International Journal on Advanced Science Engineering Information Technology
Vol.8 (2018) No. 5
ISSN: 2088-5334

Time Series Predictive Analysis based on Hybridization of Meta-heuristic Algorithms

Zuriani Mustaffa[#1], Mohd Herwan Sulaiman[*], Dede Rohidin[1], Ferda Ernawan[#], Shahreen Kasim[2]

[#]*Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia*
e-mail: zuriani@ump.edu.my; ferda@ump.edu.my

[*] *Faculty of Electrical & Electronics Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.*
e-mail: mherwan@ieee.org

[1]*School of Computing, Telkom University, 40257 Bandung, West Java, Indonesia*

[2]*Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, Malaysia*

***Abstract*— This paper presents a comparative study which involved five hybrid meta-heuristic methods to predict the weather five days in advance. The identified meta-heuristic methods namely Moth-flame Optimization (MFO), Cuckoo Search algorithm (CSA), Artificial Bee Colony (ABC), Firefly Algorithm (FA) and Differential Evolution (DE) are individually hybridized with a well-known machine learning technique namely Least Squares Support Vector Machines (LS-SVM). For experimental purposes, a total of 6 independent inputs are considered which were collected based on daily weather data. The efficiency of the MFO-LSSVM, CS-LSSVM, ABC-LSSVM, FA-LSSVM, and DE-LSSVM was quantitatively analyzed based on Theil's U and Root Mean Square Percentage Error. Overall, the experimental results demonstrate a good rival among the identified methods. However, the superiority goes to FA-LSSVM which was able to record lower error rates in prediction. The proposed prediction model could benefit many parties in continuity planning daily activities.**

***Keywords*— computational intelligence; least squares support vector machines; machine learning; meta-heuristic; optimization; swarm intelligence; time series prediction**

I. INTRODUCTION

Computational Intelligence (CI) algorithms have been extensively studied and used by researchers in addressing various real-life problems. It is becoming increasingly difficult to deny the efficiency of CI algorithms which includes machine learning and optimization techniques. These are proven by a good number of applications of these algorithms in the literature [1]-[3]. Since decades ago, the emerging of population-based optimization algorithms is undeniably encouraging. Among them are the well-known Genetic Algorithm (GA) [4] which its efficiency is undoubted.

Interestingly, since its invention, the GA is continuously studied and being a favorite in the academic community of interest. Years later, researchers started to pay attention to nature, where one after another nature-inspired optimization algorithms began to rise such as Particle Swarm Optimization (PSO) [5], Ant Coloy Optimization (ACO)) [6] Grey Wolf Optimizer (GWO) [7] and many others. These optimization algorithms which motivated based on nature-inspired concept deal with choosing the best alternative related to the provided objective function [8]. As an optimizer, these algorithms can be seen not only used as a single algorithm but also hybrid with another algorithm, such as machine learning technique namely Support Vector Machines (SVM), Artificial Neural Network (ANN) to name a few. A considerable amount of CI-based approaches has been widely presented in order to address many complex real-world issues, including prediction, has been published in the literature. Prediction model based on CI was surveyed which focused on prostate cancer [9]. In the study, cancer data of different modalities are considered. Meanwhile, a hybrid algorithm which is based on Support Vector Regression (SVR) and PSO was presented in [10] for a month ahead of the residential sector's electricity demand. The simulations results showed that the presented method was superior compared to Artificial Neural Network (ANN), Auto-Regressive Integrated Moving Average (ARIMA), Multiple Linear Regressions and a few other techniques reported in the literature. Progressing further, a study in [11] proposed kernel extreme learning machine based on MFO which is applied in medical diagnoses namely Parkinson and breast cancer cases. The findings of the study demonstrate that the proposed method can serve as an effective decision-making method for the case under study. Meanwhile, in 2014, a hybrid Support Vector Machine (SVM) optimized by PSO has been presented by [12] for real estate price prediction. In the study, the PSO is utilized as an optimization tool to optimize the SVM parameters automatically.

In this paper, hybrid predictive modeling of LSSVM with five optimization algorithms namely Moth-flame Optimization (MFO), Cuckoo Search Algorithm (CSA), Artificial Bee Colony (ABC), Firefly Algorithm (FA) and Differential Evolution (DE) is presented which is realized in weather prediction. The MFO, CSA, ABC, FA, and DE are employed individually as an optimizer to LSSVM hyper-parameters, which later are known as MFO-LSSVM, CSA-LSSVM, ABC-LSSVM, FA-LSSVM, and DE-LSSVM respectively. By employing LSSVM, one of the critical problems that need serious attention is that the generalization capability of LSSVM is highly reliant on the value sets to its hyper-parameters, namely regularization parameter and kernel parameter. Any inappropriate values set to these parameters would demote the prediction performance. These identified optimization algorithms are chosen due to several advantages which includes having a few tuning parameters [13] and its capability to address optimization problems. The identified hybrid algorithms are later realized on weather predictive analysis. It is well documented that weather prediction is vital as it would impact the lives of human being. Accurate weather prediction would help different types of businesses to plan for production, power consumption, and many other activities. Besides, it is also important for certain people who face health problems such as heat stress and asthma. By knowing the weather, proper plan and action can be taken to avoid unwanted situation to occur.

The remaining part of the paper structured as the followings: Section 2 describes on the fundamental theory of LSSVM, followed by Section 3 which presents the optimization of LSSVM based on the identified meta-heuristic algorithms while the hybridization of LSSVM with those meta-heuristic algorithms are discussed in Section 4. Meanwhile, the implemented methodology is discussed in Section 5 while results are discussed and analyzed in Section 6. Finally, Section 7 concludes the study.

II. MATERIAL AND METHOD

A. *Least Squares Support Vector Machines*

Introduced by Suykens and colleagues, LSSVM [14] is a new version of SVM which offers a better solution strategy with regards to solving the problem at hands. The improved version is proven to address the problem faced by SVM during training [15]. With the reformulation of LSSVM, it simplifies complex calculation which led to more comfortable and faster training task. With that, a simpler optimization problem can be obtained [16]. Also, LSSVM offers fewer control parameters, which are γ and σ^2, compared to three control parameters required in SVM (C, σ^2, and ε) [17]. In addition, regarding a prediction task, LSSVM is proven to be better than SVM [18].

Provided with a training set of N points $\{x_i, y_i\}^N$ where x_i and y_{ou} both are the input and output values respectively, the goal is to estimate a model based on the given equation:

$$y(x) = w^T \varphi(x_i) + b + e_i \tag{1}$$

Where w represents weight vector, $\varphi(.)$: R^n is the non-linear function, b indicates the bias and e_i served as the error between the actual and predicted output at the ith training data. X_i and $y(x)$ act as input and output, respectively. The followings can obtain the coefficient vector w and bias term b:

$$\min_{w,b,e} J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^{N} e_i^2 \tag{2}$$

Subject to the equality constraints

$$y_i = w^T \varphi(x_i) + b + e_i \,, \; i = 1, 2 \,, \ldots, N \tag{3}$$

By applying the Lagrangian multiplier to (2) produces:

$$L(w,b,e,\alpha) = J(w,e) - \sum_{i=1}^{N} \alpha_i \{ w^T \varphi(x_i) + b + e_i - y_i \} \tag{4}$$

Where αi are Lagrange multipliers and γ is the regularization parameter. Differentiating (3) with w, b, ei and αi, the conditions for optimality is as express in the following:

$$\begin{aligned}
&\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N} \alpha_i \varphi(x_i) \\
&\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{N} \alpha_i = 0 \qquad\qquad i = 1, 2, \ldots, N \\
&\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \\
&\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^T \varphi(x_i) + b + e_i - y_i = 0
\end{aligned} \tag{5}$$

The following linear equations are obtained by eliminating the w and ei:

$$\begin{bmatrix} 0 & y^T \\ y & \Omega + I/\gamma \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 1_v \end{bmatrix} \tag{6}$$

With y = [y1; …; yN], α=[α1, …, αN], I is the identity matrix and 1v =[1;…;1]. The kernel trick is applied as follows:

$$\begin{aligned}
\Omega il &= \varphi(x_i)^T \varphi(x_l) \qquad i, l = 1, \ldots, N \\
&= K(x_i, x_l)
\end{aligned} \tag{7}$$

The resulting of LSSVM model for regression in (1) becomes:

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \tag{8}$$

Where α and b are the solutions of (5). Considering the non-linear features of the employed data set, in this study, Radial Basis Function (RBF) kernel is used as it is suitable to handle such data12 and able to give good performance in many prediction cases13. It is expressed as:

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \tag{9}$$

Where σ2 is one of the adjustable parameters which related to RBF kernel. The other adjustable parameter of RBF kernel is the regularization parameter, γ which can be seen in (2).

In this study, the RBF kernel is employed considering its suitability in handling with nonlinear data [19]. Brief theory of LSSVM is presented here while a detailed description can be found in [14].

B. Meta-Heuristic Algorithms

This section is dedicated to present an optimization of LSSVM hyper-parameters based on identified meta-heuristic algorithms namely MFO, CSA, ABC, FA, and DE.

1) Moth-Flame Optimizer: Moth-flame optimization (MFO) [20] is a relatively recent nature-inspired optimization technique which is motivated by the uniqueness of moths' navigation method in nature. The navigation method is known as transverse orientation. They fly at night by maintaining a fixed angle concerning the moon where the moon is used to guide them in flying for long distances. Such guideline helps them in flying in a straight line. Nonetheless, moths are easily trapped in a deadly spiral path around artificial lights. The MFO algorithm was mathematically formulated based on this unique behavior for optimization purposes, in order to find the best possible solution(s) for a given problem. In this study, the capability of MFO is fully utilized as an optimizer to optimize the LSSVM hyper-parameters, which would increase the effectiveness of LSSVM in prediction. In the MFO algorithm, the successor solutions are represented by months while the position of the moths in space taking place the problem's variables. The moths can fly in various dimensions 1-D, 2D or even hyperdimensional space by changing their position vectors. The pseudo code of MFO-LSSVM is shown in Fig. 1.

2) Cuckoo Search Algorithm: Cuckoo Search Algorithm (CSA) is one of the promising meta-heuristic algorithms which was invented by Xin-She Yang [21]. CSA is inspired by brood parasitism of cuckoo species, which is a main mechanism of cuckoos in nature. In CSA, the quality or fitness of a solution is modeled proportional to the objective function value. In CSA, the candidate solution indicates by each egg in the nest. Meanwhile, the new solution represents by the cuckoo's egg. The aim is to serve the new and potentially better solutions to replace a less good solution in the nests [8]. The pseudo code of CSA-LSSVM is shown in Fig. 2.

```
Begin
Initialize the moths' positions
While (Iteration < MNI)//MNI: Maximum Number
  Iteration
  Update the number of flames
For (i=1; size(position_of_moths,1))
  Check if moth is out of search space, if yes, bring it back
  Calculate the fitness of the moths
End
If (iteration==1)
  Sort the first population of the moths
  Update the flames
Else
  Sort the moths
  Update the flames
End
  Update the position of the best flame obtained so far
For (i=1; size (position_of_moths,1))
  For (j=1; size(position_of_moths,2))
    If (i<=number_of_flame)
      Update the position of the moth with respect to its
      corresponding flame
    End
If(i>number_of_flame)
  Update the position of the moth with respect to only one
  flame
    End
  End
End
Output the result
End
```

Fig. 1 Pseudo Code of MFO-LSSVM

3) Artificial Bee Colony Algorithm: The ABC algorithm [13] is developed from the observation of the brilliant foraging habits of honey bees swarm. In the algorithm, the honey bee swarm consists of three groups of bees. The first group is known as an employed bee, the second group is onlooker bees, and lastly, the third group is scout bees. In nature, all of the bees are working to maximize the amount of nectar. Fig. 3 shows the pseudo code of ABC-LSSVM.

4) Firefly Algorithm: social habits of fireflies inspire Firefly Algorithm (FA) [21]. In natural life of fireflies, the flashing features are employed to attract the mating partners. Meanwhile, the movement of fireflies is driven by the resulting attraction, and the attractiveness is related to the intensity of the emitted light.

The pseudo code of FA is shown in Fig. 4.

5) Differential Evolution: Unlike the other methods, DE is classified as an Evolutionary Algorithm, the same class as Genetic Algorithm. The natural selection mechanism motivates DE [22]. In DE, the candidate solutions possess an equal chance to be evaluated. Fig. 5 shows the pseudo code of DE-LSSVM.

```
Objective function f(x), x = (x1, ..., xd)^T
Initialization
While(t<MNI)//MNI: Maximum Number Iteration
    Get a cuckoo randomly/generate solution by Levy
    flights and then evaluate hyper-parameters and
    calculate the fitness value based on training and
    validation sets using LSSVM
    Choose a nest among n (say, j) randomly
If (Fi > Fj),
Replace j by the new solution
   End
   A fraction (pa) of worse nests are abandoned and new
   ones/solutions are built/generated
   Keep best solutions (or nests with quality solutions)
   Rank the solutions and find the current best
End While
Post process results and visualization
```

Fig. 2 Pseudo Code of CSA-LSSVM

```
Cycle=1
Initialize the food source positions (solutions) xi=1,.., SN
Evaluate the nectar amount (fitness function fiti) of food
sources
REPEAT
Employed Bee (EB) Phase
For Each EB
       Produce new food source positions (solutions)
       Calculate the value fiti
   Memorized the best solution
End For
Calculate the probability values pi for the best solution

Onlooker bee (OB) phase
For each OB
   Choose the food source depending on pi
   Produce new food source (solutions)
   Calculate the value fiti
   Memorized the best solution
End For

Scout Bee (SB) Phase
Abandon its food source
Search for new ones to replace the abandoned food source
Memorized the best solution
Cycle=Cycle+1
Until MNI//MNI: Maximum Number Iteration
```

Fig. 3 Pseudo Code of ABC-LSSVM

```
Randomly initialize the population and generate N fireflies
(solutions) Xi, i=1,2,..., N
Compute the fitness value of each firefly
FEs=N
While FEs<= MAX_FEs do
   For i=1 to N do
      For j=1 to N do

      //Movement through attraction
      If f(Xj)< f(Xi) then
      Move Xi toward Xj
      Compute the fitness value of the New X
      Fes++
         End
      End
   End
End
```

Fig. 4 Pseudo Code of FA-LSSVM

```
Initialization
Set the weight F and crossover probability Cr
While (stopping criterion)
  For i=1 to n
          For each xi, randomly choose 3 distinct vectors
   xp, xr
      and xr
                Generate a new vector v
                Generate a random index Jr by
permutation
                Generate a randomly distributed number
                For j=1 to d
          For each parameter Vj,I (jth component of vi),
          update
```

$$u_{j,i}^{t+1} = \begin{cases} v_{j,i}^{t+1} \ if\ r_i \le C_r \ \ or\ \ j = J_r \\ x_{j,i}^{t} \ if\ r_i > C_r \ and\ j \ne J_r \end{cases}$$

```
          Evaluate hyper-parameters and calculate the
          fitness value based on training and validation sets
          using LSSVM
      End
      Select and update the solution
    End
  Update the counters such as t=t+1
End
Post-process and output the best solution found
```

Fig. 5 Pseudo Code of DE-LSSVM

C. Hybrid Meta-Heuristic with LSSVM Algorithm

As briefly described in the previous section, five optimization algorithms are employed distinctively as an optimization tool to LSSVM to tune the hyper-parameters of LSSVM automatically; γ and σ^2.

1) *Initialization:* Figure 6 shows the successor solution which are the two hyper-parameters of interest as illustrated in a matrix.

$$X = \begin{bmatrix} \gamma 1 & \sigma^2 2 \\ \cdots & \cdots \\ \gamma 1 & \sigma^2 2 \end{bmatrix} \begin{array}{l} \leftarrow \ \textit{Candidate solution 1} \\ \leftarrow \ \textit{Candidate solution} \ldots \\ \leftarrow \ \textit{Candidate solution 20} \end{array}$$

Fig. 6 Candidate Solution in *X*

In the proposed hybrid model, the LSSVM act in evaluating the fitness function and the optimal value of hyper-parameters of interest can be obtained after a termination criterion is hit. In this study, termination criterion s indicates by maximum number of iteration, which is set to 100 iteration. On the other hand, the objective function is guided by the selected statistical metric, Root Mean Square Error (RMSPE). The indicator of the RMSPE is, the smaller the value of the RMSPE, the better the result is.

2) Evaluation: In order to evaluate the objective function of the interested hyper-parameters, it is done based on training and validation set using LSSVM function, which is embedded in the five identified optimization algorithms. As describe previously, the objective function of the study is indicated by the selected statistical metric namely RMSPE. The aim is to find an ideal combination of γ and σ^2 which will result lowest RMSPE.

D. Experimental and performance criteria

Discussion in this section focuses on the employed data, experimental setup and the utilized performance criteria.

1) Data Acquisition and Pre-Processing: For empirical Purposes, real-time series data that associated with temperature are employed. The data collected including high, average and low values of temperatures (°C), dew points (°C), humidity, sea level pressure (hPa), visibility, winds (km/h) and precip (mm). All data are obtained from www.wunderground.com. The data frequency is on daily basis which is recorded at Sultan Abdul Aziz Shah-Subang, Kuala Lumpur, Malaysia. This data set is ranging from April 2016 to May 2017. Samples of the data are tabulated in Table 1.

TABLE I
SAMPLES OF DATA

Date	Temperature (°C)			Dew points (°C)			Humidity		
2016	High	Avg	Low	High	Avg	Low	High	Avg	Low
1 Apr	35	30	26	25	24	23	89	72	49
2 Apr	36	31	25	27	24	22	89	71	44
3 Apr	36	30	25	24	23	20	94	69	41
4 Apr	36	31	27	26	24	22	84	69	56
5 Apr	34	29	25	25	24	23	94	78	59

2) Data Normalization: For the sake of accuracy in prediction, the inputs and output will be normalized by using Zero Mean Normalization. The Zero Mean Normalization is defined as follow:

3) Input and Output Arrangement: Table 2 indicates the variables assigned to the features employed. For inputs, they are temperature, dew point, humidity, sea level pressure, visibility, wind speed, and precipitation. These inputs are fed to the prediction model to predict the temperature from day 5 ahead.

TABLE II
INPUT AND OUTPUT

	Input	Variables	Output
Temperature (°C)	Max	maxT	Temperature from day 5 ahead (temp5)
	Mean	meant	
	Min	minT	
Dew Point (°C)	Max	maxDP	
	Mean	meanDP	
	Min	minDP	
Humidity (%)	Max	maxH	
	Mean	meanH	
	Min	minH	
Sea level pressure (hPa)	Max	maxSLP	
	Mean	meanSLP	
	Min	minSLP	
Visibility (km)	Max	maxV	
	Mean	meanV	
	Min	minV	
Wind (km/h)	Max	maxW	
	Mean	meanW	
	Min	minW	
Precipitation (mm)	Sum	precip	

4) Training, Validation, and Testing: For experiment purposes, the data sets are divided into three separate subgroups. They are training, validation and testing set. The training set is important where it is used to fit the prediction model, the validation set on the other hand is employed during model assessment and to prevent overfitting as well. Meanwhile, the testing set is where a real assessment is done to evaluate on of how well the prediction model generalize.

5) Performance Evaluation Metrics: To assess the Performance of identified hybrid algorithms, two statistical indicators were employed namely Theil's U and Root Mean Square Error (RMSPE). These criteria served as an interpreter in assessing the learning and generalization of identified prediction models. The formal definition of and Theil's U and RMSPE as shown in (9) and (10):

$$Theil's U = \frac{\sqrt{\frac{1}{N}\Sigma_{n=1}^{N}(y_n - y(x_n))^2}}{\sqrt{\frac{1}{N}\Sigma_{n=1}^{N}(y_n)^2} + \sqrt{\frac{1}{N}\Sigma_{n=1}^{N}(y(x_n))^2}} \tag{9}$$

$$RMSPE = \sqrt{\left(\frac{\sum_{n=1}^{N}\left(\frac{y_n - y(x_n)}{y_n}\right)^2}{N}\right)} \tag{10}$$

where
$n = 1, 2, \ldots N$
y_n = target values
$y(x_n)$ = predicted values
N =Number of test data.

III. RESULTS AND DISCUSSION

The prediction model was realized on a real weather data in Kuala Lumpur. The data were collected from April 2016 to April 2017 as a complete time series with a time interval of 1 day.

By observing results recorded in Table 3, FA-LSSVM shows its superiority by recording the smallest value for both metrics, which are 0.0366% and 0.0172 respectively. These values are obtained by setting the γ=49.1056 and σ2=544.7772. This is followed by DE-LSSVM by producing 0.0370% for RMSPE, and 0.0174 for Theil's U. On the other hand, MFO-LSSVM rank at third when the value of γ is recorded to 1000 and σ2 is 2.2962. Finally, CSA-LSSVM and ABC-LSSVM were on the fourth and last place respectively. By referring to both RMSPE and Theil's U, both values showed a good agreement.

In finding optimal values of the hyper-parameters of LSSVM, it is important to understand that too large or too small of the values may result in over-fitting or underfitting issue, which eventually will lead to poor generalization. Based on the table, the over-fitting issues can be seen occurred to MFO-LSSVM and ABC-LSSVM when the values of γ for both hybrid algorithms reached the maximum value that has been set, which is 1000. This situation contributes to the unsatisfied results for both MFO-LSSVM and ABC-LSSVM.

TABLE III
MFO-LSSVM VS. CSA-LSSVM VS. ABC-LSSVM VS. FA-LSSVM VS. DE-LSSVM FOR WEATHER PREDICTION

	Gam	Sig2	RMSPE (%)	Theils'U
MFO-LSSVM	1000	2.2962	0.0389	0.0183
CSA-LSSVM	1.6475	12.4829	0.0413	0.0198
ABC-LSSVM	1000	5.3542	0.0421	0.0204
FA-LSSVM	49.1056	544.7772	0.0366	0.0172
DE-LSSVM	16.1491	227.7176	0.0370	0.0174

Meanwhile, Table 4 shows the prediction values by the identified hybrid algorithms during from day 363 to 372 (part of testing phase). The boldface values indicate the best performance on a respective day; the italic values represent the second best while the underlined values define the third best. Based on the listed values, it shows that the FA-LSSVM consistently be among the top 3 for most of the days.

TABLE IV
TARGET VS. MFO-LSSVM VS. CSA-LSSVM VS. ABC-LSSVM VS. FA-LSSVM VS. DE-LSSVM

Day	Target	MFO-LSSVM	CSA-LSSVM	ABC-LSSVM	FA-LSSVM	DE-LSSVM
363	28.0000	28.7326	**28.2778**	28.5397	28.3351	*28.2960*
364	28.0000	29.0686	28.9432	29.4053	*27.8044*	**27.9413**
365	28.0000	28.8042	28.5067	28.5354	**28.3543**	28.4188
366	28.0000	28.8278	28.2539	28.7269	*28.1723*	**28.1189**
367	28.0000	28.9960	28.8830	28.7778	**28.4419**	*28.4804*
368	28.0000	28.5275	**28.1705**	*28.2439*	28.5407	28.5240
369	29.0000	*28.8898*	28.5073	**29.1625**	28.5911	28.5189
370	29.0000	**28.7899**	28.4412	*28.6559*	28.3908	28.4187
371	29.0000	28.8676	29.0744	28.9515	**29.0202**	*29.0523*
372	29.0000	28.5214	28.5327	28.2336	28.6189	**28.6299**

Due to the superiority of FA-LSSVM over the other methods, a t-test was conducted to evaluate the statistical significance of the difference between the FA-LSSVM over the other identified algorithms. The t-test result is as tabulated in Table 5. From the recorded t-test value, it is demonstrated that the CSA-LSSVM, ABC-LSSVM, and DE-LSSVM are a good rival to the FA-LSSVM, except for MFO-LSSVM.

TABLE V
T-TEST

Methods	**T-test value**
FA-LSSVM – MFO-LSSVM	2.802E-05
FA-LSSVM – CSA-LSSVM	0.2991
FA-LSSVM – ABC-LSSVM	0.0636
FA-LSSVM – DE-LSSVM	0.0165

The illustration of the performance of the identified hybrid algorithms is depicted in Figure 7. From the figure, it is illustrated that most of the hybrid algorithms were incapable of producing an accurate prediction during the sudden ups and downs such as on day 380 and 410.

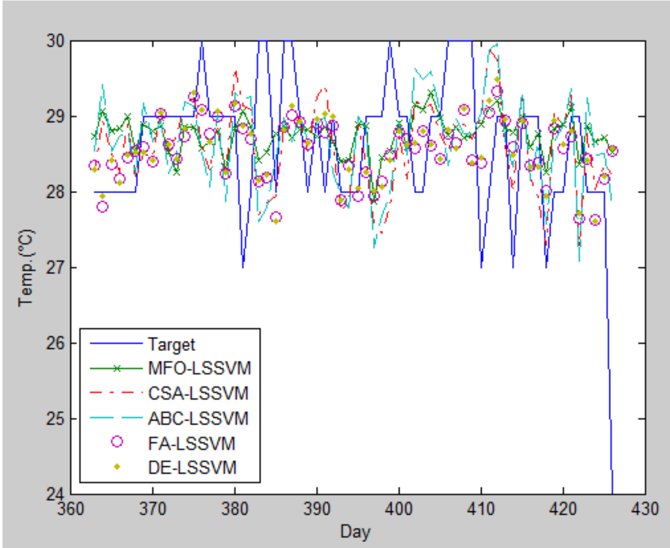

Fig. 7 MFO-LSSVM vs. CSA-LSSVM vs. ABC-LSSVM vs. FA-LSSVM vs. DE-LSSVM vs. Target

IV. CONCLUSION

This paper provides a comparative study of MFO-LSSVM, CSA-LSSVM, ABC-LSSVM, FA-LSSVM, and DE-LSSVM for weather prediction. The experimental setup was designed for 5 days ahead prediction using daily data. Findings of the study demonstrate that these hybrid algorithms are suitable for the case under study, especially the FA-LSSVM. By achieving the smallest error rates, the FA-LSSVM outperform the other identified algorithms. Nonetheless, the capability of MFO-LSSVM, CSA-LSSVM, ABC-LSSVM, and DE-LSSVM could not be underestimated for being the great rival. For further work, this research could be enhanced by referring to various works in the literature which includes in [23]-[25].

ACKNOWLEDGMENT

The authors thank the Research Acculturation Grant Scheme (RAGS) for their financial support (Grant No.RDU151408).

REFERENCES

[1] Georgina Cosma, David Brown, Matthew Archer, Masood Khan, and A. G. Pockley, "A survey on computational intelligence approaches for predictive modeling in prostate cancer," *Expert Systems with Applications*, vol. 70, pp. 1-19, 2017.

[2] Monidipa Das and S. K. Ghosh, "Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques," *Pattern Recognition Letters*, 2017.

[3] D. Petkovic, "Prediction of laser welding quality by computational intelligence approaches," *Optik-International Journal for Light and Electron Optics*, vol. 140, pp. 597-600, 2017.

[4] R. L. Haupt and S. E. Haupt, *Practical Genetic Algorithms*, Second ed. New Jersey: A John Wiley & Sons, Inc., Publication, 2004.

[5] Kennedy and R. Eberhart, "Particle Swarm Optimization," in *Proceedings of the IEEE International Conference on Neural Networks (ICNN)*, Australia, 1995, pp. pp. 1942-1948.

[6] M. Dorigo and T. Stutzle, *Ant Colony Optimization*. Cambridge: MIT Press 2004.

[7] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, 2014.

[8] Mohamad Shehab, A. T. Khader, and M. A. Al-Betar, "A survey on applications and variants of the cuckoo search algorithms," *Applied Soft Computing*, 2017.

[9] Georgina Cosma, David Browna, Matthew Archera, Masood Khanb, and A. G. Pockleyc, "A survey on computational intelligence approaches for predictive modeling in prostate cancer," *Expert Systems with Applications,* vol. 70, pp. 1-19, 2017.

[10] H. Son and C. Kim, "Short-term forecasting of electricity demand for the residential sector using weather and social variables," *Resources, Conservation and Recycling,* vol. 123, pp. 200-207, 2017.

[11] M. Wang, H. Chen, B. Yang, X. Zhao, H. Lufeng, Z. Cai, *et al.*, "Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses," *Neurocomputing,* vol. 000, pp. 1-16, 2017.

[12] X. Wang, J. Wen, Y. Zhang, and Y. Wang, "Real estate price forecasting based on SVM optimized by PSO," *Optik,* vol. 125, pp. 1439-1443, 2014.

[13] D. Karaboga, "An Idea Based on Honey Bee Swarm for Numerical Optimization," Erciyes University, Technical ReportOctober, 2005 2005.

[14] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines.* Leuven, Belgium: World Scientific Publishing Co. Pte. Ltd., 2002.

[15] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics,* vol. 42, 2012.

[16] Q. Chen, Y. Wu, and X. Chen, "Research on Customers Demand Forecasting for E-business Web Site Based on LS-SVM," in *Proceedings of the International Symposium on Electronic Commerce and Security*, 2008, pp. 66-70.

[17] P. Ou and H. Wang, "Prediction of Stock Market Index Movement by Ten Data Mining Techniques," *Modern Applied Science,* vol. 3, pp. 28-42, 2009.

[18] M. Tarhouni, K. Laabidi, S. Zidi, and M. Ksouri-Lahmari, "A Nonlinear MIMO Systems Identification Based on Improved Multi-Kernel Least Squares Support Vector Machines (Improved Multi-Kernel LSSVM)," in *Proceedings of the 8th International Multi-Conference on Systems, Signals and Devices*, 2011.

[19] S. A. Bessedik and H. Hadi, "Prediction of Flashover Voltage of Insulators using Least Squares Support Vector Machine with Particle Swarm Optimization," *Electric Power Systems Research,* vol. 104, pp. 87-92, 2013.

[20] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," *Knowledge-Based Systems*, vol. 89, pp. 228-249, 2015.

[21] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms*, Second ed. United Kingdom: Luniver Press, 2010.

[22] R. Storn and K. Price, "Differential Evolution- a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, pp. 341-359, 1997.

[23] M.A. Ismail, V. Mezhuyev, K. Moorthy, S. Kasim, A.O. Ibrahim, "Optimisation of Biochemical Systems Production using Hybrid of Newton Method, Differential Evolution Algorithm and Cooperative Coevolution Algorithm", Indonesian Journal of Electrical Engineering and Computer Science, vol.8, pp. 27-35, 2017.

[24] M.A. Ismail, V. Mezhuyev, S. Deris, M.S. Mohamad, S. Kasim, R.R. Saedudin, "Multi-objective Optimization of Biochemical System Production Using an Improve Newton Competitive Differential Evolution Method", International Journal on Advanced Science, Engineering and Information Technology, vol.7, pp.1535-1542, 2017.

[25] M.A. Ismail, S. Deris, M.S. Mohamad, M. A. Isa, A. Abdullah, M. A. Remli. S. M. Mohi-Aldeen, "A Hybrid of Optimization Method for Multi-Objective Constraint Optimization of Biochemical System Production", Journal of Theoretical and Applied Information Technology, vol.81, pp. 502-513, 2015.