo Econ. Stud.*, Vol. 66, No. 4, pp. 303–321, 2020.

[13] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network", In: *Proc. of 2017 51st Annual Conference on Information Sciences and Systems (CISS)*, 2017, pp. 1–6.

[14] J. W. Taylor, "An evaluation of methods for very short-term load forecasting using minute-by-minute British data", *Int. J. Forecast.*, Vol. 24, No. 4, pp. 645–658, 2008.

[15] F. Kaytez, M. C. Taplamacioglu, E. Cam, and F. Hardalac, "Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines", *Int. J. Electr. Power & Energy Syst.*, Vol. 67, pp. 431–438, 2015.

[16] P. H. Kuo and C. J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting", *Energies*, Vol. 11, No. 1, p. 213, 2018.

[17] T. Hong, M. Gui, M. E. Baran, and H. L. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions", in *Ieee Pes General Meeting*, 2010, pp. 1–8.

[18] M. Ghiassi, D. K. Zimbra, and H. Saidane, "Medium term system load forecasting with a dynamic artificial neural network model", *Electr. Power Syst. Res.*, Vol. 76, No. 5, pp. 302–316, 2006.

[19] B. J. Chen, M. W. Chang, and others, "Load forecasting using support vector machines: A

study on EUNITE competition 2001", *IEEE Trans. Power Syst.*, Vol. 19, No. 4, pp. 1821–1830, 2004.

[20] S. Ahmadi, A. H. Fakehi, M. Haddadi, S. H. Iranmanesh, and others, "A hybrid stochastic model based Bayesian approach for long term energy demand managements", *Energy Strateg. Rev.*, Vol. 28, p. 100462, 2020.

[21] K. Mason, J. Duggan, and E. Howley, "Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks", *Energy*, Vol. 155, pp. 705–720, 2018.

[22] J. Bedi and D. Toshniwal, "Deep learning framework to forecast electricity demand", *Appl. Energy*, Vol. 238, pp. 1312–1326, 2019.

[23] P. Jiang, R. Li, N. Liu, and Y. Gao, "A novel composite electricity demand forecasting framework by data processing and optimized support vector machine", *Appl. Energy*, Vol. 260, p. 114243, 2020.

[24] Y. Chen, X. Xu, and T. Koch, "Day-ahead high-resolution forecasting of natural gas demand and supply in Germany with a hybrid model", *Appl. Energy*, Vol. 262, p. 114486, 2020.

[25] Y. He and B. Lin, "Forecasting China's total energy demand and its structure using ADL-MIDAS model", *Energy*, Vol. 151, pp. 420–429, 2018.

[26] P. Potočnik, J. Šilc, G. Papa, and others, "A comparison of models for forecasting the residential natural gas demand of an urban area", *Energy*, Vol. 167, pp. 511–522, 2019.

[27] S. D. Leo, P. Caramuta, P. Curci, and C. Cosmi, "Regression analysis for energy demand projection: An application to TIMES-Basilicata and TIMES-Italy energy models", *Energy*, Vol. 196, p. 117058, 2020.

[28] U. B. Tayab, A. Zia, F. Yang, J. Lu, and M. Kashif, "Short-term load forecasting for microgrid energy management system using hybrid HHO-FNN model with best-basis stationary wavelet packet transform", *Energy*, Vol. 203, p. 117857, 2020.

[29] M. R. Kazemzadeh, A. Amjadian, and T. Amraee, "A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting", *Energy*, Vol. 204, p. 117948, 2020.

[30] D. Yu, "A two-step approach to forecasting city-wide building energy demand", *Energy Build.*, Vol. 160, pp. 1–9, 2018.

[31] F. Bünning, P. Heer, R. S. Smith, and J. Lygeros, "Improved day ahead heating demand forecasting by online correction methods", *Energy Build.*, Vol. 211, p. 109821, 2020.

[32] M. S. A. Musaylh, R. C. Deo, J. F. Adamowski, and Y. Li, "Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF Reanalysis atmospheric predictors in southeast Queensland, Australia", *Renew. Sustain. Energy Rev.*, Vol. 113, p. 109293, 2019.

[33] J. Bedi and D. Toshniwal, "Energy load time-series forecast using decomposition and autoencoder integrated memory network", *Appl. Soft Comput.*, Vol. 93, p. 106390, 2020.

[34] O. A. Karabiber and G. Xydis, "Forecasting day-ahead natural gas demand in Denmark", *J. Nat. Gas Sci. Eng.*, Vol. 76, p. 103193, 2020.

[35] R. Li, P. Jiang, H. Yang, and C. Li, "A novel hybrid forecasting scheme for electricity demand time series", *Sustain. Cities Soc.*, Vol. 55, p. 102036, 2020.

[36] R. Z. Homod, H. Togun, H. J. Abd, and K. S. M. Sahari, "A novel hybrid modelling structure fabricated by using Takagi-Sugeno fuzzy to forecast HVAC systems energy demand in real-time for Basra city", *Sustain. Cities Soc.*, Vol. 56, p. 102091, 2020.

[37] J. Runge, R. Zmeureanu, and M. L. Cam, "Hybrid short-term forecasting of the electric demand of supply fans using machine learning", *J. Build. Eng.*, Vol. 29, p. 101144, 2020.

[38] S. Williams and M. Short, "Electricity demand forecasting for decentralised energy management", *Energy Built Environ.*, Vol. 1, No. 2, pp. 178–186, 2020.

[39] M. A. Deif, A. A. A. Solyman, and R. E. Hammam, "ARIMA Model Estimation Based on Genetic Algorithm for COVID-19 Mortality Rates", *Int. J. Inf. Technol. Decis. Mak.*, Vol. 20, No. 6, pp. 1775–1798, 2021.

[40] M. A. Deif, A. A. A. Solyman, M. H. Alsharif, S. Jung, and E. Hwang, "A hybrid multi-objective optimizer-based SVM model for enhancing numerical weather prediction: a study for the Seoul metropolitan area", *Sustainability*, Vol. 14, No. 1, p. 296, 2021.

[41] R. E. Hammam, H. Attar, A. Amer, H. Issa, I. Vourganas, A. Solyman, P. Venu, M. R. Khosravi, and M. A. Deif, "Prediction of Wear Rates of UHMWPE Bearing in Hip Joint Prosthesis with Support Vector Model and Grey Wolf Optimization", *Wirel. Commun. Mob. Comput.*, Vol. 2022, 2022.

[42] M. A. Deif, R. E. Hammam, R. Hammam, and A. Solyman, "Adaptive Neuro-Fuzzy Inference System (ANFIS) for Rapid Diagnosis of COVID-19 Cases Based on Routine Blood

Tests", *Int. J. Intell. Eng. Syst.*, Vol. 14, No. 2, pp. 178–189, 2021, doi: 10.22266/ijies2021.0430.16.

[43] Q. I. Ahmed, H. Attar, A. Amer, M. A. Deif, and A. A. A. Solyman, "Development of a Hybrid Support Vector Machine with Grey Wolf Optimization Algorithm for Detection of the Solar Power Plants Anomalies", *Systems*, Vol. 11, No. 5, p. 237, 2023.

[44] M. A. Deif, R. E. Hammam., R. E. Hammam, and A. A. A. Solyman, "Gradient Boosting Machine Based on PSO for prediction of Leukemia after a Breast Cancer Diagnosis", *Int. J. Adv. Sci. Eng. Inf. Technol.*, Vol. 11, No. 12, pp. 508–515, 2021.

[45] M. A. Deif, R. E. Hammam, A. Solyman, M. H. Alsharif, and P. Uthansakul, "Automated Triage System for Intensive Care Admissions during the COVID-19 Pandemic Using Hybrid XGBoost-AHP Approach", *Sensors*, Vol. 21, No. 19, p. 6379, 2021.

[46] M. A. Deif, H. Attar, A. Amer, H. Issa, M. R. Khosravi, and A. A. A. Solyman, "A New Feature Selection Method Based on Hybrid Approach for Colorectal Cancer Histology Classification", *Wirel. Commun. Mob. Comput.*, Vol. 2022, 2022.

[47] M. A. Deif, H. Attar, A. Amer, I. A. Elhaty, M. R. Khosravi, and A. A. A. Solyman, "Diagnosis of Oral Squamous Cell Carcinoma Using Deep Neural Networks and Binary Particle Swarm Optimization on Histopathological Images: An AIoMT Approach", *Computational Intelligence and Neuroscience*, Vol. 2022, Article ID 6364102, 2022, doi: 10.1155/2022/6364102.

[48] E. M. O. Mokhtar and M. A. Deif, "Towards a Self-sustained House: Development of an Analytical Hierarchy Process System for Evaluating the Performance of Self-sustained Houses", *Eng. J.*, Vol. 2, No. 2, 2023.

[49] M. A. Deif, M. A. A. Eldosoky, H. W. Gomma, A. M. E. Garhy, and A. S. E. Azab, "Adaptive Neuro-Fuzzy Inference System Controller Technique for Lower Urinary Tract System Disorders", *J. Clin. Eng.*, Vol. 40, No. 3, pp. 135–143, 2015.

[50] R. E. Hammam, A. A. A. Solyman, M. H. Alsharif, P. Uthansakul, and M. A. Deif, "Design of Biodegradable Mg Alloy for Abdominal Aortic Aneurysm Repair (AAAR) Using ANFIS Regression Model", *IEEE Access*, Vol. 10, pp. 28579–28589, 2022.

[51] M. A. Deif and R. E. Hammam, "Skin lesions classification based on deep learning approach", *J. Clin. Eng.*, Vol. 45, No. 3, pp. 155–161, 2020.

[52] M. A. Deif, R. E. Hammam, S. Ahmed, A. M. Kamarposhti, S. B. Shahab, and E. H. Rania, "A deep bidirectional recurrent neural network for identification of SARS-CoV-2 from viral genome sequences", *Math. Biosci. Eng.*, Vol. 18, No. 6, p. AIMS-Press, 2021.

[53] A. Alqerem, H. Attar, W. Alomoush, and M. Deif, "The Ability of Ultra Wideband to Differentiate Between Hematoma and Tumor Occur in The Brain", In: *Proc. of 2022 International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)*, 2022, pp. 1–7.

[54] N. Baghdadi, A. S. Maklad, A. Malki, and M. A. Deif, "Reliable Sarcoidosis Detection Using Chest X-rays with EfficientNets and Stain-Normalization Techniques", *Sensors*, Vol. 22, No. 10, p. 3846, 2022.

[55]M. Alshehri, M. Kumar, A. Bhardwaj, S. Mishra, and J. Gyani, "Deep learning based approach to classify saline particles in sea water", *Water*, Vol. 13, No. 9, p. 1251, 2021.

[56] S. Mahajan, L. Abualigah, A. K. Pandit, and M. Altalhi, "Hybrid Aquila optimizer with arithmetic optimization algorithm for global optimization tasks", *Soft Comput.*, Vol. 26, No. 10, pp. 4863–4881, 2022.

[57] M. A. Deif, M. A. A. Eldosoky, A. M. E. Garhy, H. W. Gomma, and A. S. E. Azab, "Parasympathetic Nervous Signal Damping Using the Adaptive Neuro-Fuzzy Inference System Method to Control Overactive Bladder", *J. Clin. Eng.*, Vol. 40, No. 4, pp. 197–201, 2015.

[58] K. Natarajan, P. K. Bala, and V. Sampath, "Fault detection of solar PV system using SVM and thermal image processing", *Int. J. Renew. Energy Res.*, Vol. 10, No. 2, pp. 967–977, 2020.

[59] O. N. Oyelade, A. E. S. Ezugwu, T. I. A. Mohamed, and L. Abualigah, "Ebola optimization search algorithm: A new nature-inspired metaheuristic optimization algorithm", *IEEE Access*, Vol. 10, pp. 16150–16177, 2022.