

Article

A Hybrid Multi-Objective Optimizer-Based SVM Model for Enhancing Numerical Weather Prediction: A Study for the Seoul Metropolitan Area

Mohanad A. Deif ¹, Ahmed A. A. Solyman ², Mohammed H. Alsharif ³, Seungwon Jung ⁴
and Eenjun Hwang ^{4,*}

¹ Department of Bioelectronics, Modern University for Technology and Information (MTI), Cairo 11571, Egypt; Mohand.Deif@eng.mti.edu.eg

² Department of Electrical and Electronics Engineering, Istanbul Gelisim University, Avclar 34310, Turkey; aaasahmed@gelisim.edu.tr

³ Department of Electrical Engineering, College of Electronics and Information Engineering, Sejong University, Seoul 05006, Korea; malsharif@sejong.ac.kr

⁴ School of Electrical Engineering, Korea University, Seoul 02841, Korea; jsw161@korea.ac.kr

* Correspondence: ehwang04@korea.ac.kr

Abstract: Temperature forecasting is an area of ongoing research because of its importance in all life aspects. However, because a variety of climate factors controls the temperature, it is a never-ending challenge. The numerical weather prediction (NWP) model has been frequently used to forecast air temperature. However, because of its deprived grid resolution and lack of parameterizations, it has systematic distortions. In this study, a gray wolf optimizer (GWO) and a support vector machine (SVM) are used to ensure accuracy and stability of the next day forecasting for minimum and maximum air temperatures in Seoul, South Korea, depending on local data assimilation and prediction system (LDAPS; a model of local NWP over Korea). A total of 14 LDAPS models forecast data, the daily maximum and minimum air temperatures of in situ observations, and five auxiliary data were used as input variables. The LDAPS model, the multimodal array (MME), the particle swarm optimizer with support vector machine (SVM-PSO), and the conventional SVM were selected as comparison models in this study to illustrate the advantages of the proposed model. When compared to the particle swarm optimizer and traditional SVM, the Gray Wolf Optimizer produced more accurate results, with the average RMSE value of SVM for T max and T min Forecast prediction reduced by roughly 51 percent when combined with GWO and 31 percent when combined with PSO. In addition, the hybrid model (SVM-GWO) improved the performance of the LDAPS model by lowering the RMSE values for T max Forecast and T min Forecast forecasting from 2.09 to 0.95 and 1.43 to 0.82, respectively. The results show that the proposed hybrid (GWO-SVM) models outperform benchmark models in terms of prediction accuracy and stability and that the suggested model has a lot of application potentials.

Keywords: support vector machine (SVM); gray wolf optimizer (GWO); temperature forecasting; numerical weather prediction (NWP) model



Citation: Deif, M.A.; Solyman, A.A.A.; Alsharif, M.H.; Jung, S.; Hwang, E. A Hybrid Multi-Objective Optimizer-Based SVM Model for Enhancing Numerical Weather Prediction: A Study for the Seoul Metropolitan Area. *Sustainability* **2022**, *14*, 296. <https://doi.org/10.3390/su14010296>

Academic Editor: Amir Mosavi

Received: 30 November 2021

Accepted: 23 December 2021

Published: 28 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The weather has a considerable impact on the daily life of all living things, including humans and animals, and numerous industry sectors, that is why weather forecasting is one of the most regularly explored disciplines [1]. Because temperature is so closely linked to energy generation and agricultural operations, it is the most important weather factor [2,3] as low and high temperatures can affect agricultural activities, precise temperature forecasting is essential for avoiding crops damage [4]. However, because weather parameters, including temperature, are continuous, multi-dimensional, data-dense, chaotic, and dynamic, precisely predicting temperature is always difficult [5,6].

There are two types of models used in weather forecast research: physics-based models and data-based models. First, the effects of atmospheric dynamics, thermal radiation, and the influence of green spaces, lakes, and oceans were investigated numerically using physics-based weather forecasting models. Most public and commercial weather forecasting systems use physics-based models [1,7]. Data-driven models, on the other hand, predict the weather using statistics or algorithms based on machine learning.

Data-driven models have the advantage that they can recognize unexpected patterns in the weather system with no prior knowledge. However, it is possible that large amounts of data are generated and it is not understood how the models work. Physical techniques have the advantage that they can be easily understood and extrapolated from observable situations. The disadvantage is that it requires well-defined prior information and a lot of computing power [8,9]. Data-driven models for weather forecasting have recently been explored with greater intensity due to an increment in the number of features and observations.

Numerical weather prediction models (NWP), based on physical correlations of parameters and principles of atmospheric dynamics, have become an important tool for predicting numerous meteorological components, including air temperature. Because of their coarse grid resolution and imprecise physical parameterizations, NWP models tend to simplify the precise properties of terrestrial, atmospheric, and ocean systems. Due to incorrect physical parameterization, incorrect initial/boundary conditions, and domain and resolution dependence, uncertainties in NWP models lead to model distortions in air temperature predictions, despite continued advances in model performance. As a result, post-processing of the model output may be necessary to remove distortion for operational use of the models. Several statistical approaches were used to correct for bias in the air temperature data obtained by the NWP models. As a result, post-processing of the model output may be necessary to remove distortion for operational use of the models. Several statistical approaches have been used to correct bias in obtained air temperature data provided by the NWP models [10–12]. To improve forecasting effectiveness, these tactics applied to weather elements generated in NWP models in different countries.

The most widely used methods to correct bias in air temperature prediction are the Model Output Statistics (MOS) and Kalman Filter (KF) approaches. The MOS improves the prediction precision by using a statistical linear model derived between the results of the historical model and the observational data for the output of the NWP model [12]. Thanks to recent advances in computing resources, KF is now widely used to solve non-linear problems. KF first corrects the NWP model result by predicting the air temperature. The parameters of the next forecast phase are changed recurrently using the observed air temperatures [13,14]. Despite the use of a variety of machine learning techniques to reduce temperature drift, improving modeling accuracy remains a challenge. Recently, some researchers have attempted to improve predictive performance by combining the results of various machine learning algorithms in a variety of areas [15–18]. The results of all these researches show that integrating multiple machine learning models improves performance by overcoming the limitations of each classifier separately. Since machine learning methods are not affected by multicollinearity in input variables, they can process a large number of them. Unlike MOS and KF, which require bias correction to generate a model for each station, machine learning can be used to build a model that works for many stations. When spatially continuous input variables are introduced into machine learning models, the spatial distributions of predictions can be tracked.

Different learning algorithms (Support Vector Regression (SVR) and Random Forest (RF)) were used to correct the bias in the air temperature outputs of the NWP model. Eccel et al. (2007) [19] evaluated two machine learning approaches (ANN and RF) to improve the cryogenic forecasting capabilities of two NWP models, ECNWF and Local Area Model Italy (LAMI), in an Italian Alpine region. They found RF gave the best results compared to other methods, with the added benefit of being easy to automate. Yi et al. (2018) [20] improved the accuracy of air temperature from the Local Data Assimilation and Prediction

System (LDAPS) model in Seoul, South Korea, by using SVR and a linear regression model, finding that SVR showed higher correction accuracy than the linear regression model. In addition, the most widely used technology for predicting air temperature is the artificial neural network (ANN) [21]. Marzban (2003) [22] used ANN for post-processing of the Advanced Regional Prediction System (ARPS) model's hourly temperature outputs, obtaining an average 40% reduction in the mean squared error for all validated weather stations. Vashani et al. (2010) [23] found that the ANN and KF methods show better bias correction performance than the other methods for the summing accuracy of 30 weather stations in Iran, and ANN produced slightly higher accuracy than KF for longer forecast ranges. Zjavka (2016) [24] reported that a polynomial neural network could successfully bias-correct the National Oceanic and Atmospheric Administration (NOAA) meso-scale model to forecast hourly air temperature. To correct for bias in air temperature estimates from the European Center for Mid-Range Weather Forecasts (ECMWF) model, Isaksson (2018) [25] compared a deep neural network with KF. During this research, it was found that in most of the verified stations the neural network model exceeds KF in terms of error reduction. To correct for LDAPS, Dongjin Cho et al. (2020) [14] used a multi-model set (MME) and other machine learning techniques, with the MME model outperforming the other algorithms in terms of generalizability.

Using a support vector machine algorithm-based multi-objective grey wolf optimizer (GWO-SVM), this study seeks to eliminate distortion in LDAPS air temperatures, one of the NWP model outputs produced by the Korea Meteorological Administration (KMA). To our knowledge, no studies have been performed on improving the prediction of air temperature derived from the NWP model through an optimized multi-objective approach based on a machine learning algorithm. As a result, the major contributions of this paper are:

1. Developing a hybrid model (GWO-SVM) to improve the forecasting of the daily maximum and minimum air temperatures produced by the NWP model;
2. The proposed optimizer model (GWO) is compared with benchmark optimizers regarding the prediction accuracy and stability of SVM algorithm;
3. Examine the proposed model's forecasting in comparison with other machine learning approaches.

The rest of this work is arranged in the following manner. The suggested model is implemented in Section 2, which introduces the required theories, describes the collected data and prediction stages, and describes the proposed model's implementation. Section 3 presents the prediction results and significant challenges raised by this research. Finally, the conclusion is presented in Section 4.

2. Materials and Methods

2.1. Materials

The dataset was obtained from the UCI online resource [26] and used to correct for bias in the next day's maximum and minimum air temperatures expected by the Korea Meteorological Administration's LDAPS model for Seoul, South Korea. This dataset contains data for the summer from 2013 to 2017. It includes fourteen weather forecast data for the Numerical Weather Forecast (NWP), two in situ observations, and five geographic ancillary variables over Seoul, South Korea, throughout the summer. In this study, 14 forecast data from the LDAPS model, forecast data for the next day, and five auxiliary variables were used as input variables, being the maximum and minimum air temperatures of the next day (Forecast T max and Forecast T min) as objective variables (Table 1).

Table 1. Description of variables used in this study.

Variable Type	Abbreviation (unit)	Description
The variable that predicted using LDAPS	Tmax_LDAPS (°C)	Maximum air temperature
	Tmin_LDAPS (°C)	Minimum air temperature
	RHmax_LDAPS (%)	Maximum relative humidity
	RHmin_LDAPS (%)	Minimum relative humidity
	AWS_LDAPS (m/s)	Average wind speed
	LHF_LDAPS (W/m ²)	average latent heat flux
	CC1_LDAPS (%)	The average cloud cover during the next day's 6 h split (0–5 h)
	CC2_LDAPS (%)	The average cloud cover during the next day's 6 h split (6–11 h)
	CC3_LDAPS (%)	The average cloud cover during the next day's 6 h split (12–17 h)
	CC4_LDAPS (%)	The average cloud cover during the next day's 6 h split (18–23 h)
	PPT1_LDAPS (%)	The next day's precipitation averaged over six hours (0–5 h)
	PPT2_LDAPS (%)	The next day's precipitation averaged over six hours (6–11 h)
PPT3_LDAPS (%)	The next day's precipitation averaged over six hours (12–17 h)	
PPT4_LDAPS (%)	The next day's precipitation averaged over six hours (18–23 h)	
In situ data	T max_present(°C)	Present maximum air temperature
	T min_present(°C)	Present minimum air temperature
Auxiliary data	Lat_Location (°)	Latitude
	Log_Location (°)	Longitude
	ELEV_Topographic (m)	Elevation
	Slop_Topographic (°)	Slope
	SR_Topographic (wh/m ²)	Daily solar radiation

2.2. Preliminaries

2.2.1. Support Vector Machine (SVM)

SVM is a machine learning model that is commonly used. It is ideally suited to small sample sizes and has a strong statistical base [27]. In the disciplines of energy, ecology, hydrology, and economics, SVM has a wide range of applications [28–32]. In a regression issue, the training set is defined as [33,34]

$$\{(x_j, y_j) \mid x_j, y_j \in R^n, j = 1, 2, \dots, n\} \quad (1)$$

where x_j is the input and y_j is the output. The SVM model's detailed form is:

$$f(x) = \omega^T \phi(x) + c \quad (2)$$

where ω is the weighted vector, $\phi(x)$ is the nonlinear mapping function; c is the deviator.

In the SVM model, Two hyper-parameters that influence prediction performance are the kernel width and the penalty factor.

2.2.2. Multi-Objective Grey Wolf Optimizer

The grey wolf optimizer [35] is the foundation for building a multi-objective grey wolf optimizer (GWO). The GWO algorithm is a meta-heuristic algorithm based on wolf hunting behavior [9,36]. Every wolf in the herd has the potential to solve the problem. The ideal, suboptimal, and alternate solutions are represented by the four levels of the wolf swarm. Wolves approach their prey when they find it. Its position equations are:

$$\begin{cases} \vec{solve}J = |\vec{M} \cdot \vec{L}_p(p) - \vec{L}_w(p)| \\ \vec{L}_w(p+1) = \vec{L}_p(p) - \vec{N} \cdot \vec{J} \end{cases} \quad (3)$$

where the separation distance between the prey and the wolf is given by \vec{J} ; \vec{M} and \vec{N} are coefficient vectors; the position vectors of the grey wolf and the prey are \vec{L}_w and \vec{L}_p ; the current iteration is given by p .

GWO saves the top three solutions and uses Equations (4) and (5) to identify the optimum solution and continuously update the position of the grey wolf.

$$\begin{aligned} \vec{J}_\alpha &= |\vec{M}_1 \cdot \vec{L}_\alpha(p) - \vec{L}_w(p)| \\ \vec{J}_\beta &= |\vec{M}_2 \cdot \vec{L}_\beta(p) - \vec{L}_w(p)| \\ \vec{J}_\gamma &= |\vec{M}_3 \cdot \vec{L}_\gamma(p) - \vec{L}_w(p)| \\ \vec{L}_1 &= \vec{L}_\alpha(p) - \vec{N}_1 \vec{J}_\alpha \\ \vec{L}_2 &= \vec{L}_\beta(p) - \vec{N}_2 \vec{J}_\beta \\ \vec{L}_3 &= \vec{L}_\gamma(p) - \vec{N}_3 \vec{J}_\gamma \end{aligned} \quad (4)$$

$$\vec{L}_p(p+1) = \frac{1}{3}(\vec{L}_1 + \vec{L}_2 + \vec{L}_3) \quad (5)$$

where α , β , and γ are different levels of grey wolves.

The newly created individual is compared to the archived individual after each iteration. Furthermore, all individuals are categorized depending on the distance of the objective function value to avoid an overabundance of similar individuals. Second, the selection procedure of the leader wolf has shifted. Overcoming the problem of directly selecting three non-dominant solutions using the Pareto technique [37] by using roulette to choose the archive's leader wolf. Equation (6) can calculate the probability of each hypercube [38].

$$P_i = L_i^{-c} \quad (6)$$

where c is a constant; L_i is the number of Pareto optimal solutions; P_i is the probability of the hypercube.

In this paper, GWO has been used among other optimization algorithms because the advantages of GWO are as follows [35,39]: easy to implement due to its simple structure; less storage and computation requirements; faster convergence due to continuous reduction in search space; fewer decision variables; and ability to avoid local minimums. With only two control parameters to adjust the performance of the algorithm, which insures better stability and avoids complexity.

2.3. Proposed Approach

As shown in Figure 1, the prediction system includes three phases: (1) data preprocessing, (2) developing an optimization and prediction system, and (3) performance analysis.

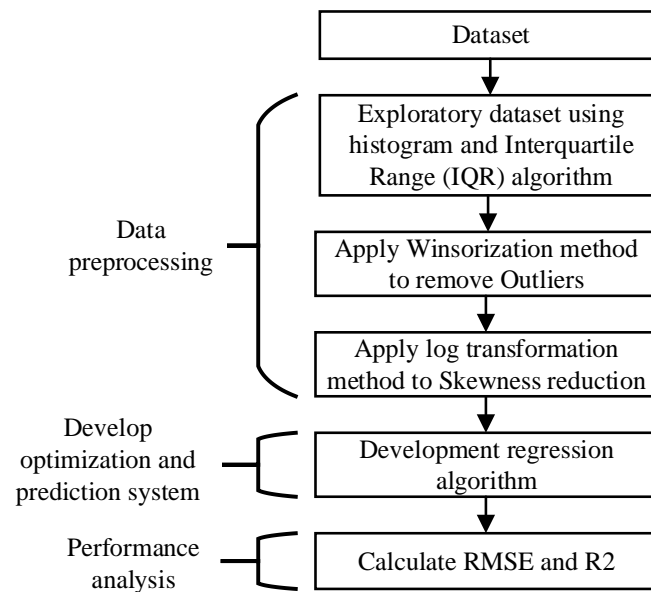


Figure 1. Overall methodology steps.

2.3.1. Data Preprocessing

The main purpose of this step is to remove the artifacts from the dataset (variables that have many missing data points, outlier data, and skewed data) to improve prediction system performance. After that, split the dataset into two parts. One part is used for training the regression model, while the other part is used for the final evaluation of the model. Data processing steps are shown in the following:

(a) Exploratory Data Analysis (EDA)

EDA is a common approach [39] for explaining the fundamental characteristics of a dataset by studying the characteristics, usually using visual methods. Histogram and the Interquartile Range (IQR) algorithm were employed to investigate and seek information about dataset artifacts.

(b) Removing the Outliers

Outliers were eliminated by applying winsorization [40], a statistical modification that reduces the impact of potential outliers by limiting extreme values in the data. This research investigates different threshold values for removing outliers from data. It will be discussed in the discussion section.

(c) Skewness Reduction

Data skewness has a significant impact on the predictive model's accuracy. Skewed data have a distribution that is pushed to one side or the other rather than being normally distributed. Wherefore, to improve accuracy, skewness should be removed from the variables. A log transformation is used to reduce skewness. A log transformation belongs to the more general family of Box–Cox transformations Box–Cox, 1964) [41], a Box–Cox transformation T^λ is defined as

$$T^\lambda(x) = \frac{x^\lambda - 1}{\lambda} \quad (7)$$

where x is a positive variable.

2.3.2. Development Regression Algorithm

After the preprocessing is complete, the processed data are entered into the SVM-GWO hybrid regression system. This model comprises two operations (training and optimization of the regression model). These two operations are synchronized and executed in the training set. Both SVM training and SVM optimization are carried out simultaneously.

When the optimization is complete, the SVM training is also completed. Academics often create an objective function to reduce the training set's prediction error in the traditional optimization problem (single-objective optimization). Since the multi-objective optimization used in this article considers both the precision and the stability of the prediction, two objective functions are defined:

$$m = \begin{cases} Obj_{Acc} = RMSE_{\text{training}} = \frac{1}{S_t} \sum_k | \frac{A_k - P_k}{A_k} | \\ Obj_{St} = \text{std}(A_k - P_k) \end{cases} \quad (8)$$

where Obj_{St} and Obj_{Acc} are the objective stability and prediction precision functions, respectively; RMSE training is the RMSE in the training set; S_t is the size of the training set's sample; A_k and P_k are the actual and predicted values at time k ; std is the population's standard deviation. Figure 2 shows the Flow chart of the hybrid regression system (SVM-GWO).

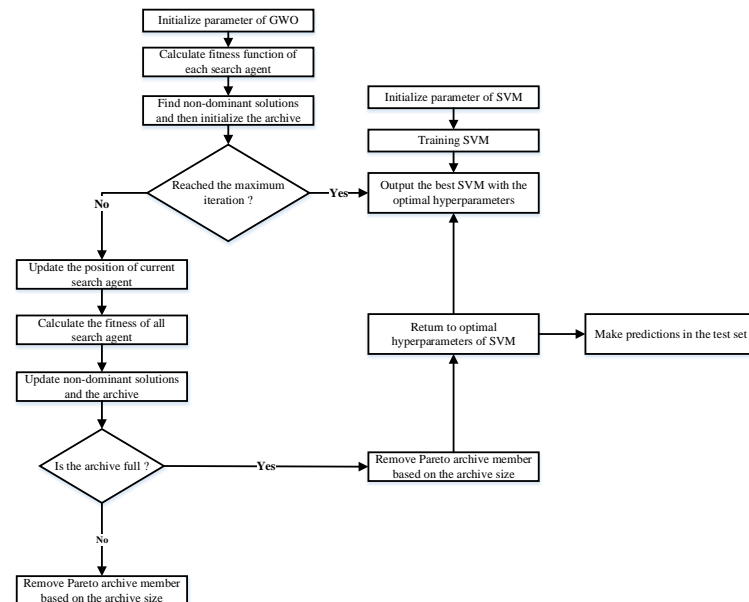


Figure 2. Flow chart of hybrid regression system.

2.3.3. Performance Analysis

The prediction performance is measured using two commonly used error metrics: RMSE (Root-mean-square error) and R^2 (The adjusted coefficients of determination). The following are their expressions:

$$R^2 = 1 - \frac{\sum_{i=1}^n (P_i - A_i)^2}{\sum_{i=1}^n A_i^2} \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - A_i)^2} \quad (10)$$

where P_i and A_i have predicted values, the actual number of the i th record, and n is the total number of records (61).

3. Results and Discussion

This part will exhibit the performance of the suggested optimization and prediction system for minimum and maximum temperature forecasts, as well as the results obtained from preprocessing data.

Figure 3 shows the box plot of variables; $T_{max_present}$, $T_{min_present}$, RH_{min_LDAPS} , T_{max_LDAPS} , T_{min_LDAPS} , and LHF_{LDAPS} . There are a lot of outliers in the figures, as you can see. The box plot is a data preparation method for detecting extreme values and outliers. It computes dispersion by dividing a rank-ordered dataset into four equal portions known as quartiles [9]. Q1, Q2, and Q3 are the values that divide each section, with Q1 and Q3 denoting the middle values in the first and second halves of the rank-ordered dataset, respectively, and Q2 denoting the median value for the entire set. The Interquartile Range (IQR) is calculated by subtracting Q3 from Q1. Data instances that fall below $Q1 + 1.5$ IQR or above $Q3 + 1.5$ IQR are considered outliers. Getting rid of these outliers enhances prediction accuracy. As a result, data that are over three standard deviations from the mean value is discarded. We used winsorization to remove outliers, and we experimented with multiple threshold values ranging from 3 to 4.8. We found the suitable values that achieved with high prediction performance without loss large dataset was 4.2 (the dataset reduced by 5.8%).

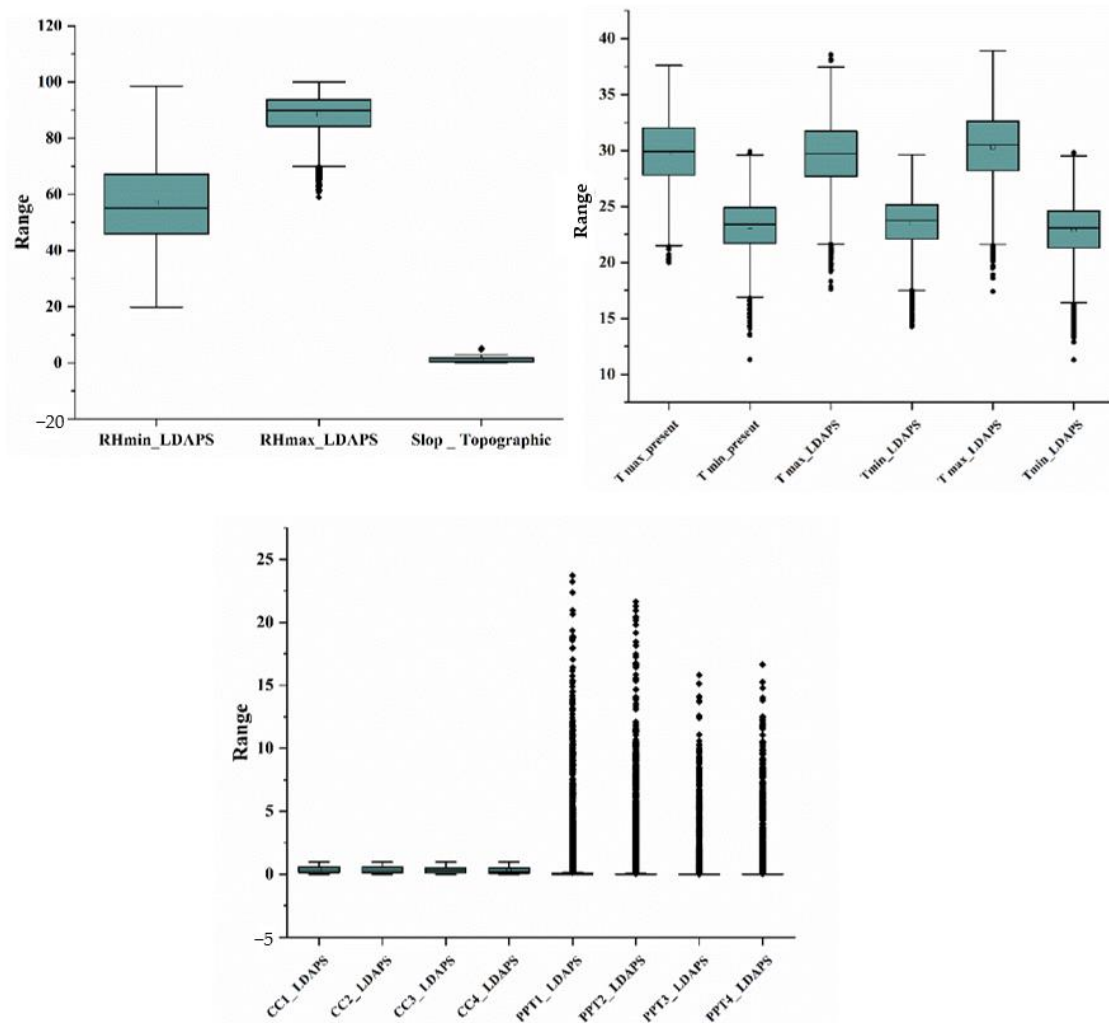


Figure 3. Box plot of dataset variables before outliers removal.

Figure 4 shows the histogram plot for T_{max_LDAPS} , T_{min_LDAPS} , $T_{max_present}$, $T_{min_present}$, RH_{max_LDAPS} , RH_{min_LDAPS} , $T_{max_present}$, and $T_{min_present}$ are almost normally distributed; $T_{max_present}$ have the highest temp as 37 °C and min as 20 °C, most of the days have maximum temperature equal to 28 °C, while $T_{min_present}$ have the highest temp as 30 °C and min as 11 °C, most of the days have minimum temperature equal to 23.

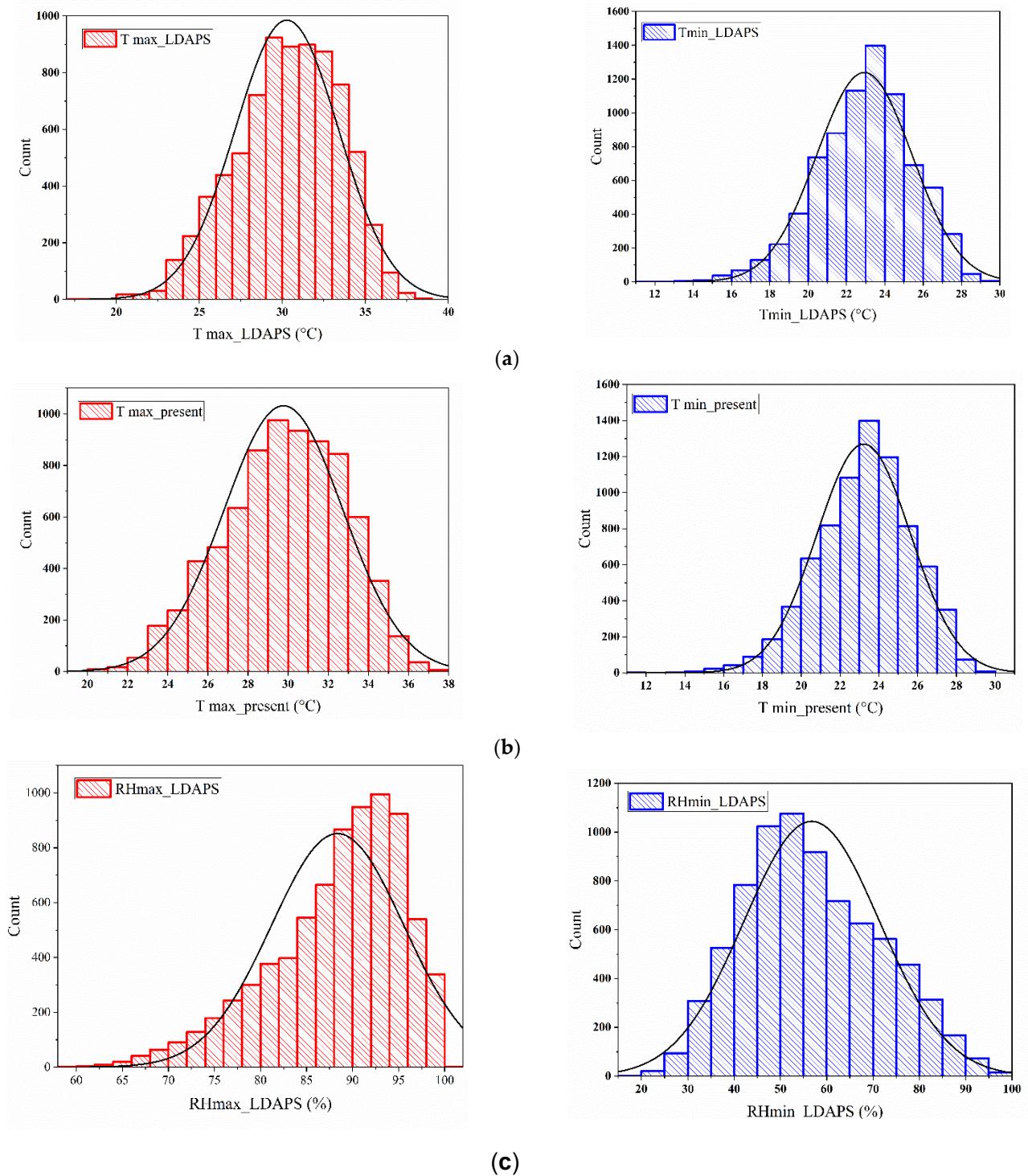


Figure 4. Histogram data distribution of (a) T_{max_LDAPS} and T_{min_LDAPS} (b) $T_{max_present}$ and $T_{min_present}$ (c) RH_{max_LDAPS} and RH_{min_LDAPS} .

T_{max_LDAPS} and T_{min_LDAPS} are almost normally distributed, with T_{max_LDAPS} having a maximum at 38.54 °C, T_{min_LDAPS} at 17 °C and for the majority of days its values lie in the range 27 to 32 °C while T_{min_LDAPS} having a maximum at 29.61 °C, minimum at 14 °C and for the majority of days its values lie in the range 23 to 26 °C. Although data of RH_{max_LDAPS} are left-skewed and data of RH_{min_LDAPS} are slightly right-skewed. RH_{max_LDAPS} for most of the days lies in the range 92 to 97 °C, while RH_{min_LDAPS} lies in the range 45 to 62 °C.

Figure 5 shows the AWS_{LDAPS} is right-skewed. It has its minimum value at 2.88 m/s and maximum values at 21.85 m/s and the majority of its values lie in the 5 m/s to 8 m/s. LHF_{LDAPS} seems to be normally distributed with slight skewness to the right. It has its minimum value at -13 and maximum values at 213 and most of its values lie between 30 to 70. $SR_{Topographic}$ variables left-skewed. It has a minimum value of 4329 and a maximum value of 5992 and the majority of its values lie in the range 5600 to 5850.

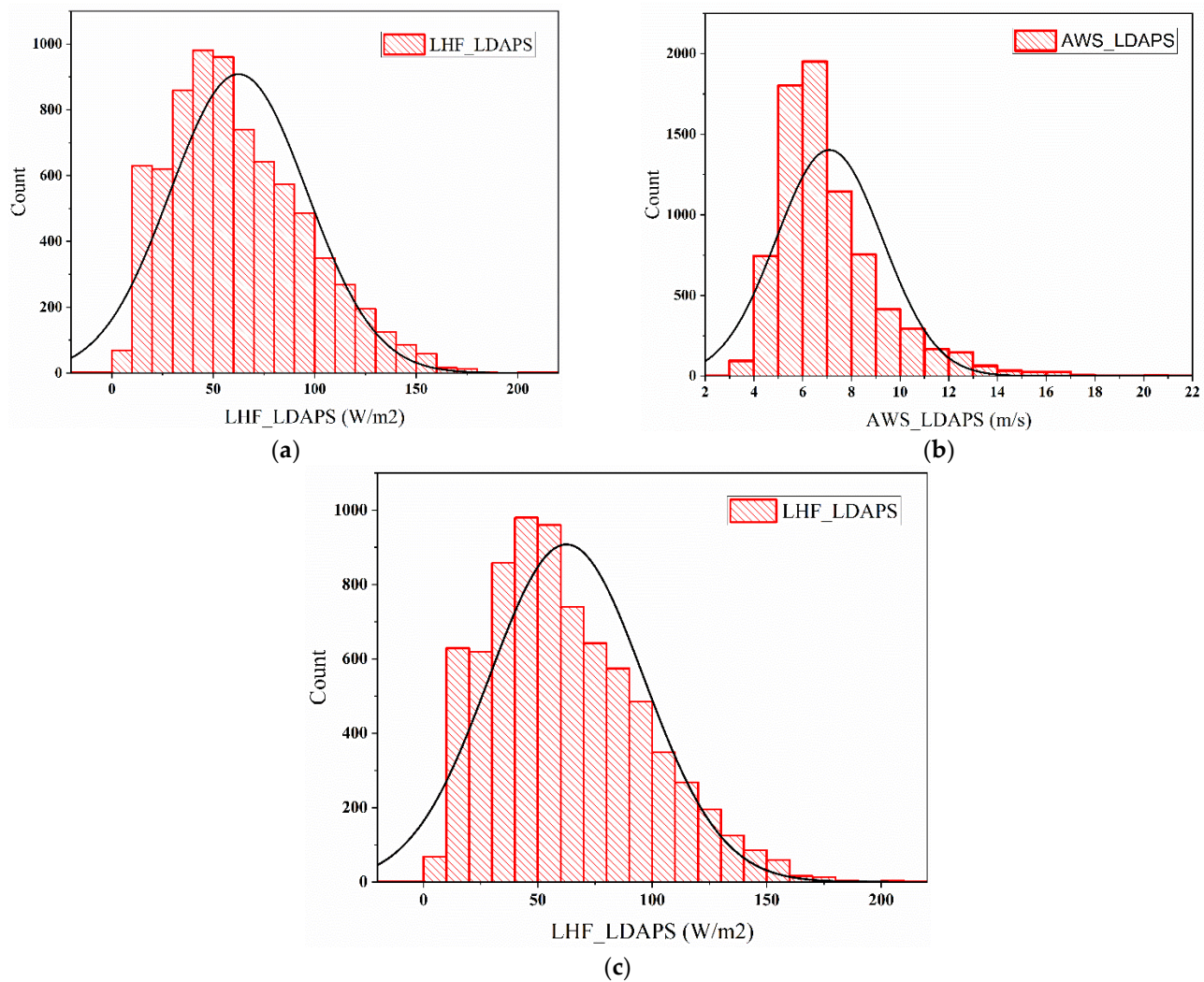
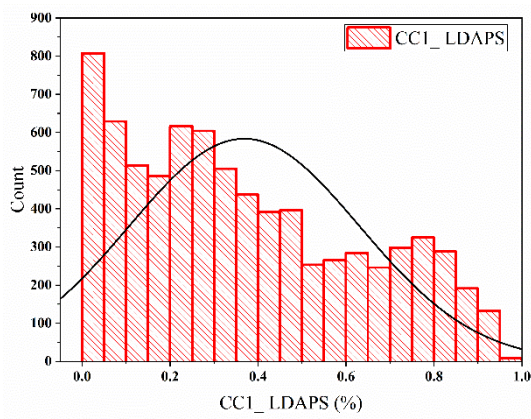


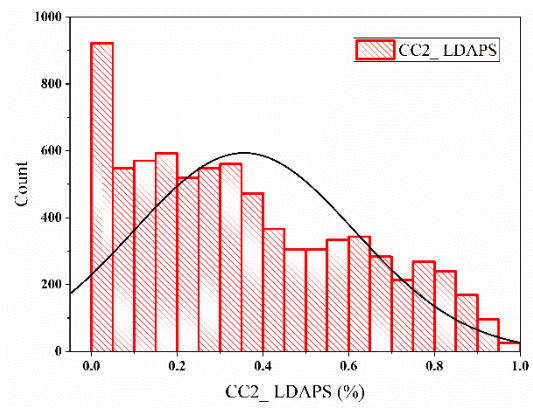
Figure 5. Histogram data distribution of (a) AWS_{LDAPS} , (b) LHF_{LDAPS} , and (c) $SR_{Topographic}$.

Figure 6. shows the cloud cover data ($CC1_{LDAPS}$, $CC2_{LDAPS}$, $CC3_{LDAPS}$, and $CC4_{LDAPS}$). All 6-h splits are right-skewed, and the majority of these splits' values are close to zero.

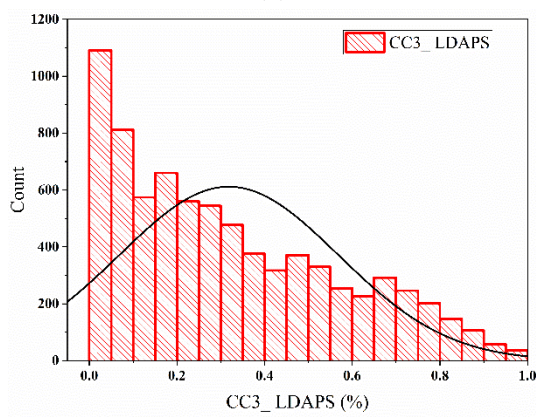
Figure 7 shows the variables' data distribution after applying a log transformation technique to remove skewness. We note the Skewness is removed compared to Figures 4–6. Removing data skewness has a significant effect on the accuracy of our predictive model.



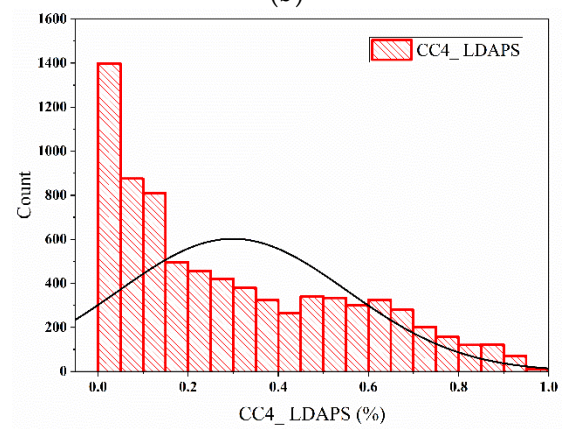
(a)



(b)



(c)



(d)

Figure 6. Histogram data distribution of the next-day 6-h split average of (a) 0–5 h, (b) 6–11 h, (c) 12–17 h, and (d) 18–23 h.

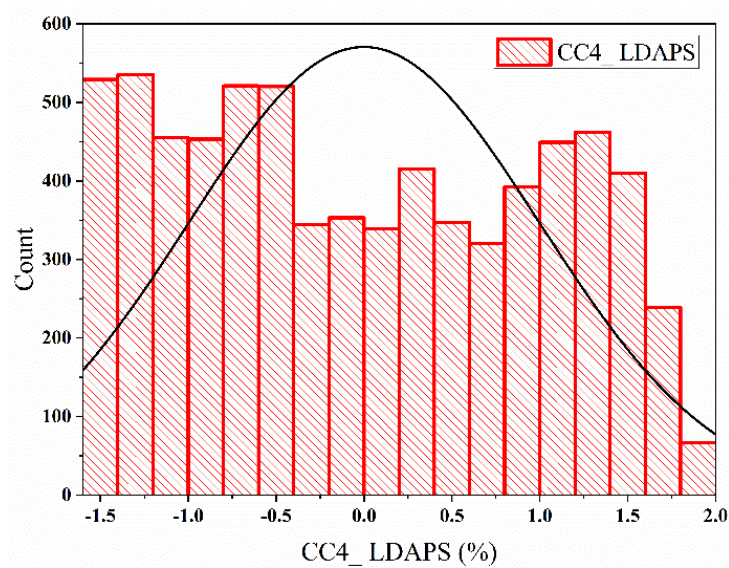


Figure 7. Cont.

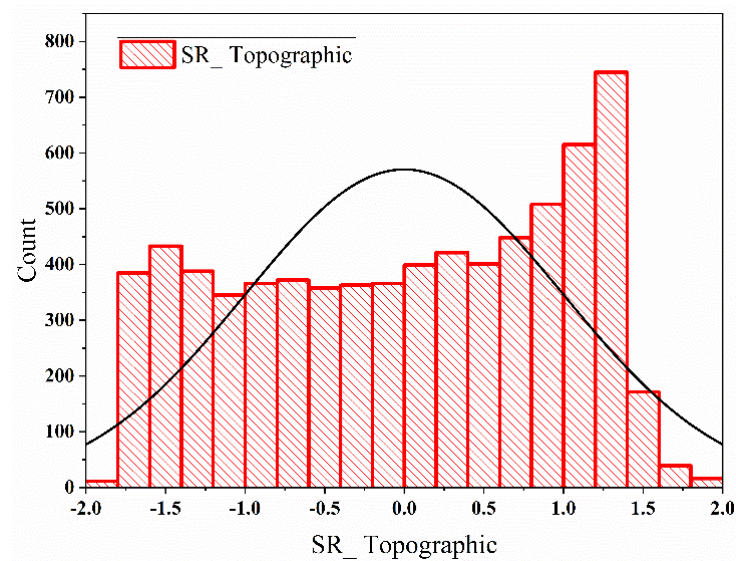


Figure 7. Illustration of example histogram data distribution for $SR_{Topographic}$ and $CC4_{LDAPS}$ variables after removing skewness.

To emphasize the benefits of the proposed model, this paper defines two models as the benchmark models for comparison with our proposed method. The first one of these models is the classic SVM model, the second model is the PSO with SVM. Reference [42] shows the theory of various models, as well as the reasons for selecting them.

The forecasting accuracy of the proposed model (SVM-GWO) and the selected benchmark models (SVM and SVM-PSO) are shown in Figure 8. All models were trained using data from 2013 to 2016, and weather data for 2017 was expected. For the prediction of next-day maximum and lowest air temperatures in 2017, the RMSE values for all models were determined.

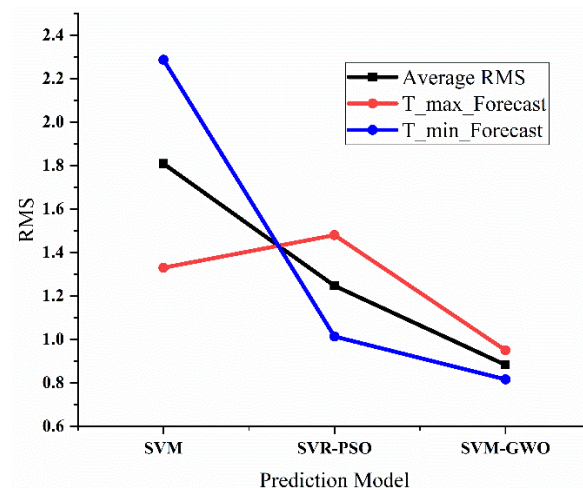


Figure 8. The impact of different optimization strategies on SVM prediction accuracy. Forecasting maximum and minimum temperatures in 2017.

In Figure 8, we can observe from the average RMS line for Tmax and Tmin Forecast, the performance of SVM prediction has improved when using PSO and GWO. When SVM is paired with GWO, the average RMSE prediction is lowered by around 51%, and when SVM is combined with PSO, the RMSE is decreased by about 31%. These findings backed up the GWO-SVM model's superior prediction capacity, leading to the conclusion that the suggested model has the best prediction accuracy of all the models tested.

In addition, a scatter plot was created and displayed in Figure 9 to show correlations between the actual observations and predictions. We observe the proposed model predicted temperatures were strongly correlated, with R^2 values of 0.91 in $T_{\max_Forecast}$ and 0.93 in $T_{\min_Forecast}$, and weakly correlated when using SVM-PSO with R^2 values of 0.56 in $T_{\max_Forecast}$ and 0.51 in $T_{\min_Forecast}$.

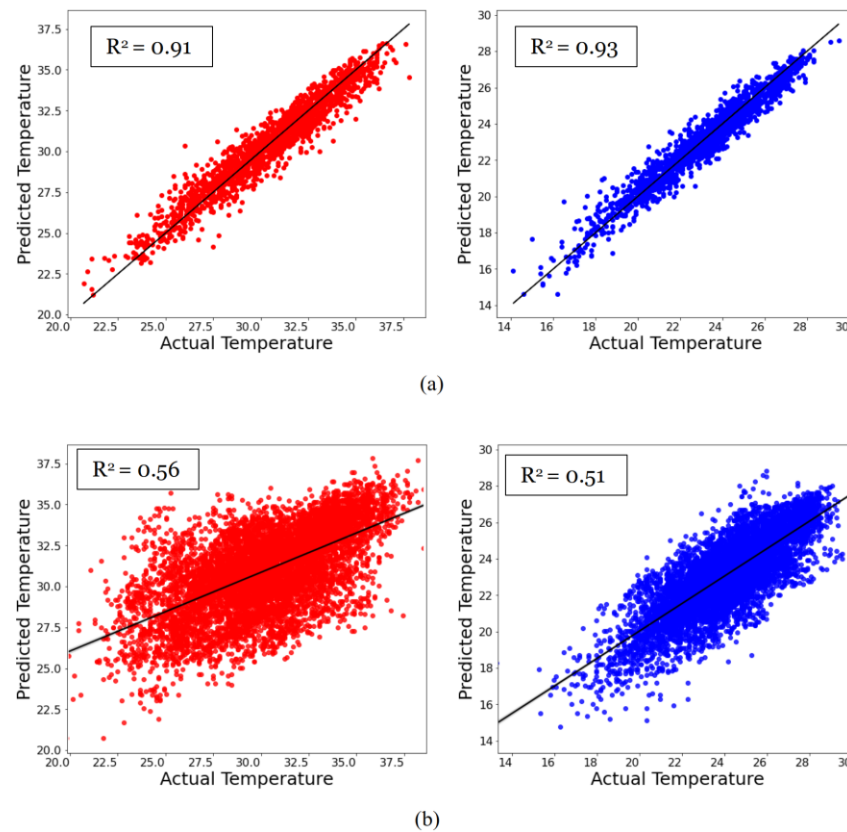


Figure 9. Scatter plot and the simple linear regression line for the proposed hybrid model, and the selected benchmark models to forecast 2017. (a) SVM-GWO and (b) SVM-PSO.

Yearly Hindcast validation for our proposed model for predicting both $T_{\max_Forecast}$ and $T_{\min_Forecast}$ was performed to support the prior findings, and it was compared to a previous work model published by Dongjin Cho et al. [14]. The author had previously predicted these values based on other machine learning models, and it was concluded that the multi-model ensemble (MME) model performed better in terms of generalization than the LDAPS model and other hindcast validation machine learning models. In addition, the (SVM-GWO) model's findings were compared to those of the LDAPS model. From 2015 to 2017, hindcast validation was performed for each year. The data were used to train prediction models that forecasted the period until the end of the year from January 1 to July 31 for each year. Table 2 shows the annual validation of three models using Hindcast.

In particular, they had the highest improved performance in 2016 for forecasting $T_{\max_Forecast}$ and 2017 for forecasting $T_{\min_Forecast}$ because the SVM-GWO model forecast had the lowest RMSE 0.93 in 2016 for $T_{\max_Forecast}$ and 0.696 in 2017 for $T_{\min_Forecast}$. Whereas the MME reference model had the lowest RMSE 1.45 for forecasting $T_{\max_Forecast}$ in 2016 and 0.84 for forecasting $T_{\min_Forecast}$ in 2017. Nevertheless, it remains, in most years, our model had a lower RMSE than the other models for forecasting both $T_{\max_Forecast}$ and $T_{\min_Forecast}$. Generally, the SVM-GWO reduces the RMSE value from 2.09 to 0.95 and 1.43 to 0.82 for forecasting both $T_{\max_Forecast}$ and $T_{\min_Forecast}$, respectively. This result confirms the results of the high predictive capacity of the hybrid model presented and leads to the conclusion that the (SVM-GWO) model offers better and more accurate predictions

compared to traditional machine learning models and can be a reliable method to predict future max and minimum air temperatures.

Table 2. RMSE comparison between the proposed model, LDAPS model, and the reference model for prediction of forecasting both $T_{max_Forecast}$ and $T_{min_Forecast}$.

$T_{max_Forecast}$			
Year	LDAPS	MME	SVM-GWO
2015	2.07	1.53	0.94
2016	2.15	1.45	0.93
2017	2.04	1.65	0.98
Average RMSE	2.09	1.54	0.95
$T_{min_Forecast}$			
Year	LDAPS	MME	SVM-GWO
2015	1.47	1.05	0.896
2016	1.43	1.03	0.856
2017	1.39	0.84	0.696
Average RMSE	1.43	0.97	0.82

Figure 10 shows the daily RMSE and R^2 time-series values of the LDAPS and the proposed model SVM-GWO were compared for the last year of the investigation period (2017). The time-series represents by day of the year (DOY) with RMSE and R^2 values for LDAPS and proposed SVM-GWO model for forecast $T_{max_Forecast}$ and $T_{min_Forecast}$. Both $T_{max_Forecast}$ and $T_{min_Forecast}$ forecasts by SVM-GWO generally showed a lower daily RMSE than the LDAPS model Figure 10a,c. This is because the time-series of the SVM-GWO corrected temperatures were closer to the observations. For the $T_{max_Forecast}$ forecast, the lowest RMSE for SVM-GWO model was DOY 227 and the highest RMSE was DOY 212. Although for the $T_{max_Forecast}$, the lowest and highest RMSE were DOY 197 and 222, respectively.

Figure 10b,d Generally, The SVM-GWO model had a higher R^2 than LDAPS and MEE mod for both $T_{max_Forecast}$ and $T_{min_Forecast}$ forecasts, As a result, the SVM-GWO model accurately simulates the temperature distribution within a metropolis.

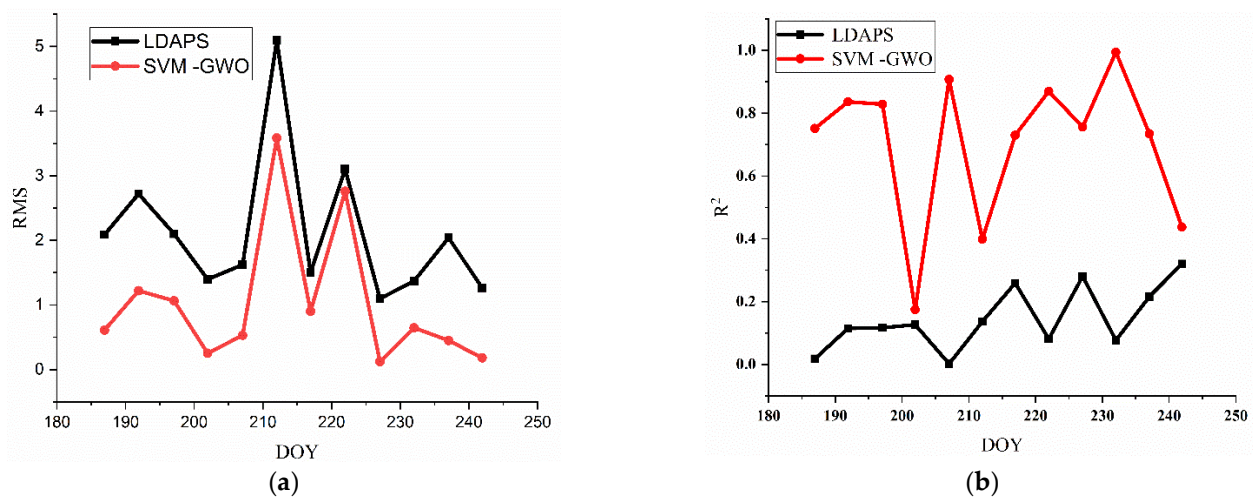


Figure 10. Cont.

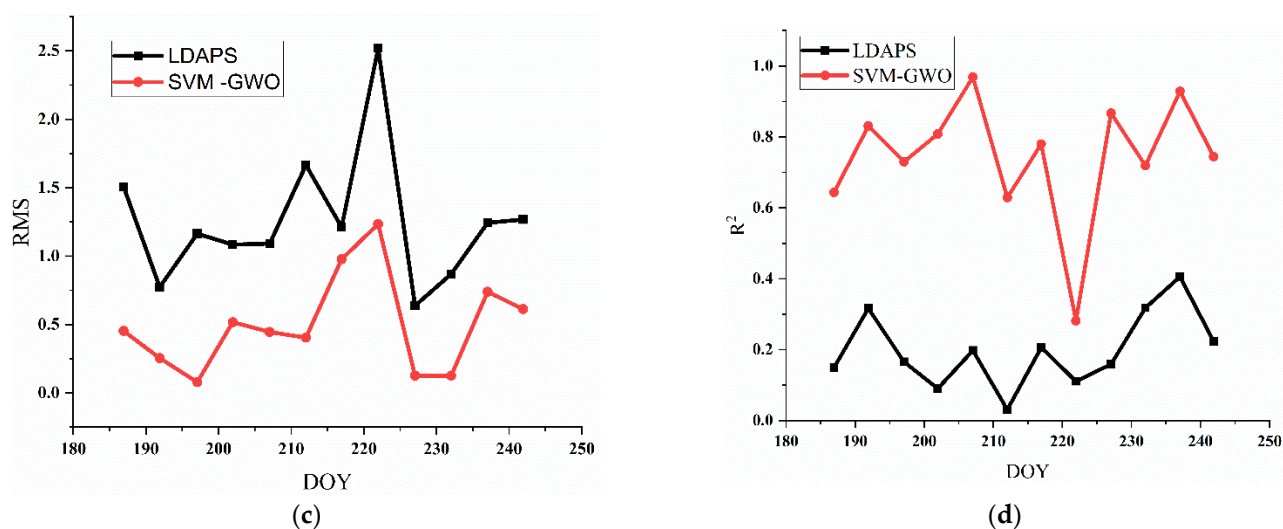


Figure 10. Time-series of the daily RMSE and R^2 values of the LDAPS and proposed a SVM-GWO model for (a,b) $T_{max_Forecast}$, (c,d) for $T_{min_Forecast}$.

4. Conclusions

The LDAPS model outputs of $T_{max_Forecast}$ and $T_{min_Forecast}$ in the Seoul Metropolitan Area are improved using a hybrid model that includes a gray wolf optimizer (GWO) and a support vector machine (SVM). The forecast models were created by using 14 LDAPS model forecast data, in situ observations' $T_{max_Forecast}$ and $T_{min_Forecast}$, and five auxiliary data as input variables. The four machine learning algorithms and the LDAPS model were evaluated using hindcast validation. When compared to the PSO and conventional SVM, the Gray Wolf Optimizer showed its strength by generating more stable and accurate results, with the average RMSE of SVM for T_{max} and T_{min} prediction lowered by roughly 51% when combined with GWO and 31% when combined with PSO. In addition, the hybrid model (SVM-GWO) improved the performance of the LDAPS model by lowering the RMSE values for $T_{max_Forecast}$ and $T_{min_Forecast}$ forecasting from 2.09 to 0.95 and 1.43 to 0.82, respectively. Despite the need for further research with other NWP models, this strategy is likely to be successful when applied to other NWP models for the study area that can deterministically predict next-day temperatures.

Author Contributions: Conceptualization, M.A.D.; Data curation, A.A.A.S. and S.J.; Formal analysis, A.A.A.S.; Funding acquisition, E.H.; Investigation, A.A.A.S., M.H.A. and E.H.; Methodology, M.A.D.; Project administration, M.H.A.; Resources, M.H.A. and S.J.; Software, M.A.D.; Supervision, E.H.; Writing—original draft, M.A.D. and A.A.A.S.; Writing—review & editing, M.H.A., E.H. and S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Korea Environment Industry & Technology Institute (KEITI) through the Exotic Invasive Species Management Program, funded by the Korea Ministry of Environment (MOE) (2021002280004).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alsharif, M.H.; Kim, J.; Kim, J.H. Opportunities and Challenges of Solar and Wind Energy in South Korea: A Review. *Sustainability* **2018**, *10*, 1822. [[CrossRef](#)]
2. Alsharif, M.H.; Younes, M.K.; Kim, J. Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea. *Symmetry* **2019**, *11*, 240. [[CrossRef](#)]

3. Alsharif, M.H.; Younes, M.K. Evaluation and Forecasting of Solar Radiation using Time Series Adaptive Neuro-Fuzzy Inference System: Seoul City as A Case Study. *IET Renew. Power Gener.* **2019**, *13*, 1711–1723. [CrossRef]
4. Durai, V.R.; Bhradwaj, R. Evaluation of statistical bias correction methods for numerical weather prediction model forecasts of maximum and minimum temperatures. *Nat. Hazards* **2014**, *73*, 1229–1254. [CrossRef]
5. Goswami, K.; Hazarika, J.; Patowary, A.N. Monthly Temperature Prediction Based On Arima Model: A Case Study In Dibrugarh Station Of Assam, India. *Int. J. Adv. Res. Comput. Sci.* **2017**, *8*.
6. Deif, M.A.; Solyman, A.A.A.; Hammam, R.E. ARIMA Model Estimation Based on Genetic Algorithm for COVID-19 Mortality Rates. *Int. J. Inf. Technol. Decis. Mak.* **2021**, 1–24. [CrossRef]
7. Candy, B.; Saunders, R.W.; Ghent, D.; Bulgin, C.E. The Impact of Satellite-Derived Land Surface Temperatures on Numerical Weather Prediction Analyses and Forecasts. *J. Geophys. Res. Atmos.* **2017**, *122*, 9783–9802. [CrossRef]
8. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef]
9. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]
10. Anadranistakis, M.; Lagouvardos, K.; Kotroni, V.; Elefteriadis, H. Correcting temperature and humidity forecasts using Kalman filtering: Potential for agricultural protection in Northern Greece. *Atmos. Res.* **2004**, *71*, 115–125. [CrossRef]
11. de Carvalho, J.R.P.; Assad, E.D.; Pinto, H.S. Kalman filter and correction of the temperatures estimated by PRECIS model. *Atmos. Res.* **2011**, *102*, 218–226. [CrossRef]
12. Stensrud, D.J.; Yussouf, N. Short-Range Ensemble Predictions of 2-m Temperature and Dewpoint Temperature over New England. *Mon. Weather. Rev.* **2003**, *131*, 2510–2524. [CrossRef]
13. Libonati, R.; Trigo, I.; DaCamara, C. Correction of 2 m-temperature forecasts using Kalman Filtering technique. *Atmos. Res.* **2008**, *87*, 183–197. [CrossRef]
14. Cho, D.; Yoo, C.; Im, J.; Cha, D. Comparative Assessment of Various Machine Learning-Based Bias Correction Methods for Numerical Weather Prediction Model Forecasts of Extreme Air Temperatures in Urban Areas. *Earth Space Sci.* **2020**, *7*, e2019EA000740. [CrossRef]
15. Chou, J.-S.; Pham, A.-D. Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength. *Constr. Build. Mater.* **2013**, *49*, 554–563. [CrossRef]
16. Healey, S.P.; Cohen, W.B.; Yang, Z.; Brewer, C.K.; Brooks, E.B.; Gorelick, N.; Hernandez, A.J.; Huang, C.; Hughes, M.J.; Kennedy, R.E.; et al. Mapping forest change using stacked generalization: An ensemble approach. *Remote. Sens. Environ.* **2018**, *204*, 717–728. [CrossRef]
17. Ren, J.; Song, K.; Deng, C.; Ahlgren, N.A.; Fuhrman, J.A.; Li, Y.; Xie, X.; Poplin, R.; Sun, F. Identifying viruses from metagenomic data using deep learning. *Quant. Biol.* **2020**, *8*, 64–77. [CrossRef] [PubMed]
18. Deif, M.A.; Hammam, R.E. Skin Lesions Classification Based on Deep Learning Approach. *J. Clin. Eng.* **2020**, *45*, 155–161. [CrossRef]
19. Eccel, E.; Ghielmi, L.; Granitto, P.; Barbiero, R.; Grazzini, F.; Cesari, D. Prediction of minimum temperatures in an alpine region by linear and non-linear post-processing of meteorological models. *Nonlinear Process. Geophys.* **2007**, *14*, 211–222. [CrossRef]
20. Yi, C.; Shin, Y.; Roh, J.-W. Development of an Urban High-Resolution Air Temperature Forecast System for Local Weather Information Services Based on Statistical Downscaling. *Atmosphere* **2018**, *9*, 164. [CrossRef]
21. Deif, M.A.; Solyman, A.A.; Kamarposhti, M.A.; Band, S.S.; Hammam, R.E. A deep bidirectional recurrent neural network for identification of SARS-CoV-2 from viral genome sequences. *Math. Biosci. Eng.* **2021**, *18*, 8933–8950. [CrossRef]
22. Marzban, C. Neural Networks for Postprocessing Model Output: ARPS. *Mon. Weather. Rev.* **2003**, *131*, 1103–1111. [CrossRef]
23. Vashani, S.; Azadi, M.; Hajjam, S. Comparative Evaluation of Different Post Processing Methods for Numerical Prediction of Temperature Forecasts over Iran. *Res. J. Environ. Sci.* **2010**, *4*, 305–316. [CrossRef]
24. Zjavka, L. Numerical weather prediction revisions using the locally trained differential polynomial network. *Expert Syst. Appl.* **2016**, *44*, 265–274. [CrossRef]
25. Isaksson, R. *Reduction of Temperature Forecast Errors with Deep Neural Networks—Reducering av Temperaturprognosfel med Djupa Neuronnätverk*; Department of Earth Sciences, Uppsala University: Uppsala, Sweden, 2018.
26. Dua, D.; Graff, C. {UCI} Machine Learning Repository. 2017. Available online: <https://archive.ics.uci.edu/ml/datasets/Bias+correction+of+numerical+prediction+model+temperature+foreca> (accessed on 10 October 2021).
27. Suykens, J.A.K.; Vandewalle, J. Least Squares Support Vector Machine Classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]
28. Kucharski, A.J.; Russell, T.W.; Diamond, C.; Liu, Y.; Edmunds, J.; Funk, S.; Eggo, R.M.; Centre for Mathematical Modelling of Infectious Diseases COVID-19 Working Group. Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *Lancet Infect. Dis.* **2020**, *20*, 553–558. [CrossRef]
29. Yan, H.; Zhang, J.; Rahman, S.; Zhou, N.; Suo, Y. Predicting permeability changes with injecting CO₂ in coal seams during CO₂ geological sequestration: A comparative study among six SVM-based hybrid models. *Sci. Total. Environ.* **2020**, *705*, 135941. [CrossRef] [PubMed]
30. Sadeghi, R.; Zarkami, R.; Sabetraftar, K.; Van Damme, P. Use of support vector machines (SVMs) to predict distribution of an invasive water fern *Azolla filiculoides* (Lam.) in Anzali wetland, southern Caspian Sea, Iran. *Ecol. Model.* **2012**, *244*, 117–126. [CrossRef]

31. Zhao, L.-T.; Zeng, G.-R. Analysis of Timeliness of Oil Price News Information Based on SVM. *Energy Procedia* **2019**, *158*, 4123–4128. [[CrossRef](#)]
32. Deif, M.A.; Solyman, A.A.A.; Alsharif, M.H.; Uthansakul, P. Automated Triage System for Intensive Care Admissions during the COVID-19 Pandemic Using Hybrid XGBoost-AHP Approach. *Sensors* **2021**, *21*, 6379. [[CrossRef](#)]
33. Liu, M.; Cao, Z.; Zhang, J.; Wang, L.; Huang, C.; Luo, X. Short-term wind speed forecasting based on the Jaya-SVM model. *Int. J. Electr. Power Energy Syst.* **2020**, *121*, 106056. [[CrossRef](#)]
34. Deif, M.A.; Hamman, R.E.; Solyman, A.A.A. Gradient Boosting Machine Based on PSO for prediction of Leukemia after a Breast Cancer Diagnosis. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2021**, *11*, 508–515. [[CrossRef](#)]
35. Mirjalili, S.; Saremi, S.; Mirjalili, S.M.; Coelho, L.D.S. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Syst. Appl.* **2016**, *47*, 106–119. [[CrossRef](#)]
36. Al-Tashi, Q.; Rais, H.M.; Abdulkadir, S.J.; Mirjalili, S.; Alhussian, H. A Review of Grey Wolf Optimizer-Based Feature Selection Methods for Classification. In *Algorithms for Intelligent Systems*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 273–286.
37. Akbari, E.; Rahimnejad, A.; Gadsden, S.A. A greedy non-hierarchical grey wolf optimizer for real-world optimization. *Electron. Lett.* **2021**, *57*, 499–501. [[CrossRef](#)]
38. Zhang, Z.; Hong, W.-C. Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads. *Knowl. Based Syst.* **2021**, *228*, 107297. [[CrossRef](#)]
39. Deif, M.; Hammam, R.; Solyman, A. Adaptive Neuro-Fuzzy Inference System (ANFIS) for Rapid Diagnosis of COVID-19 Cases Based on Routine Blood Tests. *Int. J. Intell. Eng. Syst.* **2021**, *14*, 178–189. [[CrossRef](#)]
40. Nyitrai, T.; Virág, M. The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Econ. Plan. Sci.* **2019**, *67*, 34–42. [[CrossRef](#)]
41. Box, G.E.P.; Cox, D.R. An analysis of transformations. *J. R. Stat. Soc. Ser. B* **1964**, *26*, 211–243. [[CrossRef](#)]
42. Lu, H.; Ma, X.; Ma, M. A hybrid multi-objective optimizer-based model for daily electricity demand prediction considering COVID-19. *Energy* **2021**, *219*, 119568. [[CrossRef](#)]