Coupling Normalization with Moving Window in Backpropagation Neural Network (BNN) for Passive Microwave Soil Moisture Retrieval

Soo See Chai¹ · Kok Luong Goh² · Yee Hui Robin Chang³ · Kwan Yong Sim⁴

Received: 3 June 2021 / Accepted: 6 October 2021
© The Author(s) 2021

Abstract

A common practice to capture the non-stationary characteristics of the time series data in Artificial Neural Network (ANN) is by randomly dividing the whole set of available data into training, validation and testing, i.e. the data in validation and testing are represented in the training data. Consequently, the usability of the developed model on data not represented by the training data used during the network model development process is always doubtful. In this work, we present a back-propagation neural network (BNN) model trained using one-day history data to predict soil moisture data at 1 km resolution for two future dates. Specifically, high soil moisture values were observed in the training data while the testing data were characterized by drier conditions due to minimal or no rainfall. Our model uses separate mean and standard deviation statistics values from the training and testing data, respectively, to the z-normalized data. With data pre-processed using this method, the BNN model next uses a moving window of size 4 km × 4 km to capture the spatial variability of the soil moisture throughout the 40 km × 40 km study area. The coupling of the normalization and moving window method managed to achieve average soil moisture with Root Mean Square (RMSE) of 3.67% and correlation coefficient, R^2 of 0.89. By only using the suggested normalization without the moving window method, the BNN model managed to achieve an average RMSE of barely 5.82% with $R^2 = 0.83$. When comparing with the normal practice of using the same mean and standard deviation statistics of the training data in the testing data, the retrieval accuracy of the BNN model deteriorates to 8.86% with $R^2 = 0.32$. The experiment results demonstrated that the proposed coupling method performed better in terms of both RMSE and R^2 for soil moisture retrieval.

Keywords Backpropagation · Neural network · Normalization · Passive microwave · Soil moisture

✉ Soo See Chai
sschai@unimas.my
Kok Luong Goh
klgoh@ppks.edu.my
Yee Hui Robin Chang
robincyh@uitm.edu.my
Kwan Yong Sim
ksim@swinburne.edu.my

¹ Faculty of Computer Science and Information Technology, University of Malaysia Sarawak (UNIMAS), 94300 Kota Samarahan, Sarawak, Malaysia

² Faculty of Science and Technology, i-CATS University College, 93350 Kuching, Sarawak, Malaysia

³ Faculty of Applied Sciences, Universiti Teknologi MARA, Cawangan Sarawak, 94300 Samarahan, Sarawak, Malaysia

⁴ Faculty of Engineering, Computing and Science, Swinburne University of Technology Sarawak Campus, Jalan Simpang Tiga, 93350 Kuching, Sarawak, Malaysia

1 Introduction

Surface soil moisture content is a critical parameter for various disciplines. Therefore, accurate large-scale estimation of the value is vital in improving hydrology and climatic modeling and prediction [1, 2]. Due to the high spatial-temporal variability of the surface soil moisture [3], monitoring surface soil moisture over big areas is a tough task as accurate soil moisture values will need to be obtained at fine resolution. Low frequencies microwave remote sensing has shown to be the best performing approach among the different satellite remote sensing methods for large areas integration information on soil moisture [4–6]. Among the active and passive microwave data, soil moisture is successfully retrieved from passive microwave sensors [7]. The most particular benefit of passive microwave sensors is that, on the received signal, in the lack of substantial vegetation cover, soil moisture has the dominant effects [8].

Artificial Neural Network (ANN) has proved to be the method that could determine the distribution of soil moisture most accurately [9]. Before applying the ANN model, some major issues that need to be considered include: network topology, learning parameters and data normalization methods [10–12]. In applying the ANN model for soil moisture retrieval, the available data is normally decomposed randomly into training, validation and testing set in the pre-processing stage. A few examples of such work could be found in [13–18]. The general practice is to decompose the total data into 60%, 20% and 20% for training, validation and testing. That is to say, the trained network will only need to interpolate in between the training points [19] as the validation and testing data were represented in the training data. Prediction of the neural network is only possible if the future event does not extend the input range used to train the network model. When the trained model is tested on data that is outside the range of the training points, i.e. extrapolation, the network will not be able to perform well. Yet, in real-life applications, especially in research related to weather forecasting, unpredictable conditions do happen and the use of the neural network models on these unpredictable conditions is always skeptical. To handle the different conditions, the neural network models will need to be trained with all possible data so that the trained models could produce decent accuracy. Consequently, the training data set must be large enough to include all possible data ranges. This would mean, long history data would be needed. However, this could be a problem if the target area is new and long history data on this area is not available. We termed this as an “out-of-range” problem, which would be explored using our proposed model. In our work, our model predicts the extreme soil moisture values which were not covered within the training data.

For any application, after the data is obtained, the data will need to go through pre-processing steps to acquire quality data from the raw data. Data normalization aims to eliminate the magnitudes discrepancy between the dimensional data. Moreover, data normalization also can help to avoid huge errors in prediction due to magnitudes discrepancy between the input and output data [20]. Min–max, z-normalization and decimal scaling normalization are some of the commonly used data normalization methods [21, 22]. Data normalization has been proved to enhance neural network training speed and performance [23–25]. According to [26], to normalize the data in the neural network model, the validation and testing data must be standardized using the statistics calculated from the training data. To the best of our knowledge, no research has been carried out on the effects of using separate statistics in the training and testing stages, especially in the extrapolation situation. The procedure to normalize the data in the pre-processing step in the neural network model is not reported in most of the study.

In our previous study [9], the use of the “windowing” method had been demonstrated to capture the spatial variability of the soil moisture. Different windows sizes ranging from 2 km × 2 km, 3 km × 3 km, 4 km × 4 km, 5 km × 5 km, 6 km × 6 km, 7 km × 7 km and 8 km × 8 km were used in the BNN model to retrieve the soil moisture at 1 km resolution over the target area of 40 km × 40 km. From this particular study, it was found that, the retrieval accuracy deteriorated with the increase of the window size. The optimum window size found was 4 km × 4 km. In this previous work, we managed to achieve an average accuracy of soil moisture retrieval with RMSE = 7.5% and $R^2 = 0.8$. Therefore, in this current study, using the same field experiment data, the window size of 4 km × 4 km will be used.

Therefore, this paper intends to propose an approach to solve the “out-of-range” problem described earlier by coupling the normalization with moving window in Backpropagation Neural Network (BNN). In particular, the contributions of this paper include:

- (i) Exploration of Backpropagation Neural Network (BNN) model to forecast out-of-range soil moisture values, i.e. the data used in the testing are outside the range of the training data set. Specifically, the capability of the BNN in dealing with extrapolation problem is explored.
- (ii) The effects of using the input and output normalization factors of each data set on the soil moisture prediction accuracy vs. the use of the same normalization factor calculated from the training data in the testing data.
- (iii) Coupling of the suggested normalization method with moving window technique in BNN for soil moisture retrieval.

The remaining of this paper is organized as below. The related work on BNN for soil moisture retrieval is first presented. This is followed by the data and methodology. The results obtained and the conclusion highlighting the contributions of this paper are presented at the end of the paper.

2 Literature Review

2.1 Research Background

The effectiveness of ANN models in solving remote sensing applications is well demonstrated. Since 1988, there has been a rise of ANN usage to classify remotely sensed images [27]. Among the various ANN models, the multi-layer feed-forward network with error back-propagation learning algorithm, i.e. Backpropagation neural network (BNN), is the most popular ANN model in use [28–30]. This is mostly

owing to the BNN model's excellent nonlinear mapping ability, which is critical in soil moisture time series prediction [31]. Figure 1 shows a general BNN model. The following is an overview of some of the recent research studies on soil moisture retrieval using microwave data, with an emphasis on the BNN model.

In the domain of soil moisture retrieval using microwave data, the Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS) satellites have a mission objective of $\pm 0.04 \text{ m}^3/\text{m}^3$ or 4% accuracy of the volumetric soil moisture within the first five centimeters of the soil [32]. The use of BNN for soil moisture retrieval using microwave data has shown great successes in fulfilling this mission. In a recent study on sensitivity analysis with SMAP soil moisture estimates on a monthly basis over Mainland China [33], a back-propagation artificial neural network was used to build a model utilizing data spanning more than two years. The developed model was utilized to create long-term and continuous soil moisture maps across Mainland China. The back-propagation neural network structure employed in the study was: input layer comprises of seven nodes (surface reflectivity, Normalized Difference Vegetation Index (NDVI), vegetation water content, elevation, slope, precipitation, roughness). The anticipated soil moisture values were represented by a single node in the output layer. There were two hidden levels with a total of eight nodes. The activation function selected was the hyperbolic tangent. A regression layer with no activation function was the last layer. The maximum number of training sessions was set to 6000, the error measure was set to Root Mean Square Error (RMSE), the error threshold was set to 0.001, and the learning rate was set to 0.05. Prior to training, the dataset was normalized to values between 0 and 1. The predicted soil moisture showed a strong and positive linear relationship with SMAP in-situ

soil moisture with correlation coefficient, R^2 of 0.840 and RMSE of $0.068 \text{ m}^3/\text{m}^3$.

An evaluation study of soil moisture estimation using BNN and two empirical models, namely, modified Dubois model and multiple linear relationship analysis, was conducted using backscattering coefficients extracted from Sentinel-1 images and in-situ soil surface parameter measurements [34]. It was concluded that the BNN model produced a better correlation coefficient with in-situ soil moisture values than the empirical models. The input layer of the BNN model in their study comprised of local incidence, vertical polarization, vertical and horizontal polarization, and surface roughness, with one hidden layer of 20 neurons and the soil moisture as the output. The data was normalized to the value between 0 and 1. The average RMSE value obtained using the BNN model was 4.98% with R^2 of 0.49, whereas for the multiple regression analysis, the RMSE value was 4.52% with R^2 of 0.31. The modified Dubois model only managed to produce an average RMSE of 8.65% with R^2 of 0.26.

In another study by [35], BNN was developed to retrieve soil moisture from Advanced Microwave Scanning Radiometer 2 (AMSR2). Their research work could be divided into two parts. In the first part, the BNN was developed to reproduce the soil moisture values of the SMAP satellite. The one hidden layer BNN model was trained with data from 2015 to 2017 to predict the soil moisture values in the year 2018. They reported a good accuracy with a correlation coefficient of 0.74, Root Mean Square Error (RMSE) of $0.033 \text{ m}^3/\text{m}^3$. However, when the trained BNN model was used to develop a global soil moisture data for the year 2012 to 2018 and compared the retrieved values with in-situ Soil Climate Analysis Network (SCAN) sites, they only managed to obtain retrieval results with RMSE of $0.113 \text{ m}^3/\text{m}^3$ and correlation coefficient of 0.44.

Ge, Hang, Liu and Liu [36] attempted to compare deep convolutional neural network and BNN with multi-source input including L-band Soil Moisture and Ocean Salinity (SMOS) brightness temperature (TB), C-band Advanced Scatterometer (ASCAT) backscattering coefficients, Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI) and soil temperature. Their experiment was divided into learning and testing phases. In the learning phase, the network models were trained using the simulated soil moisture values of the European Center for Medium-range Weather Forecasts Re-Analysis Interim (ERA-interim) for data from 1 January to 31 December 2015. In the testing phase, in-situ soil moisture measurements from Soil Climate Analysis Network (SCAN) from 1 January to 31 December 2016 was used. The data of the ascending and descending orbits were processed separately. From their study, they found that, both BNN and deep convolutional neural network methods are effective in

Fig. 1 General schematic diagram of a Backpropagation Neural Network model

estimating soil moisture from satellite observations and the performance of the deep convolutional neural network is slightly better than BNN. Nevertheless, the convolutional neural network has a more complex structure and it usually costs much more time than BNN to estimate soil moisture values. In their work, they reported that, the data for BNN was normalized using the Min–Max normalization method to scale all data into the range [0,1]. The average values of RMSE and correlation coefficient obtained for convolutional neural network (BNN) were 0.564 (0.531) and 0.099 (0.140) for ascending orbit and 0.530 (0.517) and 0.102 (0.156) for descending orbit.

Yao et al. [37] had developed a one hidden layer BNN model to produce a long-term soil moisture time series data for the year 2003 to 2015. The BNN model was trained using data from July 2010 to June 2011, and the entire year of 2013 with the reflectivities (Rs) of the C/X/Ku/Ka/Q bands, MVI (Microwave Vegetation Index) derived from Advanced Microwave Scanning Radiometer-Earth Observing System/ Advanced Microwave Scanning Radiometer 2 (AMSR-E/AMSR2) brightness temperature (TB data), as input. The predicted soil moisture values from the BNN model was compared with the Soil Moisture and Ocean Salinity (SMOS) Level 3 soil moisture and the results of RMSE = 0.55 m³/m³ and correlation coefficient of 0.67 were obtained. When evaluating the BNN retrieved soil moisture with in-situ soil moisture from the SCAN (Soil Climate Analysis Network) sites, the accuracy acquired was RMSE = 0.084 m³/m³ with a correlation coefficient of 0.52.

According to the review, the use of BNN for soil moisture retrieval has demonstrated good agreement with both satellite soil moisture products and in-situ measurements. However, it was evident that the retrieval accuracy would fall short of the mission's ± 0.04 m³/m³ or 4% goal. Furthermore, based on the review, the testing data was mostly partitioned from the dataset available, or a long history of soil moisture data would be required to develop a robust prediction model. In our work, we would concentrate our efforts on the challenge of "out-of-range" prediction, which involves predicting data from the same geographic region but with extreme soil moisture levels not covered by the training data. In particular, our approach to solving this problem emphasizes the use of normalization and the moving window method.

2.2 Normalization: Input and Output Vectors

The type of problem will determine the number of input and output nodes of the network model, as well as the model selected for this particular problem [38]. The input vectors for a neural network model will be either raw data, pre-processed data or a set of parameters. The benefits of data normalization included increase the speed of calculation, improve performance and results, as well as enhance the

reliability of the network model [12, 39–42]. Depending on the activation function used in the BNN model, the input vector can be normalized between 0 and 1 (for standard sigmoid function) or – 1 to 1 (for hyperbolic tangent).

A neural network model is a series of neurons. Within each neuron is a set of input, weight and a bias value. Before learning initially begins, the neural network model will randomize both the weight and bias values. The weight can be deemed as the strength of the connections while bias represents the distance between the predictions and intended value. In principle, the objective of normalization is to rescale the input vector to adjust the weight and bias corresponding to the relevant vector for obtaining the same output features [12]. Enabling the input and output values to be at the same interval via normalization is essential to determine the very high or very low values in the data [43]. These very high or very low values may result in wrong outputs of the neural network model. Therefore, normalization of input and output data has sound effects on the consistency of the results obtained.

2.3 Normalization Method: Statistical or Z-normalization

This method normalizes each input feature vector using the mean (μ) and standard deviation (σ) of each feature across the training data. The normalization is done using the following equation:

$$x'_i = \frac{x_i - \mu_i}{\sigma_i} \quad (1)$$

where: x'_i is the new x value,

x_i is the current x value,

μ_i is the mean of the current feature vector, and

σ_i is the standard deviation of the current feature vector.

This produces data where each feature has a zero mean and one-unit variance [12] and efficiently reduces the effect of an outlier in the data. Before training is commenced, all the feature vectors in the data set will undergo this normalization technique, resulting in a new training set. Once the mean and standard deviation values are computed for each feature over a set of training data, these values must be retained and used as the weight in the final system design [12].

3 Data

This study uses the National Airborne Field Experiment 2005 (NAFE'05) data in the Goulburn catchment, south-eastern Australia, in November 2005. The area monitored was a square of approximately 40 km × 40 km area. The

campaign included extensive airborne passive microwave observations together with spatially distributed and in-situ ground monitoring of soil moisture. This aircraft was equipped with the Polarimetric L-band Multibeam Radiometer (PLMR) and thermal imager. The PLMR obtains data of both H- and V-polarized brightness temperatures (T_b). The soil moisture (SM) data for the 40 km $\times$ 40 km study area at 1 km nominal resolution was derived using the L-band Microwave Emission from the Biosphere (L-MEB) model. The ground monitoring includes soil moisture profile monitoring stations, supplementary monitoring stations, spatial soil moisture mapping and supporting data. Stevens water HydraProbe® sensors were inserted vertically in the ground to measure soil moisture in the top 5 cm. In addition, supplementary sensors were deployed to provide information on the near-surface soil temperature profile. The main purpose of these continuous ground observations is to relate the air-to-ground measurements throughout the day. Figure 2 shows the soil moisture distribution of the L-MEB model product at 1 km resolution. Further detail of the data used could be found in [44].

4 Methodology

4.1 Data Partitioning

The input of the BNN model consists of T_bH , T_bV and T_s . For the training of the BNN model, the data from 7th November 2005 was used. The trained model was tested using data from 14th November 2005 and 21st November 2005. During training, the data on 7th November 2005 was

Fig. 2 L-MEB retrieved soil moisture at 1 km resolution on **a** 7th Nov, **b** 14th Nov and **c** 21st Nov 2005. The boundaries of the focus farms and the whole study area are shown as polygons

randomly subdivided into 90% for training, 5% for validation and another 5% for testing purpose. Figure 3 summarizes this data partitioning process. The main goal of this data arrangement method is to allow the trained network to predict soil moisture that is outside of the training data range. The focus of this study would be on training using the wet conditions on November 7th, 2005 to predict soil moisture values on the other future dates of November 14th and 21st, 2005, when the weather was getting drier. The trained network's ability to predict values outside of the training data range is an important feature for the network model's transferability and reusability.

4.2 Statistics Description of Training and Testing Data

The statistical properties of the H- and V-polarized brightness temperature (T_bH , T_bV), surface temperature (T_s) and soil moisture (SM) at 1 km resolution on 7th, 14th and 21st November 2005 are listed in Table 1. For each date, each feature has 1600 data points at 1 km resolution (area size 40 km $\times$ 40 km). Table 1 summarizes the fact that the study area was becoming drier. The retrieved soil moisture

Fig. 3 The data used for training, validation and testing of to form the trained ANN model. The trained ANN model is tested on the data on 14th and 21st Nov 2005 which was not used in forming the trained ANN model

Table 1 Statistics properties of the regional L-MEB soil moisture product (v/v) at 1 km resolution together with brightness temperature at H- and V-polarized (T_bH and T_bV)(K) and surface temperature (T_s)(K) data

Date	T_bH (K)	T_bV (K)	T_s (K)	Soil Moisture (v/v)
7th Nov 2005	241.5 $\pm$ 10.1	261.4 $\pm$ 7.8	297.1 $\pm$ 3.1	0.39 $\pm$ 0.12
14th Nov 2005	266.0 $\pm$ 6.5	279.3 $\pm$ 5.4	298.2 $\pm$ 6.4	0.17 $\pm$ 0.10
21st Nov 2005	271.3 $\pm$ 3.9	282.6 $\pm$ 3.1	313.8 $\pm$ 3.9	0.16 $\pm$ 0.08

The mean and standard deviation values of each data type for each date are listed

(SM) exhibits intriguing spatio-temporal dynamics, reflecting the rainfall regime experienced by the area during the field experiments. The high soil moisture values on November 7, 2005 were caused by heavy rainstorms that passed through the study area. The drier conditions on 14th and 21st November 2005 were due to little or no rainfall experienced during the field experiment. For the drier conditions on both 14th November 2005 and 21 November 2005, the TbH, TbV and Ts were showing increasing values. In addition to this, another interesting observation is that the standard deviation values of the H- and V-polarized brightness temperatures (TbH and TbV) were high for the date 7th November 2005, decreasing to the smaller on November 14, 2005, and the smallest on November 21, 2005. This trend is in accordance with the soil moisture (SM) values. It should be noted that, the data on 21st November 2005 was less spread out as compared to the data on 14th November 2005 while the data on 7th November 2005 was the most spread out among these three days. This is demonstrated by higher standard deviation values for soil moisture, TbH, and TbV.

4.3 BNN Architecture

Aside from the input and output layers, the number of hidden layers and neurons must be determined. While ANNs with two hidden layers can represent any function, there is currently no theoretical reason to use neural networks with more than two hidden layers. As a result, the ANN architecture under consideration here employs one and two hidden layers. Using too few or too many hidden neurons may jeopardize any application. Too few hidden neurons will result in underfitting, which means that the ANN will be unable to detect complex signals within the data accurately. Using too many hidden neurons, on the other hand, will result in overfitting, in which the neural network has so much information processing capacity that the limited amount of information contained in the training set is insufficient to train all the neurons in the hidden layers. Furthermore, if too many hidden neurons are used, the training time will increase. In this study, the trial-and-error method is employed to determine the optimal number of hidden layers and hidden neurons.

In our earlier study [9], the optimal window size was reported to be 4 km × 4 km. To identify the best architecture, the BNN model with varying numbers of hidden neurons for both one and two hidden layers was examined using this window size. It should be noted that only training data was used during this process.

Figure 4 and Table 2 show that a single hidden layer BNN with 2, 4, and 6 neurons had a lower Root Mean Square Error (RMSE) than two hidden layers with 2:2, 4:4, and 5:5 neurons. The RMSE increased when 8 hidden neurons were used in a single hidden layer, but began to decrease when 10 and 20 hidden neurons were used. The RMSE increased with

Fig. 4 Graph showing the trends of RMSE and R^2 obtained using one and two hidden layers in BNN with varying hidden neurons

Table 2 RMSE and R^2 obtained using single and two hidden layers of BNN with varying hidden neurons

	Hidden Neurons	RMSE	R^2
Single Layer	2	5.66	0.54
	4	5.35	0.66
	6	4.87	0.56
	8	5.31	0.43
	10	4.56	0.69
	20	2.86	0.88
	50	3.57	0.75
Two Layers	100	3.08	0.84
	2:2	6.24	0.29
	4:4	5.28	0.53
	5:5	5.60	0.74
	10:10	4.95	0.55
	20:20	3.63	0.79
	50:50	3.50	0.78
	100:100	3.25	0.82

hidden neuron 50 and decreased with hidden neuron 100 in a single layer. The pattern observed for two layers was nearly identical. This indicated that there was no compelling reason to employ two hidden layers. The first minimum achieved before the first raised of RMSE was used as the criterion for selecting the optimal number of hidden neurons. As a result, the optimal number of hidden neurons is 6 in one single layer.

4.4 Performance Evaluation Metrics

To evaluate the performance of the developed BNN model, the coefficient of determination (R^2) and Root Mean Square Error (RMSE) were used. The R^2 examines the goodness of fit and measures how well the BNN model predicts the

soil moisture values. This metric is expressed as a number between 0.0 and 1.0, with 1.0 indicating a perfect fit and consequently a highly dependable model for future projections, and 0.0 indicating that the model fails to predict the data effectively. Another commonly used performance evaluation metric for soil moisture retrieval is the RMSE. This metric, on the other hand, assesses the prediction's accuracy. In this study, the RMSE would measure the soil moisture prediction accuracy using the proposed model. The Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS) satellites have a mission objective of $0.04 \text{ m}^3/\text{m}^3$ or 4% accuracy of volumetric soil moisture within the first five centimeters of the soil in the area of soil moisture retrieval utilizing microwave data. We would evaluate our model on the basis of this mission.

4.5 Method Design

The method design is divided into three parts. In the first part, the mean and standard deviation values for each input and output feature are computed from the training data set. The calculated normalization factors are then used to normalize both validation and testing data. In the second part of the method design, the mean and standard deviation of the input and output for each feature are computed separately in the training, validation and testing stages. During training, the mean and standard deviation of each input and output feature is used to normalize the data, while in the validation and testing stage, the mean and standard deviation of each feature for each data set is used (Fig. 5). To deal with the high spatial and temporal variability characteristics of the soil moisture over the $40 \text{ km} \times 40 \text{ km}$ field, the “moving window” of the size $4 \text{ km} \times 4 \text{ km}$ suggested in our previous study [9] was developed and applied. The “moving window” slides from the first pixel until the last pixel on the image of

Fig. 5 The procedure of normalization used in part 2 of the experiment. The normalization factors used are of the feature of each data set in the testing of the trained BNN model

the $40 \text{ km} \times 40 \text{ km}$ target field. During each slide, the mean and standard deviation of the input and output vectors were calculated at $4 \text{ km} \times 4 \text{ km}$ resolution. Using the normalization factors calculated within this $4 \text{ km} \times 4 \text{ km}$ resolution, soil moisture prediction at 1 km resolution was done for the whole $40 \text{ km} \times 40 \text{ km}$ target field. This is shown in Fig. 6.

5 Results and Discussions

A BNN with the architecture of 3 inputs (TbH, TbV and Ts), 6 hidden neurons and 1 output (SM) was trained using the data on the date 7th November 2005. The 1 km L-MEB soil moisture product was used as the target values of the BNN model. The trained network will be tested using the data on 14th and 21st November 2005. The z-normalization was used to normalize the data. The results obtained for the three parts of the experiments are presented in the following sections. We would evaluate our model based on $0.04 \text{ m}^3/\text{m}^3$ or 4% accuracy of volumetric soil moisture within the first five centimeters of the soil in the area of soil moisture retrieval utilizing microwave data, which is the mission objective of the Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS) satellites.

5.1 Using Normalization Factors of the Training Data Set

In this part of the experiment, the mean μ and standard deviation σ for each feature of the input and output data

Fig. 6 The start of the $4 \text{ km} \times 4 \text{ km}$ “moving window” and its moving direction

from the training data are calculated and used in both the training and testing of the BNN model. Table 3 shows the results obtained. It is noted that, the correlation coefficient, R^2 obtained was very low, showing that, the relationships among the actual and forecast soil moisture (SM) values are not strong. Moreover, the soil moisture retrieval accuracy obtained was at an average of $RMSE=8.85\%$ and $R^2=0.32$. It should be emphasized that, the R^2 obtained for testing on 21st Nov 2005 was extremely low (0.05) which indicates that, the relationships between the actual and predicted soil moisture were very minimal. As a comparison, the results obtained during training on 7th Nov 2005 were also included in Table 3. When compared to the findings produced on November 21, 2005, the retrieval accuracy on the training data on November 7, 2005 was closer to the retrieval accuracy on November 14, 2005. This is mostly owing to the fact that the weather on the 21st of November 2005 was drier than on the 14th of November 2005, which was one week away from date of the training data. It should also be emphasized that, as previously stated, heavy rainstorms occurred across the research region on November 7, 2005, resulting in high moisture levels.

5.2 Using the Normalization Factor of Each Data Set

During this experiment, the normalization factors from the training data set were not retained. It should be noted that the same BNN model developed in the previous section was employed in this section. The normalization factors are calculated separately for each set of data during training, validation and testing. This is summarized in Fig. 5. In other words, during training, the normalization factors used were calculated from the training data set, while for the validation and testing, the normalization factors were calculated from each of the datasets. Table 4 shows the results obtained. The RMSE obtained for both dates are lower as compared to the first part of the experiment and the R^2 of the predicted and actual SM was strong. Average accuracy obtained was 5.82% with $R^2=0.83$. An important observation from this experiment is that, the R^2 values obtained were good for both dates, indicating that the relationships between the actual and predicted soil moisture values were strong.

Table 3 RMSE (%) and R^2 obtained for testing data on 14th and 21st Nov 2005 using the standard normalization procedure

	Date	RMSE (%)	R^2
Training	7th Nov 2005	8.09	0.69
Testing	14th Nov 2005	9.97	0.58
	21st Nov 2005	7.74	0.05

For comparison, the results produced for training data on November 7, 2005 are given

Table 4 RMSE (%) and R^2 obtained when the trained network is tested on data on 14th and 21st November 2005 using the normalization factor from each date

Date	RMSE (%)	R^2
14th Nov 2005	6.37	0.86
21st Nov 2005	5.26	0.80

5.3 Using the Normalization Factor of Each Data Set Coupling with a Moving Window

Although the normalization factor for each data set manages to lower the RMSE value with strong R^2 as shown in Table 4, the retrieval accuracy obtained is still considered poor as compared to the SMAP and SMOS mission of 4%. According to our previous study [9], the retrieval accuracy is poor as the soil moisture is highly spatial and temporal variable over the $40\text{ km} \times 40\text{ km}$ field. To solve this problem, a “moving window” of $4\text{ km} \times 4\text{ km}$ is used. The mean and standard deviation of the input and output vectors within this “moving window” is calculated. In practical, these values could be estimated with satellite data. For simulation purpose, in this study, an upscaling to $4\text{ km} \times 4\text{ km}$ resolution using the $1\text{ km} \times 1\text{ km}$ resolution data is done. The mean and standard deviation of the $4\text{ km} \times 4\text{ km}$ resolution is used to normalize the input and output data. The “moving window” slides from the top left of the image until the end of the image during the prediction process and this is shown in Fig. 6.

Table 5 clearly shows that, using the same BNN model built, the RMSE obtained is lower than the targeted 4% ($RMSE=3.67\%$) with high R^2 showing a very strong relationship ($R^2=0.89$) between the predicted and actual soil moisture value. In addition, the prediction on each date was stable with both RMSE achieved less than 4% and R^2 more than 0.85. More fluctuation could be seen in the prediction results in the previous two experiments.

6 Conclusions and Future Directions

The use of data-driven methods, especially ANN models has proven to be effective for inversion of soil moisture estimation, thanks to the capabilities of ANN in mapping the

Table 5 RMSE (%) and R^2 obtained using normalization factors of each data set and “moving window” method

Date	RMSE (%)	R^2
14th Nov 2005	3.92	0.91
21st Nov 2005	3.41	0.86

nonlinear and complex functions of soil moisture. The BNN is the most commonly used ANN model for soil moisture retrieval. The common method in fitting the network model is to divide the available data set into training, validation and testing subsets. This way of data splitting ensures that the validation and testing data are represented in the training data. With this, the developed network model would be useful if and only if the model was trained with all unforeseen circumstances. However, in the atmospheric disciplines, unpredicted circumstances do happen. In addition, for a totally new target, long history data might not be available. In situations like this, the model developed with the whole available data splitting method while fitting the model would fail to produce decent results. This is clearly shown in this paper whereby the retrieval accuracy of soil moisture was low when the BNN model is trained on a wet condition and tested on two future dates whereby the conditions were getting drier and drier.

Before developing the network model, one of the major steps during the pre-processing stage is data normalization. In common practice, the statistics used to normalize the data should be obtained from the training data set. Such mechanism assumes data stationary and uniformity and poses problems, particularly in extrapolation problem as in the scenarios dealing within this study. In our study, the use of separate statistics to normalize the training, validation and testing data in the BNN model was demonstrated to cope with such a situation. As a comparison, the soil moisture prediction was also done with BNN model which used the same statistics calculated from the training data to normalize the data in the validation and testing stages. It was found that, when using separate statistics to normalize the data in training, validation and testing stages, the average accuracy improves from $RMSE = 8.89\%$ ($R^2 = 0.32$) to $RMSE = 5.82\%$ ($R^2 = 0.83$).

Owing to the soil moisture high spatial–temporal variability characteristics, we adopted the “moving window” method to compute the mean and standard deviation statistics to normalize the data. Coupling the separate normalization statistics with the “moving window” method, the trained BNN model manages to capture the spatial and temporal variability of the soil moisture values. With this suggested method, the trained BNN model managed to achieve an average accuracy of $RMSE = 3.67\%$ ($R^2 = 0.89$), which is an increase of 58% from without utilizing the “moving window” technique. Another important observation from this experiment was that, the $RMSE$ and R^2 values were very consistent as compared to the results obtained from the previous two experiments. This indicated that, the suggested coupling method managed to produce stable prediction results.

Estimating soil moisture is critical for detecting early-stage water deficit conditions and developing drought situations. The two severe circumstances have a big impact on

crop production uncertainty and food security. Accurate soil moisture prediction would help with agricultural insurance, policymaking and decision-making, and crop planning, particularly in dry and semi-arid regions of the world. Surface soil moisture referred to the water content approximately 5–15 cm of the top soil layer. This thin layer of soil plays an important role in hydrology, biochemistry, ecology, agriculture, and other Earth processes. Machine learning techniques have yield promising results for soil moisture prediction. Nevertheless, there are numerous unique instances where machine learning may be used to improve remotely sensed soil moisture products. In our work, we propose a possible solution to the problem of not having enough long-term data to cover all possible scenarios, which could be a possible condition for a new geographic area for weather prediction problems. The developed model would be able to handle such unforeseen circumstances using the proposed method while maintaining the robustness of the trained model. We must admit that the study's limitation is that it used only one wet condition to train the network and tested it with two drier conditions. As a result, we propose that a further verification on more contrasting circumstances for training and testing be performed to validate the proposed methodology in this study.

Acknowledgements The authors would like to thank Faculty of Computer Science and Information Technology, UNIMAS for providing the facilities for the research.

Author Contributions All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Funding Open Access funding provided by Universiti Malaysia Sarawak. The study was funded under University Malaysia Sarawak (UNIMAS) Cross Disciplinary Grant (Grant No. F08/CDRG/1832/2019).

Availability of Data and Material The datasets for the current study are available from <https://www.nafe.monash.edu/> on reasonable request.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Priddy, K.L., Keller, P.E.: Artificial neural networks: an introduction, vol. 68. SPIE press (2005)
- Brocca, L., Melone, F., Moramarco, T., Morbidelli, R.: Spatial-temporal variability of soil moisture and its estimation across scales. *Water Resour. Res.* (2010). <https://doi.org/10.1029/2009WR008016>
- Song, Q. and N. Kasabov, WDN-RBF: weighted data normalization for radial basic function type neural networks. 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541). Vol. 3. IEEE, (2004)
- Njoku, E.G., Entekhabi, D.: Passive microwave remote sensing of soil moisture. *J. Hydrol.* **184**(1–2), 101–129 (1996)
- Wagner, W., Naeimi, V., Scipal, K., de Jeu, R., Martínez-Fernández, J.: Soil moisture from operational meteorological satellites. *Hydrogeol. J.* **15**(1), 121–131 (2007)
- Zhang, Q. and S. Sun, Weighted Data Normalization Based on Eigenvalues for Artificial Neural Network Classification. arXiv preprint arXiv:1712.08885, 2017.
- Jin, J., Li, M., Jin, L.: Data normalization to accelerate training for linear neural net to predict tropical cyclone tracks. *Math. Probl. Eng.* **2015**, 1–8 (2015)
- Singh, G., Srivastava, H.S., Mesapam, S., Patel, P.: Passive microwave remote sensing of soil moisture: a step-by-step detailed methodology using AMSR-E data over indian sub-continent. *Int. J. Adv. Remote Sens. GIS* **4**(1), 1045–1063 (2015)
- Chai, S.-S., Walker, J., Makarynskyy, O., Kuhn, M., Veenendaal, B., West, G.: Use of soil moisture variability in artificial neural network retrieval of soil moisture. *Remote Sens.* **2**(1), 166 (2010)
- Kim, D.: Normalization methods for input and output vectors in backpropagation neural networks. *Int. J. Comput. Math.* **71**(2), 161–171 (1999)
- Tai, Q.-Y., Shin, K.-S.: GA-based normalization approach in backpropagation neural network for bankruptcy prediction modeling. *J. Intell. Inf. Syst.* **16**(3), 1–14 (2010)
- Jayalakshmi, T., Santhakumaran, A.: Statistical normalization and back propagation for classification. *Int. J. Comput. Theor. Eng.* **3**(1), 1793–8201 (2011)
- Chai, S.-S., Veenendaal, B., West, G., Walker, J.P.: Backpropagation neural network for soil moisture retrieval using NAFE'05 data: a comparison of different training algorithms. *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **37**(B4), 1345–1349 (2008)
- Elshorbagy, A., Parasuraman, K.: On the relevance of using artificial neural networks for estimating soil moisture content. *J. Hydrol.* **362**(1), 1–18 (2008)
- Notarnicola, C., Angiulli, M., Posa, F.: Soil moisture retrieval from remotely sensed data: neural network approach versus Bayesian method. *IEEE Trans. Geosci. Remote Sens.* **46**(2), 547–557 (2008)
- Kornelsen, K.C., Coulbaly, P.: Root-zone soil moisture estimation using data-driven methods. *Water Resour. Res.* **50**(4), 2946–2962 (2014)
- Hassan-Esfahani, L., Torres-Rua, A., Jensen, A., McKee, M.: Assessment of surface soil moisture using high-resolution multi-spectral imagery and artificial neural networks. *Remote Sens.* **7**(3), 2627–2646 (2015)
- Rodríguez-Fernández, N.J., Aires, F., Richaume, P., Kerr, Y.H., Prigent, C., Kolassa, J., et al.: Soil moisture retrieval using neural networks: application to SMOS. *IEEE Trans. Geosci. Remote Sens.* **53**(11), 5991–6007 (2015)
- Jia, Y., Ma, J.: What can machine learning do for seismic data processing? An interpolation application. *Geophysics* **82**(3), V163–V177 (2017)
- Zhang, N., R. Li, and N. Huang. Resource Demand Predicting Based on Cluster Analysis in Cloud Computing Networks. in *Software Engineering and Information Technology: Proceedings of the 2015 International Conference on Software Engineering and Information Technology (SEIT2015)*. 2015. World Scientific.
- Ogasawara, E., L.C. Martinez, D. De Oliveira, G. Zimbrão, G.L. Pappa, and M. Mattoso. Adaptive normalization: A novel data normalization approach for non-stationary time series. in *Neural Networks (IJCNN), The 2010 International Joint Conference on*. 2010. IEEE
- Gökhan, A., Güzeller, C.O., Eser, M.T.: The effect of the normalization method used in different sample sizes on the success of artificial neural network model. *Int. J. Assess. Tools Educ.* **6**(2), 170–192 (2019)
- Puheim, M. and L. Madarász. Normalization of inputs and outputs of neural network based robotic arm controller in role of inverse kinematic model. in *2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*. 2014. IEEE
- Kwon, B.-S., R.-J. Park, S.-W. Jo, and K.-B. Song. Analysis of short-term load forecasting using artificial neural network algorithm according to normalization and selection of input data on weekdays. In *2018 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*. 2018. IEEE
- Khond, S.V.: Effect of data normalization on accuracy and error of fault classification for an electrical distribution system. *Smart Sci.* **8**(3), 117–124 (2020)
- Sarle, W.S. Neural network FAQ, part 2 of 7: Learning, May 2002. Available from: <ftp://ftp.sas.com/pub/neural/FAQ2.html>. Accessed 5 Apr 2021
- Sadeghi, M., Asanjan, A.A., Faridzad, M., Nguyen, P., Hsu, K., Sorooshian, S., et al.: PERSIANN-CNN: precipitation estimation from remotely sensed information using artificial neural networks—convolutional neural networks. *J. Hydrometeorol.* **20**(12), 2273–2289 (2019)
- Wang, W., Zhao, S., Jiao, L., Taylor, M., Zhang, B., Xu, G., et al.: Estimation of PM2.5 concentrations in China using a spatial back propagation neural network. *Sci. Rep.* **9**(1), 1–10 (2019)
- Wang, L., Wang, P., Liang, S., Zhu, Y., Khan, J., Fang, S.: Monitoring maize growth on the North China Plain using a hybrid genetic algorithm-based back-propagation neural network model. *Comput. Electron. Agric.* **170**, 105238 (2020)
- Sharifi, A.: Yield prediction with machine learning algorithms and satellite images. *J. Sci. Food Agric.* **101**(3), 891–896 (2021)
- Yang, X., Jia, S., Zhang, C.: A prediction method for soil moisture time series. In: *International Conference on Machine Learning for Cyber Security*. Springer (2020)
- Ye, N., Walker, J.P., Wu, X., de Jeu, R., Gao, Y., Jackson, T.J., et al.: The Soil Moisture Active Passive Experiments: Validation of the SMAP Products in Australia. *IEEE Transact Geosci Remot Sens* **59**(4), 2922–2939 (2020)
- Yang, T., Wan, W., Sun, Z., Liu, B., Li, S., Chen, X.: Comprehensive evaluation of using TechDemoSat-1 and CYGNSS data to estimate soil moisture over mainland China. *Remote Sens.* **12**(11), 1699 (2020)
- Li, Y., Yan, S., Chen, N., Gong, J.: Performance evaluation of a neural network model and two empirical models for estimating soil moisture based on sentinel-1 SAR data. *Prog. Electromagn. Res.* **105**, 85–99 (2020)
- Yao, P., H. Lu, S. Yue, F. Yang, H. Lyu, K. Yang, et al. Estimating Surface Soil Moisture from AMSR2 Tb with Artificial Neural Network Method and SMAP Products. in *IGARSS 2019–2019*

- IEEE International Geoscience and Remote Sensing Symposium. 2019. IEEE.
36. Ge, L., Hang, R., Liu, Y., Liu, Q.: Comparing the performance of neural network and deep convolutional neural network in estimating soil moisture from satellite observations. *Remote Sens.* **10**(9), 1327 (2018)
 37. Yao, P., Shi, J., Zhao, T., Lu, H., Al-Yaari, A.: Rebuilding long time series global soil moisture products using the neural network adopting the microwave vegetation index. *Remote Sens.* **9**(1), 35 (2017)
 38. Zhang, Q., Sun, S.: Weighted data normalization based on eigenvalues for artificial neural network classification. In: *International Conference on Neural Information Processing*. Springer (2009)
 39. Neaupane, K.M., Achet, S.H.: Use of backpropagation neural network for landslide monitoring: a case study in the higher Himalaya. *Eng. Geol.* **74**(3–4), 213–226 (2004)
 40. Sola, J., Sevilla, J.: Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Trans. Nucl. Sci.* **44**(3), 1464–1468 (1997)
 41. Eesa, A.S., Arabo, W.K.: A normalization methods for backpropagation: a comparative study. *Sci. J. Univ. Zakho* **5**(4), 319–323 (2017)
 42. Alickovic, E., Subasi, A.: Normalized neural networks for breast cancer classification. In: *International Conference on Medical and Biological Engineering*. Springer (2019)
 43. Aksu, G., Güzeller, C.O., Eser, M.T.: The effect of the normalization method used in different sample sizes on the success of artificial neural network model. *Int. J. Assess. Tools Educ.* **6**(2), 170–192 (2019)
 44. Rüdiger, C., Hancock, G., Hemakumara, H.M., Jacobs, B., Kalma, J.D., Martinez, C., et al.: Goulburn River experimental catchment data set. *Water Resour. Res.* (2007). <https://doi.org/10.1029/2006WR005837>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.