Gazi Üniversitesi
Fen Bilimleri Dergisi
PART C: TASARIM VE TEKNOLOJİ

Gazi University
Journal of Science
PART C: DESIGN AND TECHNOLOGY

GU J Sci, Part C, 12(4): 819-831 (2024)

Estimation of Hydroelectric Power Generation and Analysis of Climate Factors with Deep Learning Methods: A Case Study in Yozgat Province in Turkey

Feyza Nur ÇAKICI[1*] Süleyman Sungur TEZCAN[1] Hidir DUZKAYA[1]

[1]*Gazi University, Faculty of Engineering, Department of Electrical-Electronic Engineering, Ankara, Turkey*

Article Info

Research article
Received: 17/07/2024
Revision: 13/09/2024
Accepted: 21/09/2024

Keywords

Hydroelectric Generation
Climate Data
Deep Learning Methods
Power Forecasting

Makale Bilgisi

Araştırma makalesi
Başvuru: 17/07/2024
Düzeltme: 13/09/2024
Kabul: 21/09/2024

Anahtar Kelimeler

Hidroelektrik Üretim
İklim Verisi
Derin Öğrenme Metotları
Güç Tahmini

Graphical/Tabular Abstract (Grafik Özet)

In this study, hydroelectric power forecasting is conducted using Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and hybrid LSTM-SVR models based on climate data. / Bu çalışmada, iklim verilerine dayalı olarak Uzun Kısa Süreli Hafıza (LSTM), Destek Vektör Regresyonu (SVR) ve hibrit LSTM-SVR modelleri kullanılarak hidroelektrik güç tahmini gerçekleştirilmiştir.

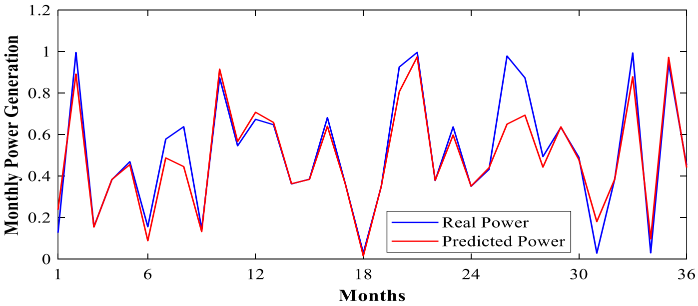

***Figure A**: Test graph of the SVR model with 12 parameters / **Şekil A:** 12 parametreli SVR modelinin test grafiği*

Highlights (Önemli noktalar)

- *Twelve climate variables obtained from the Yozgat Meteorological Directorate are being used. In this way, the impact of climate data on hydroelectric power generation is being examined./ Yozgat Meteoroloji Müdürlüğü'nden elde edilen 12 iklim verisi kullanılmaktadır. Böylece iklim verilerinin hidroelektrik güç üretimi üzerindeki etkisi incelenmektedir.*
- *The effects of climate data on hydroelectric power generation are analyzed using Pearson correlation analysis./ Pearson Korelasyon analizi ile iklim verilerinin hidroelektrik güç üretimine etkileri analiz edilmektedir.*
- *Hyperparameters are selected using the Bayesian Optimization method, and the performance of the models is enhanced./ Bayes Optimizasyon yöntemi kullanılarak hiperparametreler seçilmekte ve modellerin performansı artırılmaktadır.*

Aim (Amaç): *The aim of this study is to investigate the effect of climate data on hydroelectric power generation. / Bu çalışmanın amacı, iklim verilerinin hidroelektrik enerji üretimi üzerindeki etkisini incelemektir.*

Originality (Özgünlük): *Power prediction has been conducted using LSTM, SVR, and the LSTM-SVR models, utilizing data obtained from the Süreyyabey Hydroelectric Power Plant and the Yozgat Meteorological Directorate. / Süreyyabey hidroelektrik güç santrali ve Yozgat Meteoroloji Müdürlüğü'nden elde edilen veriler kullanılarak LSTM, SVR ve LSTM-SVR modelleri ile güç tahmini gerçekleştirilmiştir.*

Results (Bulgular): *As a result of power prediction using the SVR model with 11 and 12 climate parameters, the R-value is close to 1, while the MAE and RMSE values are observed to be close to 0./ SVR modelinin 11 ve 12 iklim parametresi ile güç tahmini sonucunda R değeri 1'e yakın olmakta, MAE ve RMSE değerleri ise 0'a yakın değerler almaktadır.*

Conclusion (Sonuç): *In this study, the SVR model has achieved the best performance in power prediction, and it has been concluded that climate data has a significant impact on hydroelectric power generation./ Bu çalışmada, SVR modeli güç tahmininde en iyi performansı elde etmiş ve iklim verilerinin hidroelektrik güç üretiminde önemli bir etkisi olduğu sonucuna varılmıştır.*

*Corresponding author, e-mail: fnur.cakici@gazi.edu.tr
DOI: 10.29109/gujsc.1517800

GU J Sci, Part C, 12(4): 819-831 (2024)

Gazi Üniversitesi
Fen Bilimleri Dergisi
PART C: TASARIM VE TEKNOLOJİ

Gazi University
Journal of Science
PART C: DESIGN AND TECHNOLOGY

http://dergipark.gov.tr/gujsc

Estimation of Hydroelectric Power Generation and Analysis of Climate Factors with Deep Learning Methods: A Case Study in Yozgat Province in Turkey

Feyza Nur ÇAKICI[1*] Süleyman Sungur TEZCAN[1] Hidir DUZKAYA[1]

[1]*Gazi University, Vocational School of Technical Sciences, Department of Design, Ankara, Turkey*

Article Info

Research article
Received: 17/07/2024
Revision: 13/09/2024
Accepted: 21/09/2024

Keywords

Hydroelectric Generation
Climate Data
Deep Learning Methods
Power Forecasting

Abstract

Hydroelectric power is a significant renewable energy source for the development of countries. However, climatic data can impact power generation in hydroelectric power plants. Hydroelectric power forecasting is conducted in this study using Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and hybrid LSTM-SVR models based on climatic data. The dataset consists of climate data from the Yozgat Meteorology Directorate in Turkey from 2007 to 2021 and power data obtained from the Süreyyabey Hydroelectric Power Plant in Yozgat. The correlation coefficient examines the relationship between climate data and monthly hydroelectric power generation. The hyper-parameters of the models are adjusted using the Bayesian Optimization (BO) method. The performance of monthly hydroelectric power prediction models is assessed using metrics such as correlation coefficient (R), root mean square error (RMSE), and mean absolute error (MAE). When trained using 11 and 12 climate parameters, the SVR model exhibits an R-value close to 1, and MAE and RMSE values close to 0 are observed. Additionally, regarding training time, the SVR model achieves accurate predictions with the shortest duration and the least error compared to other models.

Yozgat İli'nde Hidroelektrik Enerji Üretiminin Tahmini ve İklim Faktörlerinin Derin Öğrenme Yöntemleri ile Analizi: Bir Vaka Çalışması

Makale Bilgisi

Araştırma makalesi
Başvuru: 17/07/2024
Düzeltme: 13/09/2024
Kabul: 21/09/2024

Anahtar Kelimeler

Hidroelektrik Üretim
İklim Verisi
Derin Öğrenme Metotlar
Güç Tahmini

Öz

Hidroelektrik enerji, ülkelerin kalkınmasında önemli bir yenilenebilir enerji kaynağıdır. Ancak iklim verileri, hidroelektrik santrallerindeki enerji üretimini etkileyebilmektedir. Bu çalışmada, hidroelektrik enerji üretimi tahmini, iklim verilerine dayalı olarak Uzun Kısa Süreli Hafıza (LSTM), Destek Vektör Regresyonu (SVR) ve hibrit LSTM-SVR modelleri kullanılarak gerçekleştirilmiştir. Veri seti, Türkiye'deki Yozgat Meteoroloji Müdürlüğü'nden 2007-2021 yılları arasında alınan iklim verileri ve Yozgat'taki Süreyyabey Hidroelektrik Santrali'nden elde edilen enerji verilerinden oluşmaktadır. İklim verileri ile aylık hidroelektrik enerji üretimi arasındaki ilişki, korelasyon katsayısı ile incelenmiştir. Modellerin hiper-parametreleri, Bayesian Optimizasyonu (BO) yöntemi kullanılarak ayarlanmıştır. Aylık hidroelektrik enerji tahmin modellerinin performansı; korelasyon katsayısı (R), ortalama kare hatası (RMSE) ve ortalama mutlak hata (MAE) gibi metriklerle değerlendirilmiştir. 11 ve 12 iklim parametresi ile eğitilen SVR modelinde R değeri 1'e yakın olup, MAE ve RMSE değerlerinin 0'a yakın olduğu gözlemlenmiştir. Ayrıca eğitim süresi açısından SVR modeli, diğer modellere kıyasla en kısa sürede en az hata ile doğru tahminler yapmıştır.

1. INTRODUCTION (GİRİŞ)

Advancements in technology, the proliferation of industrial activities, and population growth increase electricity demand, escalating the electricity generated to meet this demand. The International Renewable Energy Agency (IRENA) predicts that global electricity demand will increase by an average of 3.4% annually until 2026 [1]. Sustainable energy sources are becoming increasingly popular to meet this growing demand and reduce greenhouse gas emissions [2]. As a result, global hydroelectric power generation in 2020 has reached 4345.99 TWh [3]. According to the IRENA report, 2022 global hydroelectric installed capacity has reached 1256 GW, excluding pumped hydro. This represents 37% of total renewable energy sources [4].

*Corresponding author, e-mail: fnur.cakici @gazi.edu.tr
DOI: 10.29109/gujsc.1517800

Hydroelectric power plants are affected by climate change because they use natural water resources for energy production [5-7]. Due to the negative effects of climate change, hydroelectric power capacity decreased by more than 2% globally, and hydroelectric power generation decreased by 4.5% in Turkey in 2023 [1]. Floods, droughts, rainfall acceleration or increase, changes in rainfall timing, and temperature changes are a few of these effects [2]. As a result of a decrease in precipitation, the turbine efficiency is affected by a reduction of flow [8]. In contrast, excessive precipitation may damage turbine blades due to increased sediment load, even though increasing precipitation increases power generation [6]. Increased temperatures lead to increased evaporation from reservoir surfaces, resulting in a decrease in water levels.

The Representative Concentration Pathway (R.C.P.) climate scenarios are used to examine the impact of climate change on hydroelectric power generation. RCP 2.6, RCP 4.5, and RCP 8.5 scenarios illustrate future climate conditions. Data derived from these scenarios are analyzed using machine learning, deep learning, and statistical methods to estimate hydroelectric power generation. In the RCP 4.5 and RCP 8.5 scenarios, carbon emissions are projected to reach 4.5 W/m^2 and 8.5 W/m^2 by 2100, respectively. In contrast, the RCP 2.6 scenario anticipates a reduction in carbon emissions. Based on these carbon emission projections, the RCP 2.6, RCP 4.5, and RCP 8.5 scenarios are considered best-case, moderate, and worst-case scenarios, respectively [9-11].

The studies evaluating the effects of climate change on hydroelectric energy production from dams in China, conducted by Huang et al. (2021), state that the hydroelectric energy production of the Liyuan Dam decreased by 163.3 MW and 188.3 MW under the RCP2.6 and RCP8.5 climate scenarios, respectively, relative to the base period. It is also projected that the increase in electricity demand by 91.42 MW could lead to power outages [9]. Meanwhile, the study by Huangpeng et al. (2021) predicts that the electricity production of the Jinanqiao Hydroelectric Power Plant may decrease by 10.74%, 16.38%, and 22.25% by 2050 under the RCP2.6, RCP4.5, and RCP8.5 climate scenarios, respectively [12].

According to findings from various studies in Africa, Boadi et al. (2019) indicate that the power generation of the Aksombo Hydroelectric Power Plant in Ghana has been influenced by rainfall fluctuations, accounting for 21% between 1970 and 1990, while 72.4% has been influenced by ENSO (El Niño Southern Oscillation) and lake levels. These results suggest Ghana must explore alternative energy sources for electricity generation [13]. Hamududu et al. (2016) project that power generation in the Zambezi River Basin will decrease by 8%, 18%, and 28% by 2020, 2050, and 2080, respectively, due to rising temperatures and reduced rainfall [14]. Uamusse et al. (2020) suggest that hydroelectric power plants in Mozambique will experience reduced power generation due to drought, negatively impacting the economy. The study predicts a 20% decrease in power generation for the Cahora Bassa hydroelectric power plants by 2100 [15].

From the studies in Iran, Behesti et al. (2019) stated that the hydroelectric power production of the Karun III Dam is affected by the flow. Accordingly, hydroelectric power production is estimated to increase by 26.7% to 40.5% between 2020 and 2049 and 17.4% to 29.3% between 2070 and 2099 [5]. In the study by Wang et al. (2021), it is indicated that the decrease in flow of the Aras Dam on the border of Iran and the Republic of Azerbaijan would result in a reduction in hydroelectric power generation. Additionally, according to the RCP2.6, RCP4.5, and RCP8.5 climate scenarios, this dam's annual average hydropower generation is projected to decrease by 3.36 MW, 4.62 MW, and 6.64 MW, respectively [11].

One of the studies conducted by different countries, Khaniya et al. (2020) in its research, Samanalawawe Hydroelectric Power Station in Sri Lanka is estimated to increase hydroelectric power generation by 7.29% and 10.22% from 2020 to 2050, respectively, according to RCP4.5 and RCP8.5 climate scenarios. This study reveals that Samanalawawe Dam is not affected by climate change [16]. In another study, Shrestha et al. (2021) predict that the hydroelectric power generation of the Kulekhani Hydroelectric Power Plant in Nepal will decrease by between 0.5% and 13%. The decrease is attributed to the temperature increase and rainfall fluctuations [17].

One of the studies on hydroelectric power prediction in Turkey is the work by Karakuş (2023), who employed a new deep hybrid model to forecast power generation and Net Head for the Hirfanlı Hydroelectric Power Plant in Kırşehir. This study utilized climate factors and data obtained from hydroelectric plants. The successful performance of the hybrid model in predicting hydroelectric power generation and Net Head has been confirmed through statistical analyses. Additionally, the article emphasizes that this model will be used to adjust energy consumption [18]. In Ceribaşı's study, energy predictions for the Adasu Regulator and Hydroelectric Power Plant and Pamukova Hydroelectric Power Plant in the Sakarya Basin, Turkey, are made both in the short and long term. The study emphasizes that these predictions demonstrated strong performance and underscores the importance of short and long-term forecasts in energy planning [19].

This study aims to investigate the effect of climate data on hydroelectric power generation. The study utilized monthly climate data from the Yozgat Meteorology Directorate in Turkey from 1 January 2007 to 31 December 2021. Following the preparation phase of the climate data, the Bayesian Optimization technique is employed to enhance prediction performance. For hydroelectric power forecasting, Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and hybrid LSTM-SVR models are applied, and the accuracy of these models is evaluated using statistical measurement parameters.

2.MATERIALS AND METHODS (MATERYAL VE METOD)

This section provides information about Süreyyabey hydroelectric power plants and climate datasets. The min-max normalization technique is employed for data preprocessing. The relationship between climate data is being examined through Karl Pearson correlation coefficient. Bayesian optimization is used to tune the hyper-parameters of the RNN and SVR models to improve the prediction accuracy. RNN, SVR, and hybrid RNN-SVR models are used in hydroelectric power prediction.

2.1.Study Area And Data Set (Çalışma Alanı ve Veri Seti)

Süreyyabey Dam, located on the Çekerek Stream in the Yozgat province, has a reservoir area of 41.34 km² [20], with a riverbed elevation of 103 meters [21]. The annual energy production of the Süreyyabey Dam is projected to be 51 GWh [21]. The data set consists of climate data from the Yozgat Meteorology Directorate and power generation data from the Süreyyabey hydroelectric power plant for 15 years. The parameters of the data set are indicated in Table 1.

Table 1. Daily climate and power generation data for the Süreyyabey hydroelectric power plant
(Süreyyabey hidroelektrik santralleri için günlük iklim ve enerji üretim verileri)

Feature	Units	Range
Precipitation Amount	Mm	(0)-(93.139)
Open Surface Evaporation Amount	mm	(0)-(12.2)
Sunshine Duration	h	(-1.6)-(20.25)
Average Air Temperature	℃	(-12.1)-(40.762)
Maximum Air Temperature	℃	(-9.2)-(37.4)
Minimum Air Temperature	℃	(-17.0)-(22.2)
Average Wind Direction	°	(0.3)-(359.8)
Average Wind Speed	m/s	(0.7)-(12.2)
The Direction of Maximum Wind	°	(1.0)-(360)
Maximum Wind Speed	m/s	(1.9)-(25.3)
Time of Maximum Wind	h	(-1.38)-(23.58)
Average Humidity	%	(0)-(100)
Power Generation	kWh	(89.0)-(331.02)

Figure 1 illustrates the total hydroelectric power generation data between 2007 and 2021. In the graph, hydroelectric power generation reached its lowest level at 13.40 GWh in 2009 and its highest level at 44.10 GWh in 2012. There is a downward trend in hydroelectric power generation from 2012 to 2015. However, power generation shows an increasing trend in 2016, 2017, 2019, and 2021 but has yet to reach the highest level seen in 2012.

2.2.Min-Max Normalization (Min-Mak Normalizasyon)

Data sets in machine learning, deep learning, regression, and classification problems often have different ranges [22]. These differences prevent the model from achieving better results. Therefore, the min-max normalization technique transforms the data set into a standard format [23]. This method normalized the data values to the [0,1] range, eliminating scale differences [22]. Equation (1) provides the mathematical formula for the min-max normalization technique.

$$X' = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{1}$$

The normalization of climate data converts different scales of data into standardized data.

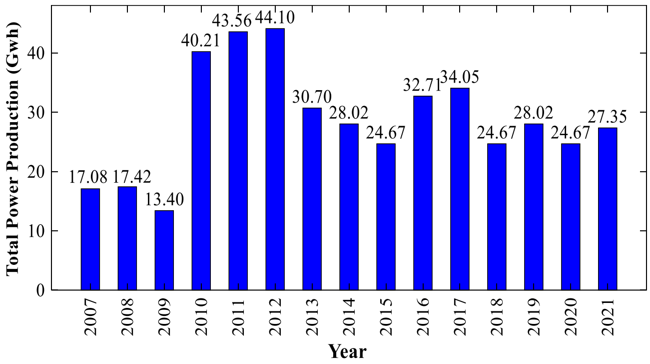

Figure 1. Total power generation values (GWh) for Süreyyabey Hydroelectric Power Plant (Süreyyabey Hidroelektrik Santrali için toplam güç üretim değerleri (GWh))

2.3.Correlation Coefficient (Korelayon Katsayısı)

Correlation is a data analysis technique used to describe the strength of the relationship between any two variables in a dataset [24]. Additionally, it examines whether the relationship between these variables is linear. To determine the relationship between two variables, two methods are commonly used: the Karl Pearson correlation coefficient and the Spearman rank correlation coefficient. The Pearson correlation coefficient examines whether there is a linear relationship between two independent variables. The correlation coefficient is a scientific measure that takes values ranging from +1 to -1. There are specific criteria for interpreting the correlation coefficient. As a result of these criteria, "0" indicates that there is no linear relationship between the variables, "+1" indicates a powerful positive relationship, and "-1" means a very strong negative relationship [25]. Equation (2) shows the equality of the correlation coefficient [26].

$$r_{y\hat{y}} = \frac{\sum(\hat{y}_i - \bar{\hat{y}}_i)(y_i - \bar{y})}{[\sum(y_i - \bar{\hat{y}})^2]^{1/2}[\sum(y_i - \bar{y})^2]^{1/2}} \tag{2}$$

where $\hat{y}_i$ denotes the predicted value of y, and represents the mean of the y values. Equation (3) is used to calculate the mean value [27].

$$\bar{U} = \frac{1}{n}\sum_{i=1}^{n} U_i \tag{3}$$

2.4.Bayesian Optimization (BO) (Bayes Optimizasyonu)

Hyper-parameter tuning is a crucial aspect of deep learning and machine learning algorithms. In deep learning, hyper-parameters such as learning rate, number of iterations, number of hidden layers, batch size, and activation functions are used [28]. Hyper-parameters affect the model's performance, thereby increasing the accuracy of the predicted data. Hyper-parameters are used in the training process and the development of the model [29]. The learning rate is used to adjust the speed at which the weights in the model are updated [30].

In the literature, manual tuning, grid search, random search, and Bayesian Optimization (BO) methods are preferred for hyper-parameter tuning [31]. Manual hyper-parameter tuning is more challenging and requires expertise compared to other methods. As the size of the dataset increases, manual tuning can significantly impact the performance and accuracy of the models during training. Therefore, using automated search methods for hyper-parameter tuning is more efficient [32]. Grid search is an easy-to-implement method; however, its efficiency decreases as hyper-parameters increase [31]. Random search has advantages such as being more efficient by not searching for unnecessary hyper-parameters for the model, having lower costs than grid search, and being efficient in high-dimensional datasets [32]. However, its disadvantage is that it conducts unnecessary analyses as the search space increases since it only utilizes previously found good performances [31]. Unlike other optimization methods, Bayesian Optimization (BO) utilizes prior knowledge of the objective function to minimize loss effectively, leading to improved prediction accuracy [28]. BO utilizes a probabilistic model for solving complex problems [33]. In BO, Gaussian process models and acquisition functions tune hyper-parameters.

In LSTM, SVR, and LSTM-SVR models, hyper-parameters such as batch size, activation function, epoch, optimization algorithm, hidden layer structure, learning rate, kernel function, box constraint, epsilon, and kernel scale are tuned using the Bayesian Optimization (BO) method.

2.5.Long Short-Term Memory (Uzun Kısa Süreli Bellek)

Long Short-Term Memory (LSTM) addresses the vanishing gradient problem encountered in RNN models [34]. LSTM consists of four fundamental components: the forget gate, input gate, output gate, and cell state [35]. With its three gates, input, output, and forget, LSTM has a longer data retention time than RNN. The cell state is used to store and update information in the gates. The forgetting gate processes data that will not be considered, the input gate determines the data to be added, and the output gate identifies the data to be outputted [36]. LSTM utilizes sigmoid and hyperbolic tangent functions. Sigmoid functions determine which data to forget and which to remember, whereas hyperbolic tangent functions regulate cell data [37]. Figure 2 illustrates the structure of the LSTM model.

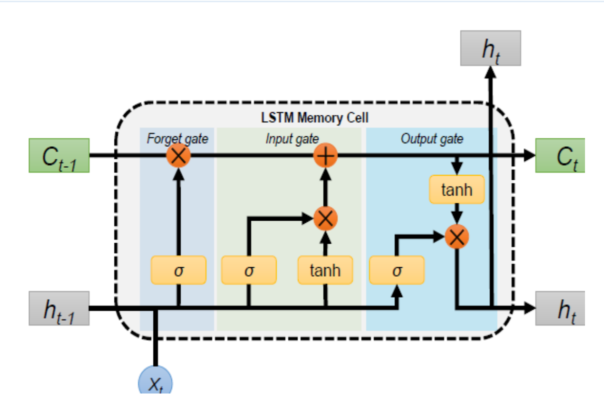

Figure 2. The structure of the LSTM model [38]
(LSTM modelin yapısı)

2.6.Support Vector Regression (SVR) (Destek Vektör Regresyon)

Support Vector Regression (SVR) is modeled as a supervised machine learning algorithm capable of predicting nonlinear data. Its computation is based on the linear regression model [39]. SVR aims to minimize the error between predicted and actual values. Its advantages are achieving good results even with small data sizes and overcoming complex problems [40]. Furthermore, parameter selection plays an important role in the performance of SVR [31]. Several disadvantages are associated with this method, including increased computation time with growing datasets [30] and difficulties in selecting the appropriate kernel function for optimal solutions [40]. To improve the accuracy of the SVR model, it is crucial to adjust the type of kernel function and its parameters. Improper adjustment of these values can lead to issues such as over-fitting or under-fitting [41].

2.7.Performance Metrics (Performans Metrikler)

To evaluate the methods used and the predicted hydroelectric power using monthly climate data, as well as to compare the performance results, the measurement parameters of correlation coefficient (R), root mean square error (RMSE), and mean absolute error (MAE) are utilized.

The mean absolute error (MAE) is used to calculate the difference between the predicted power data and the actual power data [33]. The MAE ranges from 0 to $+\infty$, with values closer to 0 indicating that the predicted values are closer to the actual data [42]. Equations (4) and (5), respectively, depict the mathematical formulas of MAE and RMSE [31]. RMSE also ranges between 0 and $+\infty$, similar to MAE. As RMSE approaches MAE, model error robustness improves [42].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y-\hat{y}\right| \qquad (4)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y-\hat{y})^2} \qquad (5)$$

where y represents the actual data, the predicted data, and the total number of data points in the dataset.

3.RESULTS (Sonuçlar)

This section discusses the comparison results of the LSTM, SVR, and hybrid LSTM-SVR models used for power predictions of the Süreyyabey Hydroelectric Power Plant. Climate data are obtained from the Yozgat Weather Directorate, and power production data from the Süreyyabey Hydroelectric Power Plant are utilized. The study initially is conducted as a correlation analysis of twelve climate variables. Table 2 presents the correlation analysis coefficients of the monthly climate data.

The correlation coefficients of climate data with power production from strong to weak in Table 2. When examining the correlation coefficients of monthly climate data in Table 2, it is observed that there is a positive relationship between precipitation amount, average wind direction, the direction of maximum wind, average wind speed, maximum wind speed, sunshine duration, and power production. Conversely, climate variables such as average humidity, maximum air temperature, minimum air temperature, time of maximum wind, open surface evaporation amount, and average air temperature negatively correlate with power production.

Table 2. The correlation analysis results of the monthly climate data (Aylık iklim verilerinin korelasyon analizi sonuçları)

Parameter Values	Correlation Analysis
Precipitation Amount	0.3717
Average Wind Direction	0.2108
The Direction of Maximum Wind	0.1774
Average Wind Speed	0.1570
Maximum Wind Speed	0.1249
Sunshine Duration	0.0742
Average Humidity	-0.1956
Maximum Air Temperature	-0.1846
Minimum Air Temperature	-0.0855
Time of Maximum Wind	-0.0743
Open Surface Evaporation Amount	-0.0514
Average Air Temperature	-0.0138

The correlation analysis results are used as inputs for LSTM, SVR, and hybrid LSTM-SVR models for predicting power generation. In the hybrid LSTM-SVR model, the outputs of LSTM are used as input data for the SVR model.

Table 3. Hyper-parameters and their values for LSTM, SVR and LSTM-SVR (LSTM, SVR ve LSTM-SVR için Hiperparametreler ve Değerleri)

Models	Hyper-parameters	Values
LSTM	Batch size	19
	Activation Function	ReLU
	Epoch	149
	Optimization Algorithm	Adam
	Hidden Layer Structure	6
	Learning Rate	0.047997
SVR	Kernel Function	Gaussian
	Box Constraint	5.6355
	Epsilon	0.0028193
	Kernel Scale	1.3571
LSTM-SVR	Kernel Function	Gaussian
	Box Constraint	960.81
	Epsilon	0.0010132
	Kernel Scale	9.9369
	Batch size	29
	Activation Function	ReLU
	Epoch	277
	Optimization Algorithm	Adam
	Hidden Layer Structure	2
	Learning Rate	0.024174

The study randomly partitioned the dataset into 80% training and 20% test data in MATLAB. Table 3 shows the hyper-parameters and their values for the LSTM, SVR, and LSTM-SVR models.

Table 4 compares R, RMSE, and MAE values obtained after 10 iterations for LSTM, SVR, and LSTM-SVR models across 12 parameters in monthly data. Table 5 compares 11 parameters, excluding average air temperature, for LSTM, SVR, and hybrid LSTM-SVR models. The R-value of 0.96183 obtained for the SVR model with 12 parameters and 0.96463 obtained for 11 parameters indicates a strong relationship between climatic data and hydroelectric power generation. Additionally, the lower RMSE value of the SVR model compared to other models suggests that the model makes predictions with small errors. Furthermore, the MAE value of the SVR model is close to 0 when compared to actual hydroelectric power generation values, indicating that the model operates with minimal errors. The SVR model completes its training process in the shortest time when comparing the training times of the models in the study.

Table 4. The performance metrics of LSTM, SVR, and LSTM-SVR models for 12 parameters (LSTM, SVR ve LSTM-SVR modellerinin 12 parametre için performans metrikleri)

Model	R	RMSE	MAE	Training Time
LSTM	0.8625	0.14132	0.10951	6.0711
SVR	0.96183	0.08811	0.052819	0.15625
LSTM-SVR	0.83162	0.14576	0.12129	LSTM: 8.4592 - SVR:4.1605

Table 5. The performance metrics of LSTM, SVR, and LSTM-SVR models for 11 parameters (LSTM, SVR ve LSTM-SVR modellerinin 11 parametre için performans metrikleri)

Model	R	RMSE	MAE	Training Time
LSTM	0.75665	0.14964	0.12243	6.5335
SVR	0.96463	0.083663	0.051532	0.15625
LSTM-SVR	0.8400	0.16902	0.14056	LSTM: 10.3138 - SVR:4.0729

Figures 3, 4, and 5 show the real power and power graphs predicted using the SVR, LSTM, and LSTM-SVR models for 12 parameters. The graphs depict 36 test data points, representing 20% of the 180 monthly power data points. The blue lines indicate the actual power values of the test data, while the red lines represent the power values predicted by the models. Figure 3 demonstrates that the SVR model, utilizing 12 climate parameters, predicts the actual power with less error. Although the LSTM model has an R-value of 0.8625 for the 12 climate parameters, its MAE value is lower than the LSTM-SVR model, indicating that it predicts the actual power with less error. This observation is clearly illustrated in Figure 4, showcasing the LSTM model's performance.

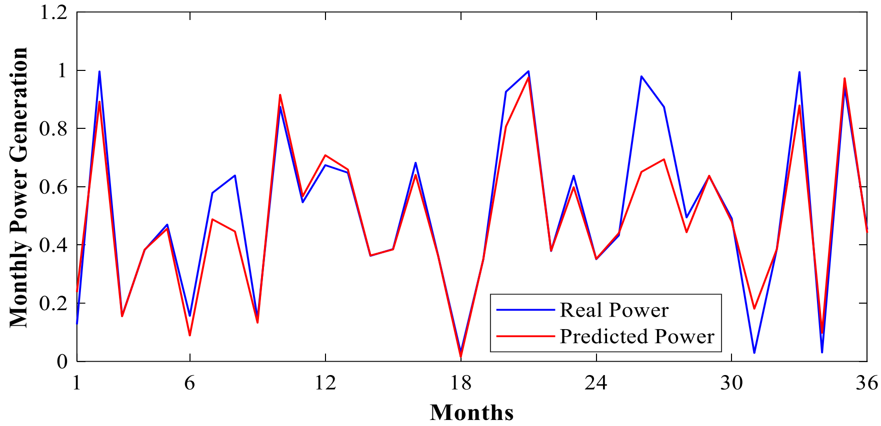

Figure 3. Test graph of the SVR model with 12 parameters (12 parametreli SVR modelinin test grafiği)

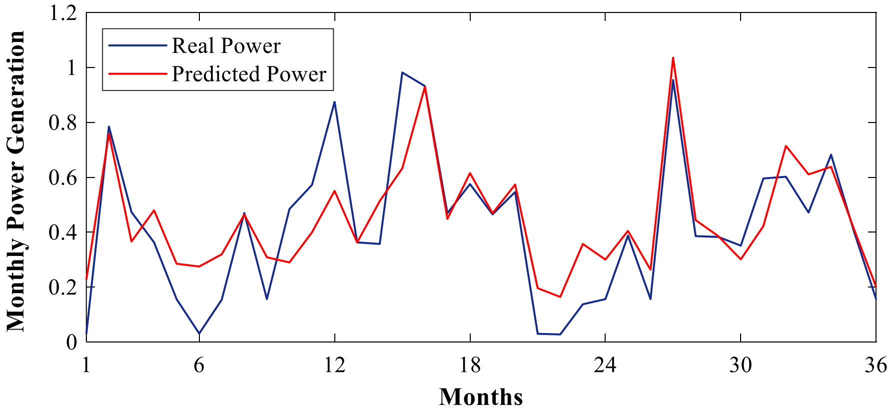

Figure 4. Test graph of the LSTM model with 12 parameters (12 parametreli LSTM modelinin test grafiği)

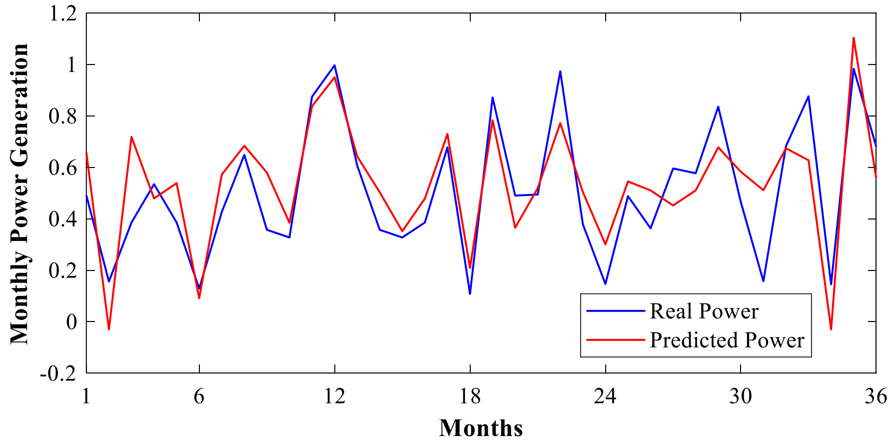

Figure 5. Test graph of the LSTM-SVR model with 12 parameters (12 parametreli LSTM-SVR modelinin test grafiği)

Figures 6, 7, and 8 show the real power and power graphs predicted using the SVR, LSTM, and LSTM-SVR models for 11 parameters. Figure 6 illustrates the relationship between actual and predicted power data when the average air temperature parameter, among the total of 12 climate parameters, is excluded from the SVR model. Figures 7 and 8, however, observe a more pronounced discrepancy between actual and predicted power values.

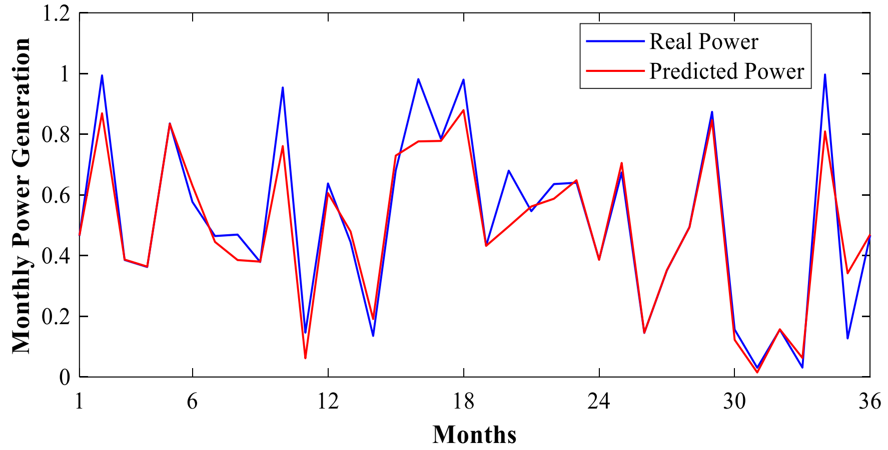

Figure 6. Test graph of the SVR model with 11 parameters (11 parametreli SVR modelinin test grafiği)

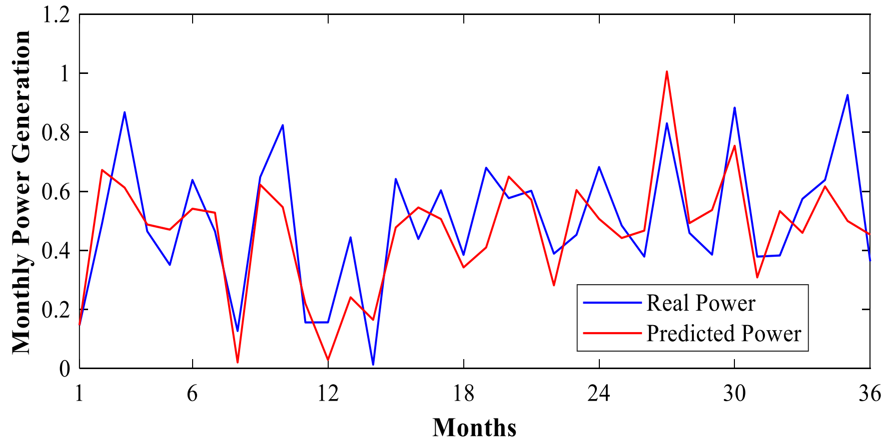

Figure 7. Test graph of the LSTM model with 11 parameters (11 parametreli LSTM modelinin test grafiği)

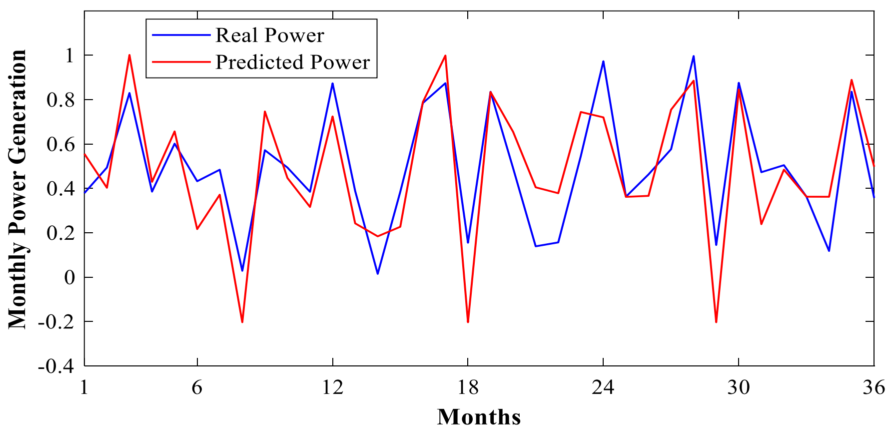

Figure 8. Test graph of the LSTM-SVR model with 11 parameters (11 parametreli LSTM-SVR modelinin test grafiği)

4.DISCUSSION (TARTIŞMA)

In the study, the impact of climate data on power generation is examined by calculating the Pearson correlation coefficient. According to the correlation analysis results, the highest correlation within monthly climate data is found between precipitation amount and hydroelectric power generation, with a correlation coefficient of 0.3717. The correlation coefficient for average air temperature is calculated as -0.0138, indicating no significant relationship between average air temperature and hydroelectric power generation. The correlation analysis reveals that the average air temperature has a value close to 0. This value indicates that the effect of average air temperature on hydroelectric power generation is weak.

The close R values of the 12 and 11 parameters presented in Tables 4 and 5 of the SVR model indicate that similar parameters are used in both models. However, it is determined that the exclusion of average air temperature has a minimal effect on the overall performance of the model. According to the results of the correlation analysis, average air temperature is found to have a weak effect on hydroelectric power generation. Therefore, it is concluded that average air temperature has a minimal impact on the performance of the SVR model.

In other studies, various models have been developed to predict hydroelectric power generation for hydroelectric plants located in different countries. These models examine the impact of different climatic parameters on hydroelectric power generation. In the study by Ekanayake et al. (2021), rainfall data, minimum and maximum temperature, and evaporation parameters collected from six locations in the catchment area of the Samanalawewa reservoir in Sri Lanka have been used. Pearson and Spearman correlation analysis has revealed that rainfall data from only one location has a minimal impact on energy production. Gaussian Process Regression (GPR), Support Vector Regression (SVR), Multiple Linear Regression (MLR), and Power Regression (PR) models have been used to predict hydroelectric power generation. The performance of these models has been evaluated using different statistical methods, including the correlation coefficient (R), root mean square error (RMSE), mean absolute percentage error (MAPE), the ratio of RMSE to the standard deviation of measured data (RSR), BIAS, and the Nash number. In the study, the R values for the SVR, GPR, MLR, and PR models are 0.87, 0.92, 0.60, and 0.67, respectively. It has been emphasized that the GPR model provided the most accurate predictions, with its R-value being the closest to 1 among the models. Compared to our study, the prediction accuracy of the SVR model increases with the number of parameters [39].

Javed et al. (2020) conducted a study on hydroelectric power prediction using temperature and rainfall data from the Tarbela Dam in Pakistan. The study employed MLR, K-Nearest Neighbour (K-NN), SVR, Random Forest (RF), and LSTM models. The RF model achieved the lowest prediction error when using temperature data, with an MAE of 2.47 and an RMSE of 3.98. While the accuracy of the LSTM model improves with an increased dataset size, this study did not perform as well, yielding an MAE of 4.39 and an RMSE of 6.89. According to the paper, the SVR model exhibited the highest error rates, with an MAE of 9.24 and an RMSE of 10.75 [43].

5. CONCLUSIONS (SONUÇLAR)

In this study, the monthly power prediction of the Süreyyabey Hydroelectric Power Plant has been conducted using LSTM, SVR, and a hybrid LSTM-SVR model. The statistical measurement parameters R, RMSE, and MAE have been employed to determine the best model for power prediction. A dataset consisting of 12 climatic variables obtained from the Yozgat Meteorological Directorate has been utilized, from 2007 to 2021 monthly. In this dataset, climate and power data have been normalized and ranked from strong to weak relationships through correlation analysis. The selection of hyper-parameters in the models has been performed using the BO method. The BO method is used to quickly find the optimal results for hyper-parameters, thereby preventing models from overfitting or under-fitting.

The SVR model achieves R values of 0.96183 and 0.96463 on the 12 and 11 climate datasets, respectively, which are closest to 1 compared to other models. Additionally, the SVR model's RMSE values are 0.08811 and 0.083663, while the MAE values are 0.052819 and 0.051532. These RMSE and MAE values are close to 0 compared to other models, indicating that the SVR model exhibits more accurate prediction performance. According to the performance results in Tables 4 and 5, the SVR model is observed to predict hydroelectric power generation most accurately.

This study emphasizes the importance of climate data in hydroelectric power prediction. In future studies, climate scenarios can be utilized for hydroelectric power plants or various energy types. Hydroelectric power can be predicted using hybrid deep learning and machine learning models.

DECLARATION OF ETHICAL STANDARDS
(ETİK STANDARTLARIN BEYANI)

The author of this article declares that the materials and methods they use in their work do not require ethical committee approval and/or legal-specific permission.

Bu makalenin yazarı çalışmalarında kullandıkları materyal ve yöntemlerin etik kurul izni ve/veya yasal-özel bir izin gerektirmediğini beyan ederler.

AUTHORS' CONTRIBUTIONS (YAZARLARIN KATKILARI)

Feyza Nur ÇAKICI: He performed the simulation studies, analyzed the results, and carried out the writting process.

Benzetim çalışmalarını gerçekleştirdi, sonuçlarını analiz etti ve yazım işlemini gerçekleştirmiştir.

Süleyman Sungur TEZCAN: He performed the simulation studies, analyzed the results, and carried out the writting process.

Benzetim çalışmalarını gerçekleştirdi, sonuçlarını analiz etti ve yazım işlemini gerçekleştirmiştir.

Hidir DUZKAYA: He performed the simulation studies, analyzed the results, and carried out the writting process.

Benzetim çalışmalarını gerçekleştirdi, sonuçlarını analiz etti ve yazım işlemini gerçekleştirmiştir.

CONFLICT OF INTEREST (ÇIKAR ÇATIŞMASI)

There is no conflict of interest in this study.

Bu çalışmada herhangi bir çıkar çatışması yoktur.

REFERENCES (KAYNAKLAR)

[1] International Energy Agency, Electricity 2024: Analysis and Forecast to 2026, (2024) 1-170.

[2] Hasan, M. M., & Wyseure, G., Impact of climate change on hydropower generation in Rio Jubones Basin, Ecuador, Water Science, and Engineering, 11(2), 157-66, (2018).

[3] Ullah, A., Topcu, A., Doğan, M. & Imran, M., Exploring the nexus among hydroelectric power generation, financial development, and economic growth: Evidence from the most significant 10 hydroelectric power-generating countries, Energy Strategy Reviews, 52, 101339, (2024).

[4] I. W. E. T. O. C. Pathway, World Energy Transitions Outlook 2023: 1.5°C Pathway, Abu Dhabi, 2023.

[5] Beheshti, M., Heidari, A., & Saghafian, B., Susceptibility of hydropower generation to climate change: Karun III Dam case study, Water, 11(5), 1025, (2019).

[6] Lumbroso, D., Woolhouse, G., & Jones, L., A review of the consideration of climate change in the planning of hydropower schemes in sub-Saharan Africa, Climatic Change, 133, 621-633, (2015).

[7] Zhao, X., Huang, G., Li, Y., & Lu, C., Responses of hydroelectricity generation to streamflow drought under climate change, Renewable and Sustainable Energy Reviews, 174, 113-141, (2023).

[8] De Souza Dias, V., Pereire da Luz, M., Medero, G. M., & Tarley Ferreira Nascimento, D., An overview of hydropower reservoirs in Brazil: Current situation, future perspectives and impacts of climate change, Water, 10(5), 592, (2018).

[9] Huang, J., Cang, J., Zhou, Z., & Gholinia F., Evaluation effect climate parameters change on hydropower production and energy demand by RCPs scenarios and the Developed Pathfinder (DPA) algorithm, Energy Reports, 7, 5455-5466, (2021).

[10] Fang, W., Chen, Y., & Xue, Q., Survey on research of RNN-based spatio-temporal sequence prediction algorithms, Journal on Big Data, 3(3), 97-110, (2021).

[11] Wang, H., Wu, X., & Gholinia, F., Forecasting hydropower generation by GFDL-CM3 climate model and hybrid hydrological-Elman neural network model based on Improved Sparrow Search Algorithm (ISSA), Concurrency and Computation: Practice and Experience, 33(24), e6476, (2021).

[12] Huangpeng, Q., Huang, W., & Gholinia, F., Forecast of the hydropower generation under influence of climate change based on RCPs and Developed Crow Search Optimization Algorithm, Energy Reports, 7, 385-397, (2021).

[13] Boadi, S. A., & Owusu, K., Impact of climate change and variability on hydropower in Ghana, African Geographical Review, 38(1), 19-31, (2019).

[14] Hamududu, B. H., & Killingtveit, Å., Hydropower production in future climate scenarios; the case for the Zambezi River, Energies, 9(7), 502, (2016).
[15] Uamusse, M. M., Tussupova, K., & Persson, K. M., Climate change effects on hydropower in Mozambique, Applied Sciences, 10(14), 4842, (2020).
[16] Khaniya, B., Karunanayake, C., Gunathilake, M. B., & Rathnayake, U., Projection of future hydropower generation in samanalawewa power plant, Sri Lanka, Mathematical Problems in Engineering, 1-11, (2020).
[17] Shrestha, A., Shrestha, S., Tingsanchali, T., Budhathoki, A., & Ninsawat, S., Adapting hydropower production to climate change: A case study of Kulekhani Hydropower Project in Nepal, Journal of Cleaner Production, 279, 123483, (2021).
[18] Karakuş, M. Ö., Impact of climatic factors on the prediction of hydroelectric power generation: a deep CNN-SVR approach, Geocarto International, 38(1), 2253203, (2023).
[19] Babacan, H. T., & Yüksek, Ö., Investigation of climate change impacts on daily streamflow extremes in Eastern Black Sea Basin, Turkey, Physics and Chemistry of the Earth, Parts A/B/C, 134, 103599, (2024).
[20] Benzer, S., New Record Of The Kizilirmak Killifish (Aphanius Marassantensis Pfleiderer, Geiger & Herder, 2014) From Süreyyabey Dam Lake In Yeşilirmak Basin, Mugla Journal of Science and Technology, 4(1), 41-45, (2018).
[21] Yılmaz, D., Süreyyabey Barajı Dolusavak Kazı Şevlerinde Yaşanan Zemin Hareketlerinin Analizi, Muş Alparslan Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 4(1), 23-31, (2023).
[22] Salcedo-Sanz, S., Cornejo-Bueno, L., Prieto, L., Paredes, D., & García-Herrera, R., Feature selection in machine learning prediction systems for renewable energy applications, Renewable and Sustainable Energy Reviews, 90, 728-741, (2018).
[23] Rao, S., Poojary, P., Somaiya, J., & Mahajan, P., A comparative study between various preprocessing techniques for machine learning, Int. J. Eng. Appl. Sci. Technol, 5(3), 2455-2143, (2020).
[24] Gogtay, N.J., & Thatte, U. M., Principles of correlation analysis, Journal of the Association of Physicians of India, 65(3), 78-81, (2017).
[25] Ratner, B., The correlation coefficient: Its values range between+ 1/− 1, or do they?, Journal of targeting, measurement and analysis for marketing, 17(2), 139-142, (2009).
[26] Asuero, A. G., Sayago, A., & González, A., The correlation coefficient: An overview, Critical reviews in analytical chemistry, 36(1), 41-59, (2006).
[27] Duzkaya, H., & Beroual, A., Statistical analysis of AC dielectric strength of natural ester-based ZnO nanofluids, 14(1), 99, (2020).
[28] Victoria A. H., & Maragatham, G., Automatic tuning of hyperparameters using Bayesian optimization, Evolving Systems, 12(1), 217-223, (2021).
[29] Yu, T., & Zhu, H., Hyper-parameter optimization: A review of algorithms and applications, arXiv preprint arXiv:2003.05689, (2020).
[30] Shrestha, A., & Mahmood, A., Review of deep learning algorithms and architectures, IEEE Access, 7, 53040-53065, (2019).
[31] Yang, L., & Shami, A., On hyperparameter optimization of machine learning algorithms: Theory and practice, Neurocomputing, 415, 295-316, (2020).
[32] Wu, J., Chen, X. Y., Zhang, H., Xiong, L. D., Lei, H., & Deng, S. H., Hyperparameter optimization for machine learning models based on Bayesian optimization, Journal of Electronic Science and Technology, 17(1), 26-40, (2019).
[33] Habtemariam, E. T., Kekeba, K., Martínez-Ballesteros, M., & Martínez-Álvarez, F., A Bayesian optimization-based LSTM model for wind power forecasting in the Adama district, Ethiopia, Energies, 16(5), 2317, (2023).
[34] Sherstinsky, A., Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network, Physica D: Nonlinear Phenomena, 404, 132306, (2020).
[35] Sahoo, B. B., Jha, R., Singh, A., & Kumar, D., Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting, Acta Geophysica, 67(5), 1471-1481, (2019).
[36] Wang, Y., Zhu, S., & Li, C., Research on multistep time series prediction based on LSTM, in 2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), 1155-1159 (2019).
[37] Jailani, N. L. M, Dhanasegaran, J. K., Alkawsi, G., Alkahtani, A. A., Phing, C. C., Baashar, Y., & Tiong, S. K., Investigating the power of LSTM-based models in solar energy forecasting, Processes, 11(5), 1382, (2023).

[38] Fan, H., Jiang, M., Xu, L., Zhu, H., Cheng, J., & Jiang, J., Comparison of long short term memory networks and the hydrological model in runoff simulation, Water, 12(1), 175, (2020).

[39] Ekanayake, P., Wickramasinghe, L., Jayasinghe, J. J. W., & Rathnayake, U., Regression-based prediction of power generation at samanalawewa hydropower plant in Sri Lanka using machine learning, Mathematical Problems in Engineering, 1-12, (2021).

[40] Gzar, D. A., Mahmood, A. M., & Abbas, M. K., A Comparative Study of Regression Machine Learning Algorithms: Tradeoff Between Accuracy and Computational Complexity, Mathematical Modelling of Engineering Problems, 9(5), (2022).

[41] Wu, J.,& Cheng, E., A novel hybrid particle swarm optimization for feature selection and kernel optimization in support vector regression, in 2010 International Conference on Computational Intelligence and Security, 189-194, (2010).

[42] Ly, H. -B., Nguyen, T, -A., & Pham, B. T., Estimation of soil cohesion using machine learning method: A random forest approach, Advances in civil engineering, 1, 8873993, (2021).

[43] Javed, U., Fraz, M. M., Mahmood, I., Shahzad, M., & Arif, O., Forecasting of electricity generation for hydro power plants, in 2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET), 2020: IEEE, 32-36 (2020).