



# EVALUATION OF ESTIMATION PERFORMANCE FOR SOIL MOISTURE USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORK

Engin PEKEL<sup>1,\*</sup> 

<sup>1</sup> Faculty of Engineering, Department of Industrial Engineering, Hitit University, North Campus, Corum, Turkey

## ABSTRACT

Soil plays a vital role in the climate system. This paper performs a hybrid method that comprises particle swarm optimization (PSO) and artificial neural network (ANN) to estimate soil moisture (SM) by considering different parameters that include air temperature, time, relative humidity and soil temperature. Besides, this paper investigates the effects of the parameters of PSO-ANN by using from the response surface. PSO algorithm involves changing the weights of ANN. Paper chooses the coefficient of determination and mean absolute error to measure the performance of the performed hybrid PSO-ANN. The numerical results show that hybrid PSO-ANN is applied to estimate SM successfully.

**Keywords:** Artificial neural network, Estimation, Particle swarm optimization, Soil moisture

## PARÇACIK SÜRÜ OPTİMİZASYONU VE YAPAY SİNİR AĞLARINI KULLANARAK TOPRAK NEMİ İÇİN TAHMİNLEME PERFORMANSININ DEĞERLENDİRİLMESİ

### ÖZET

Toprak, iklim sisteminde hayati bir rol oynar. Bu makale, hava sıcaklığı, zaman, bağıl nem ve toprak sıcaklığını içeren farklı parametreleri göz önüne alarak toprağın nemini (TM) tahmin etmek için parçacık sürüsü optimizasyonundan (PSO) ve yapay sinir ağından (YSA) oluşan bir hibrit yöntem uygular. Ayrıca, bu makale tepki yüzeyinden yararlanarak PSO-ANN parametrelerinin etkilerini araştırmaktadır. PSO algoritması, YSA'nın ağırlıklarını değiştirme sürecinde yer almaktadır. Korelasyon katsayısı ve ortalama mutlak hata, uygulanan hibrit PSO-ANN'nin performansını ölçmek için seçilmiştir. Sayısal sonuçlar, PSO-ANN hibritinin TM'yi başarılı bir şekilde tahmin etmek için uygulandığını göstermektedir.

**Anahtar kelimeler:** Yapay sinir ağları, Tahminleme, Parçacık sürü optimizasyon, Toprak nemi

## 1. INTRODUCTION

The soil is a fundamental part of the climate system because it meets the ever-increasing demand for food and other requirements of the world population [1]. Soil moisture (SM) is the water kept in the soil and precipitation, temperature, soil characteristics, and more affect it. The sustainability and the health of the crops highly depend on SM. Therefore, one of the most significant factors of the climate system is soil moisture.

Several techniques exist to measure soil water and comprise of the field sampling, the electromagnetic measurement, and the remote sensing technology. The electromagnetic measurement calculates the value of SM by using the electromagnetic properties of SM. The remote sensing technology gathers surface information in large areas and brings forth the fundamental spatial resolution [2]. The field sampling is a drying process and finds the values of SM by calculating the difference of soil weight before and after drying. Above methods may take serious time to measure the value of SM.

Pattern recognition and spatial data processing use several machine-learning methods, including artificial neural network (ANN), support vector machine (SVM), k-nearest neighbors algorithm (k-NN), and decision tree classification. The prediction applies to ANN, SVM, and particle swarm optimization [3]. However, the researchers have not performed hybridization of particle swarm optimization (PSO) and ANN in estimating of SM by considering the papers.

\* Sorumlu yazar / Corresponding author, e-posta / e-mail: pekerc@hotmail.com, enginpekel@hitit.edu.tr  
Geliş / Recieved: 19.02.2019 Kabul / Accepted: 05.12.2019 doi: 10.28948/ngumuh.529418

E. Pekel

The paper researches the literature by considering “soil moisture” and “particle swarm optimization” keywords and finds 37 published papers in total. Thirty-seven papers consist of thirty-two articles and five proceedings papers. None of the article papers deals with the estimation of SM by considering hybridization of PSO-ANN. Two papers [4,5] only perform the PSO algorithm to investigate the parameters of SM.

The main contributions of the paper are as follows:

- The paper illustrates that the application of PSO-ANN to estimate SM presents a high coefficient of determination ( $R^2$ ) value and low mean absolute error (MAE).
- The performed method provides an accurate estimation in a short time.

The paper organizes the remainder of sections as follows. Section 2 presents the structure and implementation of the performed method that includes PSO, ANN, and PSO-ANN. Section 3 presents the dataset of the case study and explains the input and output parameters. Section 4 provides the numerical results of the proposed algorithm by considering the different parameter settings. Section 5 presents the concluding remarks and directions for future research.

## 2. METHODOLOGY

This section presents the methodology of PSO, ANN, hybrid PSO-ANN algorithms and its tuning parameters.

### 2.1. PSO Algorithm

PSO is one of the population-based stochastic optimization technique developed by the researchers [6]. PSO algorithm performs the behaviors of bird flocking. PSO algorithm aims to have all the particles locate the optima. It consists of a collection of particles that move around the search space affected by their own best past location and the best past location of the whole swarm.

Algorithm 1 illustrates the pseudo-code of the PSO algorithm. PSO algorithm begins by generating the initial particles and assigns the initial velocities to the initial particles. In each iteration, the two best values of the fitness function update each particle. One of them is the fitness value the algorithm has achieved so far, and the algorithm stores this value as the best solution. The other is the best value that any particle in the population reaches.

#### Algorithm 1 Pseudo code of the PSO algorithm

```

Initialize each particle
Population ←  $\Theta$ 
Output P_g_best ←  $\Theta$ 
For (k=1 to Population_size)
  P_velocity ← Random_velocity()
  P_position ← Random_position (Population_size)
  P_p_best ← P_position
  Evaluation
  P_g_best ← P_p_best
End
While(termination_criteria)
  For each particle
    P_velocity ← update_velocity(P_velocity, P_g_best, P_p_best)
    P_position ← update_position(P_position, P_velocity)
    Evaluation
    P_p_best ← P_position
    Evaluation
    P_g_best ← P_p_best
  End
End
Return(P_g_best)

```

**EVALUATION OF ESTIMATION PERFORMANCE FOR SOIL MOISTURE USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORK**

$v_i(t + 1)$  is the new velocity for the  $i$ . particle.  $C_1$  and  $C_2$  parameters are weighting coefficients for the local and global best positions, respectively.  $p_i(t)$  is the  $i$ . particle's position at iteration  $t$ .  $p_i^{best}$  is the  $i$ . particle's best-known position and  $p_{gbest}$  is the best position known to the swarm. The rand variable generates random values between 0 and 1.

$$v_i(t + 1) = v_i(t) + \left( C_1 * rand * \left( p_i^{best} - p_i(t) \right) \right) + \left( C_2 * rand * \left( p_{gbest} - p_i(t) \right) \right) \tag{1}$$

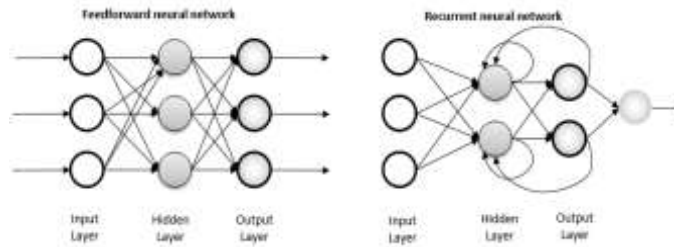
Equation (1) shows the calculation of updating the velocity of the particle.

$$p_i(t + 1) = p_i(t) + v_i(t + 1) \tag{2}$$

Equation (2) shows the calculation of updating the position of the particle. Particle positions are used as weights between neurons and inputs in PSO-ANN.

**2.2. ANN**

ANN is a novel computer architecture that relates to traditional computers by imitating biological neural networks. It provides to perform simple computational operations such as fundamental logic elements to solve with mathematical vagueness, complex, nonlinear, and stochastic problems. The fundamental logic of ANN is the weighted directed graphs where directed edges are connections between input and output. Besides, artificial neurons are nodes. There are many types of ANN model such as dynamic, static, and memory ANN. However, this paper presents several ANN models that the researchers mainly perform. The researcher divide network architecture into two groups: feed-forward network and feedback networks. Figure 1 shows the feed-forward and recurrent neural network architectures.



**Figure 1.** ANN structures

**2.3. PSO-ANN**

The ANN is a novel computer architecture that relates to traditional computers by imitating biological neural networks. It allows using simple computational operations such as fundamental logic elements to deal with complex, mathematically vague, and nonlinear or stochastic problems. Recently, the researchers have made significant efforts to improve the result of the fitness functions. One of the significant efforts is the hybridization of ANN with another meta-heuristic method.

This research paper hybridizes PSO and ANN to improve the estimation quality of SM. The main part of PSO-ANN algorithm is the application method that calculates the network connection weights. Algorithm 2 shows the pseudo-code of the hybridized PSO-ANN.

**Algorithm 2** Pseudo code of the hybridized PSO-ANN

```

Initialize initial weights ( $w(t = 0)$ )
 $t \leftarrow 0$ ;
Initialize each particle
Population  $\leftarrow \Theta$ 
Output  $P_{g\_best} \leftarrow \Theta$ 
For ( $k=1$  to Population_size)
   $P\_velocity \leftarrow \text{Random\_velocity}()$ 
   $P\_position \leftarrow \text{Random\_position}(\text{Population\_size})$ 
   $P\_p\_best \leftarrow P\_position$ 
  Evaluation
  
```

```

P_g_best←P_p_best
End
While(termination_criteria)
  For each particle
    P_velocity←update_velocity(P_velocity, P_g_best, P_p_best)
    P_position←update_position(P_position, P_velocity)
    Evaluation
    P_p_best←P_position
    Evaluation
    P_g_best←P_p_best
    Update weights (P_position)
  End
End
t←t+1;

```

PSO-ANN method uses Logarithmic-Sigmoidal transfer function in the hidden and output layers.

### 3. CASE STUDY

The University of Toronto publishes the environmental datasets to the public. This paper uses the dataset taken from the website of the University of Toronto Mississauga campus. The data comprises three different soil areas. The pond, field, and forest data generate the areas of the data. This paper uses forest dataset to estimate SM and examine the estimation results by performing different decision tree parameters. Data loggers equipped with sensors gather data hourly at these sites and it observes the inputs including time (day), soil temperature (degrees Celsius), air temperature (degrees Celsius), and relative humidity (%) and the output defined as SM ( $m^3/m^3$ ) [7]. This research paper divides the dataset into two parts for training the proposed algorithm and testing its performance. The initial dataset consists of 300 instances that include four inputs and one output. Training and testing stage keep 80% and 20% of the dataset at random, respectively.

### 4. EXPERIMENTAL RESULTS

This research paper experiments PSO-ANN to estimate SM by satisfying the best fitness performance. This section provides all characteristics of all features and numerical results of the estimation experiment. A computer that has a 32-bit Windows 7 operating system, 2.9 GHz processor, and 4-GB memory has run the method of PSO-ANN. The performed method has been implemented in the Matlab. The total time that consists of training and testing stages to estimate SM takes almost 30 seconds. Stopping criterion is the iteration number, and it is set up to 100 because increasing the number of iterations does not lead to better results.

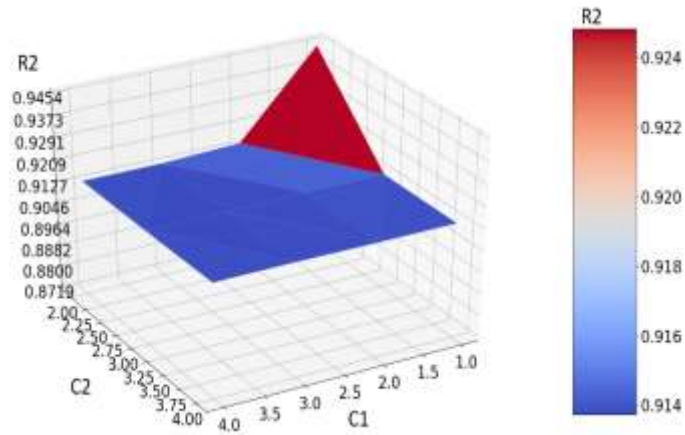
#### 4.1. Training Results

This section presents the numerical results of Field soil area regarding  $R^2$  and MAE fitness criteria. Although  $R^2$  offers some useful information regarding the model, the interpretation should not be based only on the value of  $R^2$ . High  $R^2$  value sometimes does not provide good estimation results, therefore, this paper uses  $R^2$  and MAE to evaluate the performance of the PSO-ANN. C1 and C2 parameters are the learning parameters that are described in section 3. HN is the hidden neuron number. All combinations ( $4 \times 3 \times 19 = 228$ ) are experimented by considering C1 (from 1 to 4), C2 (from 2 to 4) and HN (from 2 to 20) parameters.

##### 4.1.1. Numerical results of $R^2$ as fitness criterion

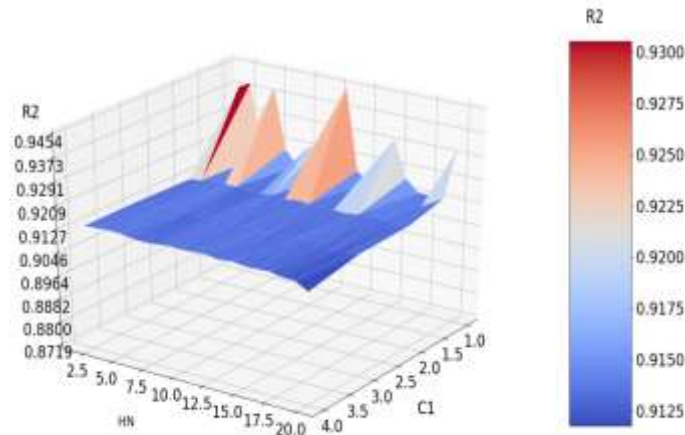
$R^2$  is a measure used to estimate future outcomes based on the known-values. It aims to fit the model that predicts how well-observed outcomes.  $R^2$  can take any rates between 0 to 1. Numerical results provide  $R^2$ -values by considering three different parameters of PSO-ANN. The paper examines C1-C2, C1-HN, and C2-HN pair-wise to investigate the effects of the parameters because the interaction of parameters may lead to severe changes in the value of fitness criterion.

*EVALUATION OF ESTIMATION PERFORMANCE FOR SOIL MOISTURE USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORK*



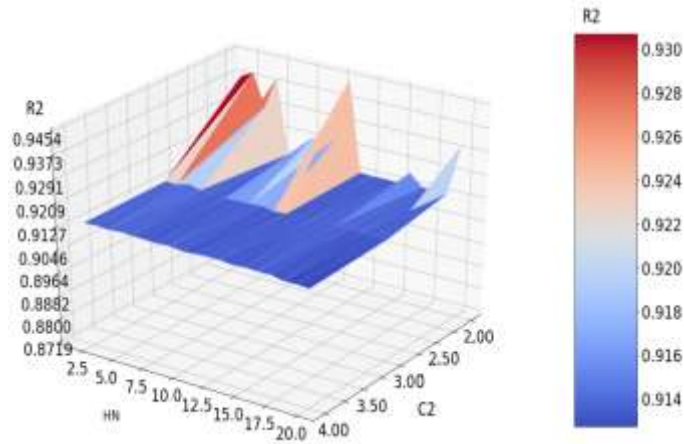
**Figure 2.** Mean value of  $R^2$  regarding  $C1$  and  $C2$  parameters

Figure 2 shows the mean value of  $R^2$  regarding  $C1$  and  $C2$  parameters by using the response surface plot. While  $C1$  and  $C2$  parameters increase, the value of  $R^2$  generally shows a downward inclination. The values of  $C1$  and  $C2$  equal to 1 and 2, respectively. These values provide the highest  $R^2$ -value by considering the best ten mean values of  $R^2$ .



**Figure 3.** Mean value of  $R^2$  regarding  $C1$  and  $HN$  parameters

Figure 3 shows the mean value of  $R^2$  for  $C1$  and  $HN$  parameters by using the response surface plot. While the value of  $HN$  changes, the value of  $R^2$  displays an uncertain pattern that provides different values. The values of  $HN$  change between 2 and 20 by considering the best ten mean values of  $R^2$ . However, eleven is the best to estimate the values of  $SM$  in the training stage because it provides the highest  $R^2$ -value.



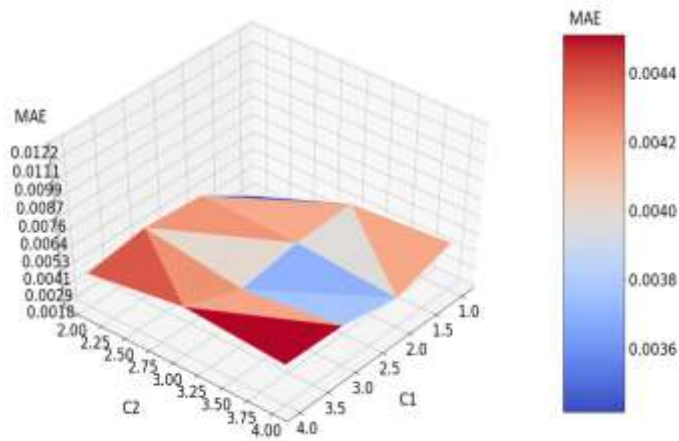
**Figure 4.** Mean value of  $R^2$  regarding C2 and HN parameters

Figure 4 shows the mean value of  $R^2$  for C2 and HN parameters by using the response surface plot. While the values of HN and C2 parameters change, the value of  $R^2$  displays an uncertain pattern. Figure 4 is similar to Figure 3, and it supports the unpredictable model considering different HN parameters. While HN parameter is greater than 11, most of the  $R^2$ -values declines.

The values of C1, C2, and HN that offer the highest  $R^2$  value equal to 1, 2, and 11, by considering Figure 2, Figure 3, and Figure 4, respectively. Implementation of PSO-ANN provides the highest  $R^2$  value (0.94) by considering three parameters.

**4.1.2. MAE fitness criterion**

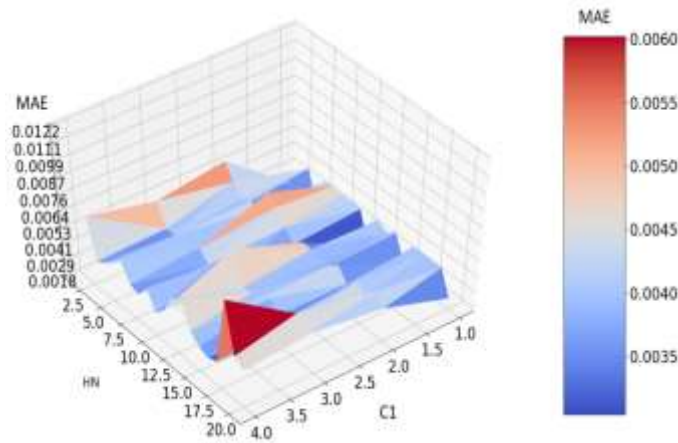
MAE is one of the most common criteria to calculate accuracy between the observed and the estimated variables. MAE calculates the absolute difference between two continuous (the observed and the estimated) variables. The calculated MAE equals to a positive value, and PSO-ANN aims to minimize this performance criterion. This section provides the numerical results of MAE by considering three different parameters of PSO-ANN.



**Figure 5.** Mean value of MAE regarding C1 and C2 parameters

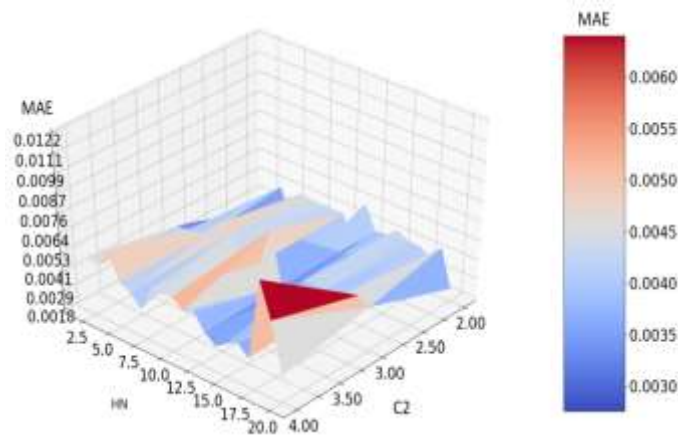
Figure 5 shows the mean value of MAE on C1 and C2 parameters by using the response surface plot. As C1 and C2 parameters increase, the mean value of MAE shows an upward and downward inclination. Downward and increase of C1 and C1 parameters do not offer a stable model. Considering the best ten mean values of MAE, the values of C1 and C2 that equal to 1 and 2, respectively offer the least MAE.

*EVALUATION OF ESTIMATION PERFORMANCE FOR SOIL MOISTURE USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORK*



**Figure 6.** Mean value of MAE regarding C1 and HN parameters

Figure 6 shows the mean value of MAE regarding C1 and HN parameters by using the response surface plot. As the value of HN changes, the value of MAE displays an unstable pattern that comprises different values. The values of HN change between 2 and 20 by considering the best ten mean values of MAE. However, eleven is the best to predict the values of SM in the training stage by providing the least MAE.



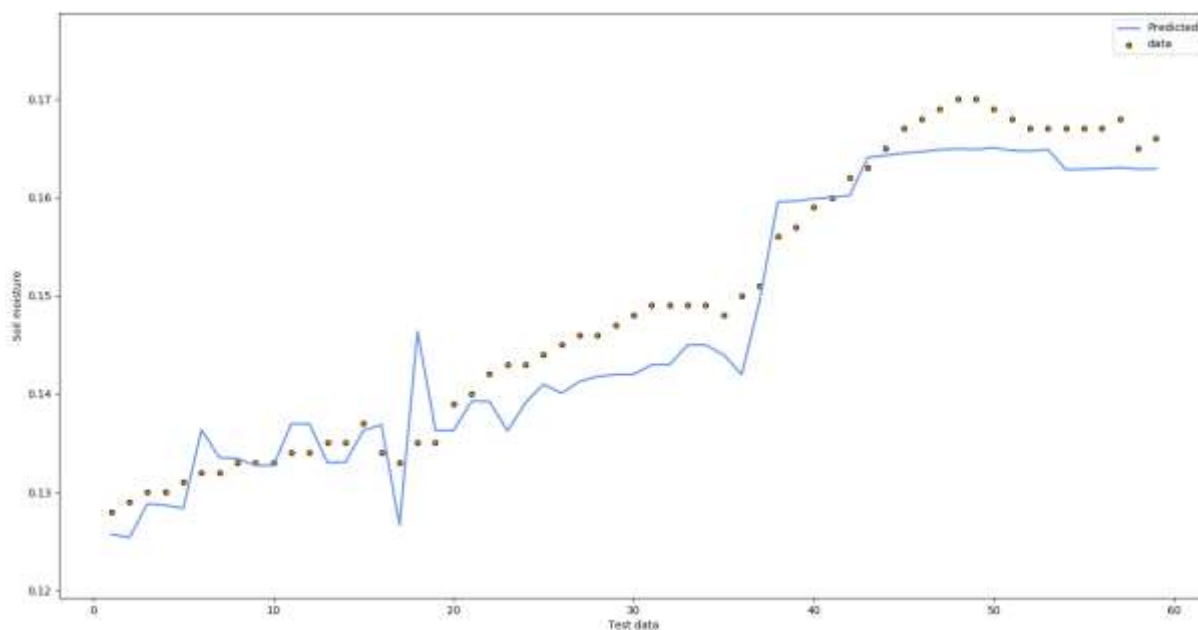
**Figure 7.** Mean value of MAE regarding C2 and HN parameters

Figure 7 shows the mean value of MAE for C2 and HN parameters by using the response surface plot. While the values of HN and C2 parameters change, the value of MAE does not display a stable pattern. Downward and increase of HN and C1 parameters do not provide valuable information to evaluate the change.

Considering Figure 5, Figure 6, and Figure 7, the values C1, C2, and HN that offer the least MAE equal to 1, 2 and 11, respectively. Implementation of PSO-ANN finds the least MAE value (0.0017) regarding three parameters. The best values of parameters are similar by considering MAE and  $R^2$  fitness criteria. Two different fitness criteria (MAE and  $R^2$ ) support each other, and this illustrates that found numerical results are more reasonable.

**4.2. Testing Results**

The proposed algorithm PSO-ANN predicts SM using the dataset, kept for the testing stage, by implementing the optimal parameters of the algorithm. The optimal parameters, found in the training stage, also provide the best values in the testing stage. The proposed algorithm finds the values of  $R^2$  and MAE as 0.923 and 0.0033, respectively.



**Figure 8.** The plot of observed data and the predicted value

Figure 8 shows the graphical illustration of observed data and predicted value. The predicted values display several spikes. Considering the overall predicted values of SM, the optimal parameters of PSO-ANN offer a decent pattern to estimate the values of SM in the testing stage.

## 5. CONCLUSION

This paper addresses the estimation of SM by performing PSO-ANN that the article papers do not previously address. The following part addresses the main contributions. The first contribution is the application of PSO-ANN to estimate SM presents a high  $R^2$  value and low MAE value. It indicates that the performed method satisfies two different fitness criteria by offering the best values. The second contribution is the accurate prediction in a short time. Besides, another important contribution is the estimation of SM by performing PSO-ANN by considering the different parameters.

The numerical results show that the performed PSO-ANN may handle the data by estimating SM with satisfying fitness criteria. The obtained results of MAE and illustrate that the optimal parameters of PSO-ANN offer a decent estimation performance in the training and the testing stage. The proposed algorithm finds C1, C2, and HN as 1, 2, and 11, respectively. The results of the training stage provide the values of  $R^2$  and MAE as 0.945, 0.0017, respectively, and the results of the testing stage present the values of  $R^2$  and MAE as 0.923 and 0.0033, respectively.

The researcher may conduct further researches to improve the fitness criteria of the estimation of SM. For instance, the researcher may perform supervised learning methods, including kernel ridge regression, deep learning, SVM, neural network, and other hybrid methods, to estimate SM values and investigate the effects of parameters of supervised learning methods.

## REFERENCES

- [1] G. Shukla, R. D. Garg, H. S. Srivastava and P. K. Garg, "An effective implementation and assessment of a random forest classifier as a soil spatial predictive model" *International Journal of Remote Sensing*, vol. 39, no. 8, Jan, pp. 2637-2669, 2018.
- [2] Y. Qu, X. Qian, H. Song, Y. Xing, Z. Li and J. Tan, "Soil Moisture Investigation Utilizing Machine Learning Approach Based Experimental Data and Landsat5-TM Images: A Case Study in the Mega City Beijing" *Water*, vol. 10, Apr, pp. 423, 2018.



*EVALUATION OF ESTIMATION PERFORMANCE FOR SOIL MOISTURE USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORK*

- [3] V. Moosavi, A. Talebi, M. H. Mokhtari and M. R. Hadian, “Estimation of spatially enhanced soil moisture combining remote sensing and artificial intelligence approaches” *International journal of remote sensing*, vol. 37, no. 23, Oct, pp. 5605-5631, 2016.
- [4] D. Kundu, R. W. Vervoort and F. F. van Ogtrop, “The value of remotely sensed surface soil moisture for model calibration using SWAT” *Hydrological Processes*, vol. 31, no. 15, May, pp. 2764-2780, 2017.
- [5] Q. Yang, H. Zuo and W. Li, “Land Surface Model and Particle Swarm Optimization Algorithm Based on the Model-Optimization Method for Improving Soil Moisture Simulation in a Semi-Arid Region” *Plos One*, vol. 11, no. 3, Mar, 2016.
- [6] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory” *Proceedings of the Sixth International Symposium Micro Machine and Human Science*, 1995, pp. 39-43.
- [7] University of Toronto, “Environmental datasets” Available: <https://www.utm.utoronto.ca/geography/resources/environmental-datasets>. [Accessed Feb. 04, 2019]

