

Received December 13, 2020, accepted December 24, 2020, date of publication December 30, 2020, date of current version January 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3048028

Soil Temperature Prediction Using Convolutional Neural Network Based on Ensemble Empirical Mode Decomposition

HUIBOWEN HAO^{ID}, FANHUA YU, AND QINGLIANG LI^{ID}

School of Computer Science and Technology, Changchun Normal University, Changchun 130032, China

Corresponding author: Qingliang Li (liqingliang@ccsfu.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant U1811464a; in part by the Jilin Province Science and Technology Development Program under Grant 2020C019-3, Grant 2019C039-1, and Grant 2019C054-8; in part by the Jilin Province Science and Technology Developing Scheme under Grant 20180201086SF; in part by the Provincial Science and Technology Innovation Special Fund Project of Jilin Province under Grant 61604019; in part by the Jilin Provincial Education Department under Grant JJKH20190499KJ; and in part by the Natural Science Foundation of Changchun Normal University under Grant 2019006.

ABSTRACT Soil temperature plays an important role in agriculture, industry and other fields. Accurate soil temperature prediction can help improve productivity and avoid risks in many fields. At present, many machine learning methods have been applied to soil temperature prediction such as support vector regression (SVR), artificial neural network (ANN), long short-term memory neural network (LSTM) and others. In this article, we propose a machine learning model called convolutional neural network based on ensemble empirical mode decomposition (EEMD-CNN) to predict soil temperature. In this model, ensemble empirical mode decomposition (EEMD) is used to decompose original soil temperature series into several intrinsic mode functions (IMFs). After decomposition, the original series are combined with IMFs to get new two-dimension input data as the input of the convolutional neural network (CNN). By comparing the results which is predicted by the trained model with the original soil temperature series and other four models of persistence forecast (PF), backpropagation neural network (BPNN), LSTM and EEMD-LSTM these, the result shows that EEMD-CNN has the better performance than other four models. EEMD-CNN shows good performance not only on predicting next day's soil temperature but also on predicting several days delay's temperature also has good performance. It is concluded that the proposed EEMD-CNN model in this study is a suitable tool for soil temperature prediction.

INDEX TERMS Soil temperature prediction, machine learning, ensemble empirical mode decomposition (EEMD), convolutional neural network (CNN).

I. INTRODUCTION

Soil temperature acts as important roles in physical, chemical, and biological processes. It is relevant to soil science, agriculture, hydrology, meteorology, environmental science and many other research fields [1], [2]. Soil temperature affects the balance of heat energy between the atmosphere and the land surface [3]. Also, it changes several pivotal processes in soil such as soil ventilation, evaporation and transpiration, root development and plant growth, and the activity of microorganisms. Soil temperature performs an essential role in agriculture, because seeds germinate at the right temperature and in a certain temperature range. Thus, with higher

soil temperature in the range, the growth rate of crops will be elevated [4]. In addition, in some other fields such as water resources and hydrologic engineering, soil temperature is an important factor. Moreover in the field of atmospheric science, the change of soil temperature have obvious effects on the decomposition of the organic matter, which leads to an increment of the carbon dioxide (CO₂) to the atmosphere [5]. Therefore, a model which can forecast soil temperature accurately is widely needed. Soil temperature forecasting at different depths has different contributions. For example, Zeynoddin *et al.* [6] pointed out that the soil temperature of 5 cm depth is significant for seeds germinations. And the soil temperature of 10 cm depth influences the activities of most vital ecosystems. For the soil temperature of 20 cm depth, it has more important impact on root absorption activities.

The associate editor coordinating the review of this manuscript and approving it for publication was Yongming Li^{ID}.

Common methods for obtaining soil temperature include the following two ways: directly measure the soil temperature and indirectly get soil temperature through prediction model. The direct measure method is achieved by inserting high-precision temperature sensors in the soil. Unlike direct measurement, indirect method can be achieved by many kinds of prediction models, which were trained by the historical values at any depth and time. Comparing with the direct measure method, predictive models could get the future soil temperature in advance even if there is a partial error with the real value at that time. If the error of the prediction model is small enough and the model has high accuracy, the predicted temperature can be used as a guide. In recent years, many kinds of prediction models aiming on accuracy soil temperature forecasting have been proposed. These approaches can be divided as follow: statistical methods and machines learning models. When using prediction models to forecast soil temperature, different models will use different environmental factors, such as air temperature, solar radiation, relative humidity, atmospheric pressure, wind speed and other surface characteristics [7], [8]. Statistical models, also named Box-Jenkins models [9], are widely used for time series and predicting subsequent series. Therefore, statistical models use historical time-based temperature series to predict future soil temperature. The commonly used statistical models are the Auto-Regressive Moving Average (ARMA)[10] model and the Auto-Regressive Integrated Moving Average (ARIMA) model. These models are often used to predict time-based series. But temperature series is not stationary enough for these models. Even though the data has been manipulated by calculus of differences, these two models are not suitable for long-term soil temperature forecasting.

Recent years, the most widely used methods to forecast soil temperature are machine learning, including support vector regression (SVR) [11]–[13], random forest (RF) [14], [15], artificial neural networks (ANN) [16]–[20] and other machine learning models [21]–[24]. SVR, developed based on the support vector machine (SVM), is the regression version of SVMs, which is one of the most important applications of function approximation and is widely used for temperature data prediction. RF is a non-linear statistical ensemble method which was proposed by Breiman. It used “bagging” to ensemble a collection of decision trees with controlled variance to serve as a prediction model. Gradient boosting decision tree (GBDT) [25] is a similar method to RF, but GBDT model is fit on the residual of the former trees to reduce the biases instead of reduce the variances. And eXtreme gradient boosting (XGBoost) [26] is another tree boosting system which performs better than GBDT and had already succeeded in many prediction cases. Among these models, different structures of ANN has been widely used in the prediction of soil temperature such as backpropagation neural networks (BPNN) [27], multilayer perceptron (MLP), radial basis neural networks (RBNN), multilayer perceptron (MLP), extreme learning machine (ELM) and other neural networks. Feng *et al.* [28] used GRNN, BPNN, RF and ELM

to predict the half-hourly soil temperature at four depths in China. The recurrent neural network (RNN) [29] is neural sequence model and performs well in processing time-based series data [30], [31]. But RNN could not solve the problem of missing information over long time series. As an improvement of RNN, long short-term memory neural network (LSTM) [32]–[36] was proposed. LSTM is improved and applied for the time series forecasting due to its ability of learning long time series without vanishing gradient problem, so LSTM model is used for time series forecasting, especially suitable for longer series predictions. After these models, some ensemble approaches were proposed and succeeded in handling time series prediction [37]–[39]. Zhang *et al.* [40] proposed a model which added ensemble empirical mode composition (EEMD) to LSTM, called EEMD-LSTM. With the help of EEMD method, EEMD-LSTM performed better than LSTM on forecasting soil temperature.

In this article, we propose a model using EEMD and convolutional neural networks (CNN), called EEMD-CNN, to forecast soil temperature using time-based temperature series to solve the problem of accurately predicting soil temperature in some cases when other environmental factors are incomplete or missing. CNN is usually used for image classification, speech recognition natural language processing and other tasks, the input of CNN usually is two-dimensional or three-dimensional features. Therefore, the one-dimensional series is not suitable as input to CNN. EEMD-CNN just solved the problem of input dimension and convolutional kernel extracts more subtle changes. To prove the performance of the EEMD-CNN, BPNN, LSTM and EEMD-LSTM, these machine learning models which are also appropriate for the time-based temperature series, are applied and compared with our model. Comparing EEMD-CNN with BPNN and LSTM, the input data of EEMD-CNN not only includes the original soil temperature series, but also contains the IMFs which are processed by EEMD. And the difference between EEMD-CNN and EEMD-LSTM is the structure of the neural network.

Besides these models, we also used persistence forecast (PF) [41], a simplest possible persistence forecast method, which can be comprehended as today’s soil temperature is tomorrow’s forecast. For a more comprehensive comparison of our model’s performance, we used data from there areas of each has three soil temperature series from different depths. For each depth, we used models to forecast the soil temperature of one day delay, three days delay and five days delay. Mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE) and r-squared (r^2 -score) were used as evaluation metrics to evaluate the performance of four models. The objectives of our study are: (1) to use the evaluation metrics to comprehensively evaluate the prediction results of the four models (2) to compare the performances between these models and (3) to demonstrate the superiority of the EEMD-CNN when using time-based soil temperature series to predict the future soil temperature.

TABLE 1. Statistical properties of the meteorological information and soil temperature data at different depth for three areas.

Study Area	Depth	Maximum (°C)	Minimum(°C)	Mean(°C)	Variance(°C)
Lägern	5 CM	26.876	-1.888	9.901	36.033
	10 CM	22.193	-0.181	9.604	28.996
	30 CM	18.815	1.260	9.130	20.372
Oensingen	5 CM	27.422	-3.444	10.689	58.177
	10 CM	26.229	-2.051	10.659	51.758
	30 CM	24.735	-0.041	10.755	43.733
Fluhli	5 CM	21.822	-0.35	8.679	38.497
	10 CM	21.727	-0.044	8.792	37.467
	25 CM	20.222	0.789	8.942	32.577

The rest of the paper is organized as follows. Section 2 describes the study area, introduces the EEMD-CNN and other models that are used for comparison. The performance criteria are also introduced in this section. Section 3 presents the results and discussion. Section 4 gives the conclusion.

II. MATERIALS AND METHODS

A. STUDY AREA

The study areas of this article are Lägern (47°28'54"N, 8°23'41"E), Oensingen (47°17'N, 7°43'E) and Fluhli (46°53'N, 8°1'E) in Switzerland. The experimental data set is downloaded from FLUXNET (<https://fluxnet.fluxdata.org/>) on a daily time scale to test the models on predicting soil temperature in three areas. The information about the data is shown in Table 1.

B. CONVOLUTIONAL NEURAL NETWORK BASED ON ENSEMBLE EMPIRICAL MODE DECOMPOSITION (EEMD-CNN)

1) ENSEMBLE EMPIRICAL MODE DECOMPOSITION (EEMD)

Empirical mode decomposition (EMD) was proposed by Huang *et al.* [42], [43] to adaptively decompose a complex signal into series of intrinsic mode functions (IMFs) according to the signal characteristics. The decomposed IMFs contain local characteristic signals of different time scales of the original signals. The essence of the EMD is to identify all the intrinsic oscillatory modes contained in the signal through the characteristic time scale. EMD is self-adaptability because it is based on the local characteristics of signal sequence time scale.

To decompose the IMFs from the original signal, the process of EMD can be described as follows:

1. Find all the extreme points of the signal $x(t)$.
2. Use the cubic spline fitting to fit the envelope lines $e_{\max}(t)$ and $e_{\min}(t)$ of the upper and lower extreme points, and find the average value $m(t)$ of two envelope lines, $x(t)$ minus $m(t)$ is equal to $h(t)$, $h(t) = x(t) - m(t)$.
3. Judge whether $h(t)$ is IMF according to the preset criteria.
4. If not, then replace $x(t)$ with $h(t)$ and repeat the above steps until $h(t)$ satisfies the criterion, then $h(t)$ is the IMF $C_k(t)$.

5. Every time an IMF is obtained, it is deducted from the original signal, repeat the above steps until the last part of the signal $r_n(t)$ is just a monotone sequence or constant sequence.

In this way, after decomposition by the EMD, the original signal $x(t)$ is decomposed into several IMFs and a residual:

$$x(t) = \sum_{k=0}^N C_k(t) + r_n(t) \quad (1)$$

EMD has many advantages, but EMD has some problems and shortcomings at the same time, such as mode mixing and endpoint effects. Endpoint effects means that different ways of handling endpoint effects in the EMD decomposition process will bring different results, and mode mixing may cause serious aliasing in the time-frequency distribution leading to degrading the decomposition accuracy. On the basis of EMD, ensemble empirical mode decomposition (EEMD) [44] is proposed. EEMD takes advantage of the uniform distribution of white noise spectrum, adds white noise to the signal to be analyzed, so that the signal of different time scales can be automatically separated to the reference scale that suits them. EEMD mainly adds white noise to the signal to supplement some missing scales and has a good performance in signal decomposition. EEMD can be described as follows:

1. Add normally distributed white noise to the original signal.
2. Take the signal with white noise as a whole, and use EMD to decompose the signal into IMFs.
3. Repeat step 1 and step 2, add different normally distributed white noise sequence to the original signal each time.

4. Integrate and average the obtained IMF as the result.

Comparing with fourier transform, wavelet decomposition and other methods, EEMD is intuitive, indirect, posterior and adaptive. And EEMD is an improved algorithm of EMD, which can effectively solve the problems of mode mixing and endpoint effects existing in EMD. Due to the characteristics of zero-mean noise, the added white noise will cancel each other after multiple average calculations, so that the calculation result of the integration and averaging can be directly regarded as the result. Therefore, EEMD is chosen to decompose raw time-series data.

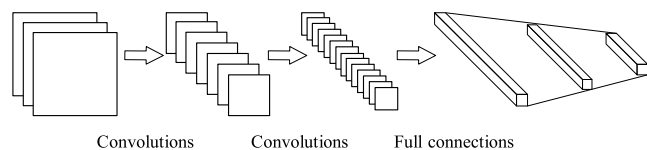


FIGURE 1. The structure of the convolutional neural network (CNN).

2) CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN [45]–[48] is a kind of feedforward neural network which contains convolution computation and is one of the representative algorithms of deep learning. Usually, the hidden layer of CNN is composed of many convolutional layers and pooling layers. Therefore, CNN can make better use of the input data of the two-dimensional or three-dimensional structure. When the two-dimensional or three-dimensional data is input into CNN model, the convolution kernel will shorten the length and width of the original input data and increase the number of channels of the data to extract the features of the original data. Then reshape the three-dimensional data into one-dimensional, and the one-dimensional data goes through the full-connect layers, we get the output of CNN. There are usually several pooling layers in the CNN model to reduce dimension, reduce computation and reduce memory consumption. However, in our model, the input data used is not complex, so the pooling layer is not used. CNN are often used for image recognition, image classification and natural language processing.

3) EEMD-CNN

In this study, EEMD-CNN is the method we proposed to predict soil temperature. Time-based temperature series is a one-dimensional data which couldn't be the input data of two-dimensional convolutional neural network. But the nonstationary and nonlinear temperature series can be decomposed by EEMD into several IMFs and residue item. The decomposed IMFs can be regarded as a supplement to the original one-dimensional data. Extend the original data with IMFs to make the input two-dimensional data have the same number of rows and columns. The workflow chart of the EEMD-CNN model is shown in Figure 2. The process of training the EEMD-CNN model are as follow:

1. Use EEMD to decompose the soil temperature series into several IMFs and residue item.
2. Combine the 7-days continuous temperature sequence with IMF1 to IMF6 of corresponding time. The obtained 2-dimensional (7*7) array as the new input data.
3. Use the next day's temperature of 7-days continuous temperature sequence as the target to train the EEMD-CNN model.
4. Input the test set into the trained model, compare the original temperature series with the predict series. Evaluate the performance of the EEMD-CNN model using several statistical evaluation metrics.

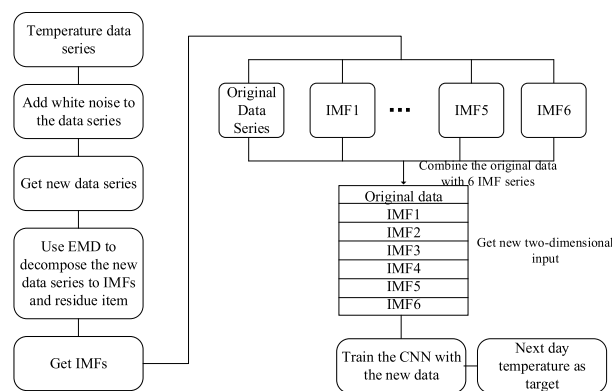


FIGURE 2. The framework of the proposed EEMD-CNN model.

C. COMPARATIVE EXPERIMENTAL METHOD

1) PERSISTENCE FORECAST (PF)

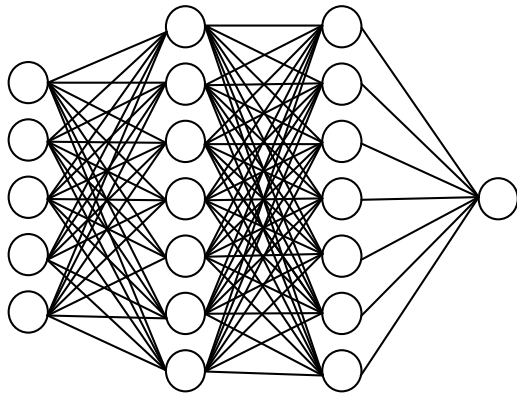
In this study, the time-based temperature series are spaced by days. PF [41] is a kind of simple forecast method which regards first day's temperature is next day's forecast. Since the temperature difference between two consecutive days is not large, it is feasible to predict next day's temperature directly from the temperature of the previous day. This is also the prediction method with the lowest calculation cost. If the prediction results that are obtained by machine learning methods perform worse than the persistence forecast's performance. It would be better to use today's temperature to predict tomorrow's instead of machine learning methods. It can be regarded as the minimum standard to evaluate the feasibility of machine learning methods. That means any soil temperature prediction model needs to beat the PF.

2) BACKPROPAGATION NEURAL NETWORK (BPNN)

Artificial neural networks (ANN) which imitates the structure of animal nervous system is one of the widely applied machine learning models. And backpropagation algorithm is the often-used algorithm to train network. The ANN trained by backpropagation algorithm is called BPNN. Usually, BPNN consists of an input layer, one or more hidden layers and an output layer. Data is input from the input layer passing through the hidden layers. Finally, we get the predicted result in the output layer. The error between the output result and the label updates the weight of each neuron according to backpropagation algorithm. In this study, BPNN utilize time-based temperature series to predict the future temperature.

3) LONG SHORT-TERM MEMORY NEURAL NETWORK BASED ON ENSEMBLE EMPIRICAL MODE DECOMPOSITION (EEMD-LSTM)

The long short-term memory neural network (LSTM) is a kind of recurrent neural network (RNN) which is successfully used on many fields such as natural language processing, speech recognition, handwriting recognition and time series prediction. In RNN, the transfer between the two adjacent cells only includes h-state (hidden state). RNN is not for long



Input layer Hidden layers Output layer

FIGURE 3. The structure of the backpropagation neural network (BPNN).

sequences because the recent data affects more on the h-state and the long-term data's impacts will be weakened. And LSTM solves this problem with input gate, output gate and forget gate. Three kinds of gates were added into LSTM cell, that makes a c-state (cell-state) run through the whole network. The c-state records a long-term memory and three kinds of gates decide what will be recorded into long-term memory and what will be delivered to next cell as the short-term memory in h-state.

Like EEMD-CNN, IMFs which are decomposed by EEMD will be combined with the original data and be used as the inputs of the EEMD-LSTM to predict the future temperature.

D. MODEL TRAINING AND TEST

In this study, nine temperature series from three areas and three depth are used as the input data to train models to predict the temperatures one day later, three days later and five days later. Each series will be split into two parts. The first part, including 80% data, is used for training models. The remaining 20% data is used as the testing set.

To evaluate the performance of EEMD-CNN and compare EEMD-CNN with other models, several statistical evaluation metrics including mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE) and r-squared (r2-score) are applied to assess model performance. The evaluation metrics are defined as follow (y_i is the true value of the testing set, \hat{y}_i is the prediction value of the model):

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \tag{3}$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \tag{4}$$

$$R2 - score = 1 - \frac{\sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2}{\sum_{i=0}^{n-1} (y_i - \bar{y}_i)^2} \tag{5}$$

III. RESULTS AND DISCUSSION

In this section, BPNN, LSTM, EEMD-LSTM, persistence forecast and the proposed EEMD-CNN models are used to

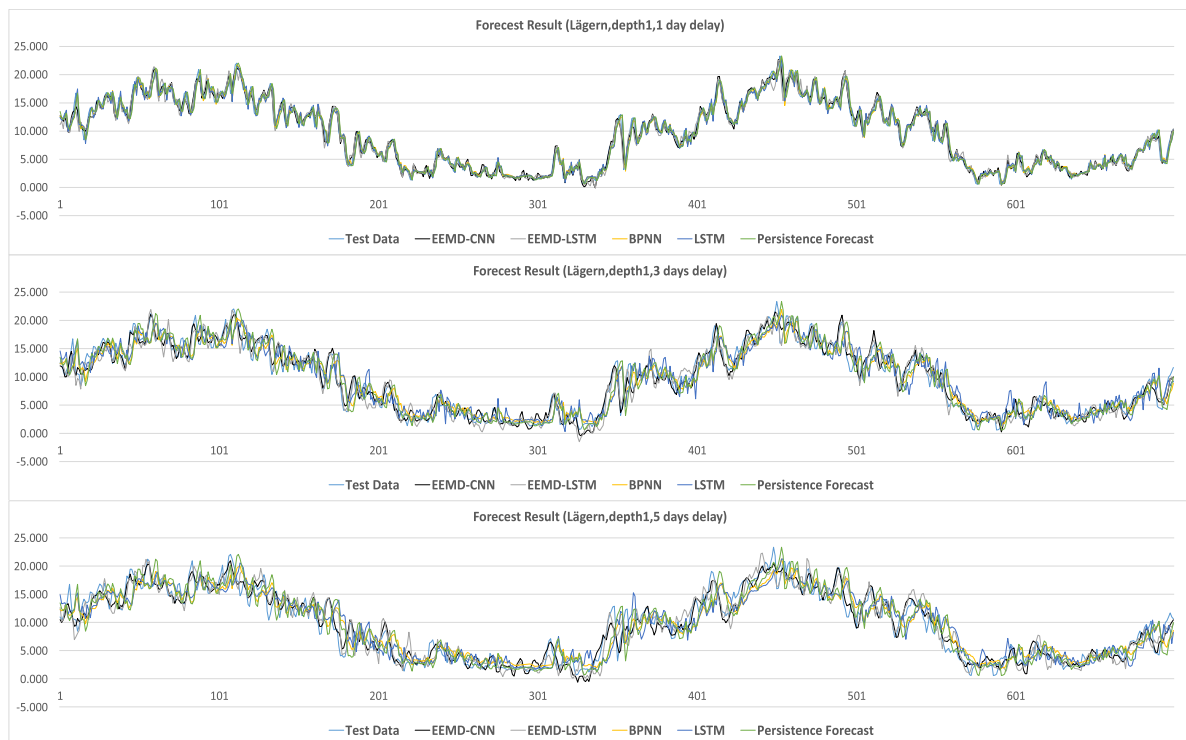


FIGURE 4. Performance comparison of the estimated values of temperature (°C) of the depth1(5cm) at Lägern among EEMD-CNN, EEMD-LSTM, BPNN, LSTM, persistence forecasting.

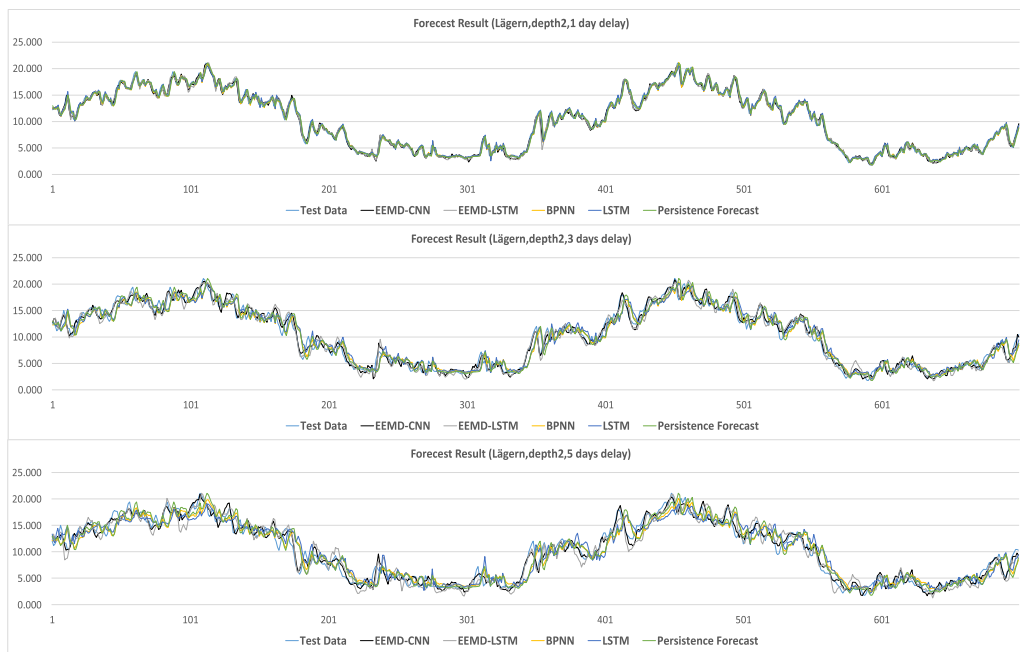


FIGURE 5. Performance comparison of the estimated values of temperature (°C) of the depth2(10cm) at Lägern among EEMD-CNN, EEMD-LSTM, BPNN, LSTM, persistence forecasting.

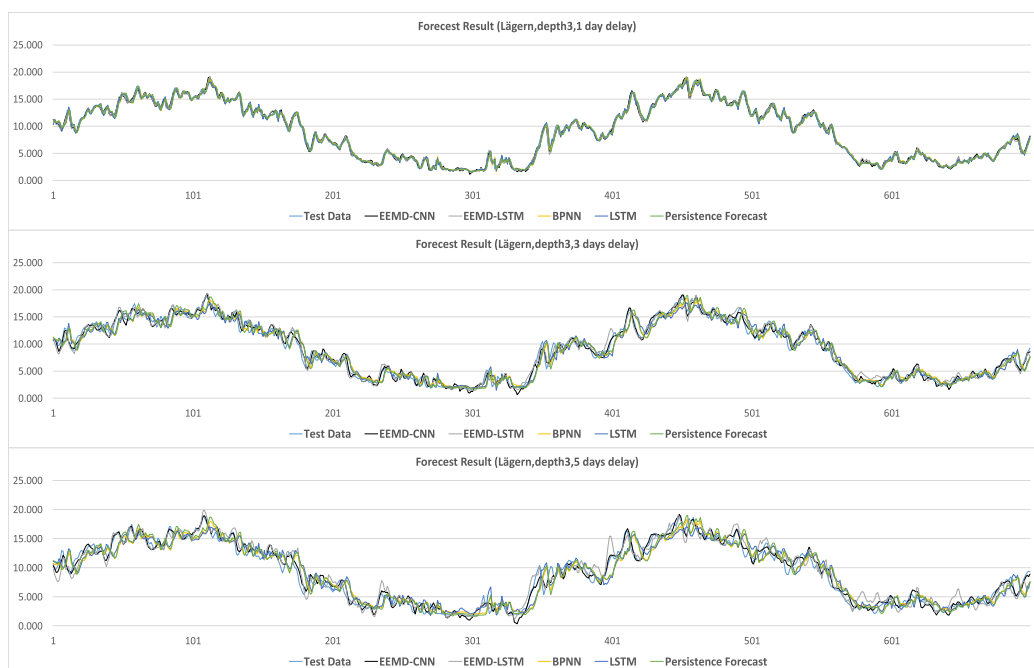


FIGURE 6. Performance comparison of the estimated values of temperature (°C) of the depth3(30cm) at Lägern among EEMD-CNN, EEMD-LSTM, BPNN, LSTM, persistence forecasting.

forecast temperature of three areas. Each area’s result including the data from three depths and the series of each depth are compared in three delays. The performance on Lägern’s soil temperature data is shown in Table 2, performance on Oensingen’s is in Table 3 and performance on Fluhli’s is in Table 4.

A. EXPERIMENTAL RESULTS IN LÄGERN

In the experiment in Lägern, the temperature series used in the model comes from depths of 5 centimeters, 10 centimeters

and 30 centimeters underground from Lägern. There are 3782 daily records (08/24/2004-12/31/2014) at each depth. We extract 3500 consecutive data from the series and split 3 series into train set and test set. Process data according to different models and train models with train set. Then use the trained models and test set to forecast temperature, the estimation results are showed in Table 2. As shown in Table 2, when forecasting the soil temperature in the case of one day delay, the models we used all have good performance, and

TABLE 2. Estimation results using five different models at Lägern. The results include three different delay tests at three depths.

Depth	Delay	Model	RMSE	MSE	MAE	R2-Score
Depth1	1day	EEMD-CNN	0.688	0.474	0.539	0.986
		EEMD-LSTM[40]	0.841	0.708	0.647	0.979
		BPNN[27]	1.098	1.205	0.812	0.964
		LSTM[32-36]	1.154	1.331	0.854	0.963
		PF[41]	1.199	1.436	0.882	0.958
	3days	EEMD-CNN	1.046	1.093	0.829	0.968
		EEMD-LSTM	1.365	1.865	1.072	0.944
		BPNN	2.141	4.583	1.658	0.864
		LSTM	2.289	5.241	1.768	0.846
		PF	2.666	7.105	2.052	0.848
	5days	EEMD-CNN	1.229	1.510	0.981	0.955
		EEMD-LSTM	1.781	3.174	1.401	0.910
		BPNN	2.440	5.955	1.934	0.823
		LSTM	2.551	6.506	1.984	0.811
		PF	2.667	7.112	2.054	0.801
Depth2	1day	EEMD-CNN	0.410	0.168	0.329	0.994
		EEMD-LSTM	0.484	0.234	0.381	0.992
		BPNN	0.610	0.372	0.464	0.987
		LSTM	0.623	0.388	0.469	0.987
		PF	0.691	0.478	0.523	0.984
	3days	EEMD-CNN	0.697	0.486	0.553	0.983
		EEMD-LSTM	0.958	0.919	0.752	0.970
		BPNN	1.407	1.979	1.113	0.932
		LSTM	1.427	2.036	1.112	0.930
		PF	1.814	3.29	1.404	0.927
	5days	EEMD-CNN	0.879	0.686	0.772	0.974
		EEMD-LSTM	1.308	1.711	1.024	0.944
		BPNN	1.740	3.029	1.371	0.895
		LSTM	1.791	3.208	1.382	0.888
		PF	1.817	3.303	1.408	0.889
Depth3	1day	EEMD-CNN	0.372	0.139	0.296	0.995
		EEMD-LSTM	0.400	0.162	0.314	0.994
		BPNN	0.498	0.248	0.378	0.990
		LSTM	0.504	0.378	0.254	0.990
		PF	0.572	0.328	0.430	0.987
	3days	EEMD-CNN	0.624	0.389	0.498	0.984
		EEMD-LSTM	0.857	0.736	0.671	0.976
		BPNN	1.209	1.461	0.956	0.942
		LSTM	1.240	1.537	0.958	0.941
		PF	1.255	1.575	0.972	0.938
	5days	EEMD-CNN	0.735	0.541	0.582	0.978
		EEMD-LSTM	1.109	1.230	0.854	0.949
		BPNN	1.527	2.333	1.203	0.907
		LSTM	1.583	2.505	1.223	0.900
		PF	1.582	2.503	1.229	0.903

the EEMD-CNN performs best of all models. In the estimation results of three days delay, performances of all models decline, but EEMD-CNN has the least decline in performance. In the performances of five days delay, EEMD-CNN still gets the best performance. Even at three different depths, EEMD-CNN's performances of five days delay are all better than the second good model EEMD-LSTM's performances

of three days delay. Figure 4–6 show the comparison of models' prediction results of three depth's in three kinds delay. Figure 7 shows the comparison of three kinds delay on forecasting the soil temperature of the same day. Figure 7 presents that even under the influence of delay, EEMD-CNN has more stable and accurate predictions. EEMD-LSTM is affected more by delay, the longer the delay, the greater the



FIGURE 7. Performance comparison of the same model at the same day of depth1(5cm) in Lägeren among test data and the estimated values of temperature (°C) of 1 day delay, 3 days delay and 5 days delay.

error. For BPNN and LSTM, the delay makes the prediction results of the two models deviate significantly. It can be seen

from the figure that two models try to fit the origin temperature, but the delay directly leads to the elevation of error.

TABLE 3. Estimation results using five different models at Oensingen. The results include three different delay tests at three depths.

Depth	Delay	Model	RMSE	MSE	MAE	R2-Score
Depth1	1day	EEMD-CNN	0.722	0.522	0.565	0.990
		EEMD-LSTM	1.026	1.053	0.794	0.981
		BPNN	1.112	1.236	0.865	0.977
		LSTM	1.171	1.373	0.905	0.975
		PF	1.155	1.334	0.898	0.976
	3days	EEMD-CNN	1.011	1.023	0.812	0.981
		EEMD-LSTM	1.567	2.458	1.231	0.955
		BPNN	2.165	4.688	1.732	0.914
		LSTM	2.414	5.830	1.898	0.893
		PF	2.235	4.993	1.774	0.909
	5days	EEMD-CNN	1.185	1.404	0.946	0.974
		EEMD-LSTM	1.943	3.777	1.529	0.931
		BPNN	2.649	7.017	2.085	0.872
		LSTM	2.881	8.302	2.249	0.848
		PF	2.756	7.595	2.153	0.682
Depth2	1day	EEMD-CNN	0.409	0.168	0.325	0.996
		EEMD-LSTM	0.450	0.202	0.352	0.996
		BPNN	0.513	0.263	0.394	0.994
		LSTM	0.512	0.262	0.397	0.994
		PF	0.589	0.346	0.450	0.992
	3days	EEMD-CNN	0.746	0.556	0.599	0.988
		EEMD-LSTM	1.093	1.198	0.874	0.973
		BPNN	1.291	1.667	1.018	0.963
		LSTM	1.407	1.982	1.097	0.956
		PF	1.336	1.786	1.052	0.960
	5days	EEMD-CNN	0.952	0.907	0.767	0.980
		EEMD-LSTM	1.341	1.801	1.060	0.960
		BPNN	1.733	3.003	1.376	0.933
		LSTM	1.895	3.594	1.491	0.920
		PF	1.765	3.114	1.400	0.931
Depth3	1day	EEMD-CNN	0.293	0.086	0.235	0.998
		EEMD-LSTM	0.316	0.100	0.254	0.998
		BPNN	0.352	0.124	0.265	0.997
		LSTM	0.350	0.122	0.263	0.997
		PF	0.449	0.354	0.202	0.995
	3days	EEMD-CNN	0.635	0.403	0.502	0.989
		EEMD-LSTM	1.029	1.060	0.836	0.972
		BPNN	1.082	1.171	0.865	0.969
		LSTM	1.081	1.168	0.857	0.969
		PF	1.154	1.331	0.911	0.965
	5days	EEMD-CNN	0.780	0.609	0.620	0.984
		EEMD-LSTM	1.314	1.728	1.079	0.955
		BPNN	1.548	2.397	1.204	0.937
		LSTM	1.568	2.458	1.239	0.936
		PF	1.595	2.543	1.243	0.934

Figure 8 is scatterplots of the predicted values and observed values, it shows degree of fitting between the estimated values of five models and observed values. It can be achieved that EEMD-CNN has the best performance under the same delay among five models from Figure 8. And with the increase in delay, the accuracy advantage of prediction becomes more obvious.

B. EXPERIMENTAL RESULTS IN OENSINGEN AND FLUHLI
In the experiment in Oensingen and Fluhi. The estimation results of these two areas are showed in Table 3 and Table 4. As shown in the two tables, the five models that are trained from the soil temperature data of these two areas, the performance of the estimation results obtained from these models is roughly the same as the models' performance in Lägern.

TABLE 4. Estimation results using five different models at Fluhli. The results include three different delay tests at three depths.

Depth	Delay	Model	RMSE	MSE	MAE	R2-Score
Depth1	1day	EEMD-CNN	0.466	0.218	0.360	0.994
		EEMD-LSTM	0.549	0.301	0.413	0.992
		BPNN	0.676	0.457	0.486	0.988
		LSTM	0.672	0.452	0.484	0.988
		PF	0.747	0.558	0.528	0.985
	3days	EEMD-CNN	0.772	0.596	0.608	0.984
		EEMD-LSTM	1.069	1.144	0.788	0.969
		BPNN	1.511	2.284	1.139	0.937
		LSTM	1.545	2.389	1.147	0.934
		PF	1.571	2.469	1.160	0.933
	5days	EEMD-CNN	0.942	0.888	0.733	0.976
		EEMD-LSTM	1.349	1.822	1.011	0.950
		BPNN	1.851	3.425	1.417	0.906
		LSTM	1.996	3.983	1.487	0.891
		PF	1.947	3.789	1.442	0.897
Depth2	1day	EEMD-CNN	0.375	0.141	0.289	0.996
		EEMD-LSTM	0.452	0.204	0.342	0.994
		BPNN	0.498	0.248	0.358	0.993
		LSTM	0.495	0.245	0.357	0.993
		PF	0.585	0.343	0.415	0.990
	3days	EEMD-CNN	0.690	0.476	0.550	0.986
		EEMD-LSTM	0.938	0.880	0.728	0.974
		BPNN	1.243	1.545	0.942	0.955
		LSTM	1.267	1.606	0.936	0.953
		PF	1.286	1.654	0.951	0.952
	5days	EEMD-CNN	0.814	0.663	0.654	0.981
		EEMD-LSTM	1.195	1.430	0.923	0.958
		BPNN	1.578	2.491	1.217	0.927
		LSTM	1.656	2.744	1.218	0.920
		PF	1.628	2.650	1.213	0.923
Depth3	1day	EEMD-CNN	0.276	0.077	0.210	0.998
		EEMD-LSTM	0.314	0.098	0.237	0.997
		BPNN	0.354	0.125	0.252	0.996
		LSTM	0.352	0.124	0.253	0.996
		PF	0.435	0.190	0.319	0.994
	3days	EEMD-CNN	0.548	0.301	0.428	0.990
		EEMD-LSTM	0.687	0.473	0.517	0.985
		BPNN	0.960	0.922	0.726	0.971
		LSTM	0.952	0.907	0.706	0.971
		PF	1.011	1.022	0.750	0.968
	5days	EEMD-CNN	0.667	0.445	0.526	0.986
		EEMD-LSTM	0.888	0.790	0.668	0.975
		BPNN	1.302	1.695	0.995	0.946
		LSTM	1.314	1.728	0.977	0.945
		PF	1.342	1.801	1.001	0.943

In Table 3 and Table 4, in these models which are trained by the data at same depth and with the same delay, EEMD-CNN performs best under the same data conditions. This rank is the same as the performance of EEMD-CNN in Lägern.

C. ANALYSIS

In this study, after training all models using data from three areas, it can be seen that EEMD-CNN has the best

performance both in the prediction of different temperature series and in the prediction of the temperature of a few days later. EEMD-CNN is the best choice in both cases. This result is probably because of the convolution kernel used in the convolutional neural network. In BPNN, the neuron used is the basic neural network cell. It is a common neuron in the feed-forward artificial neural network architecture. In LSTM and EEMD-LSTM, the neuron used is recurrent cell. This

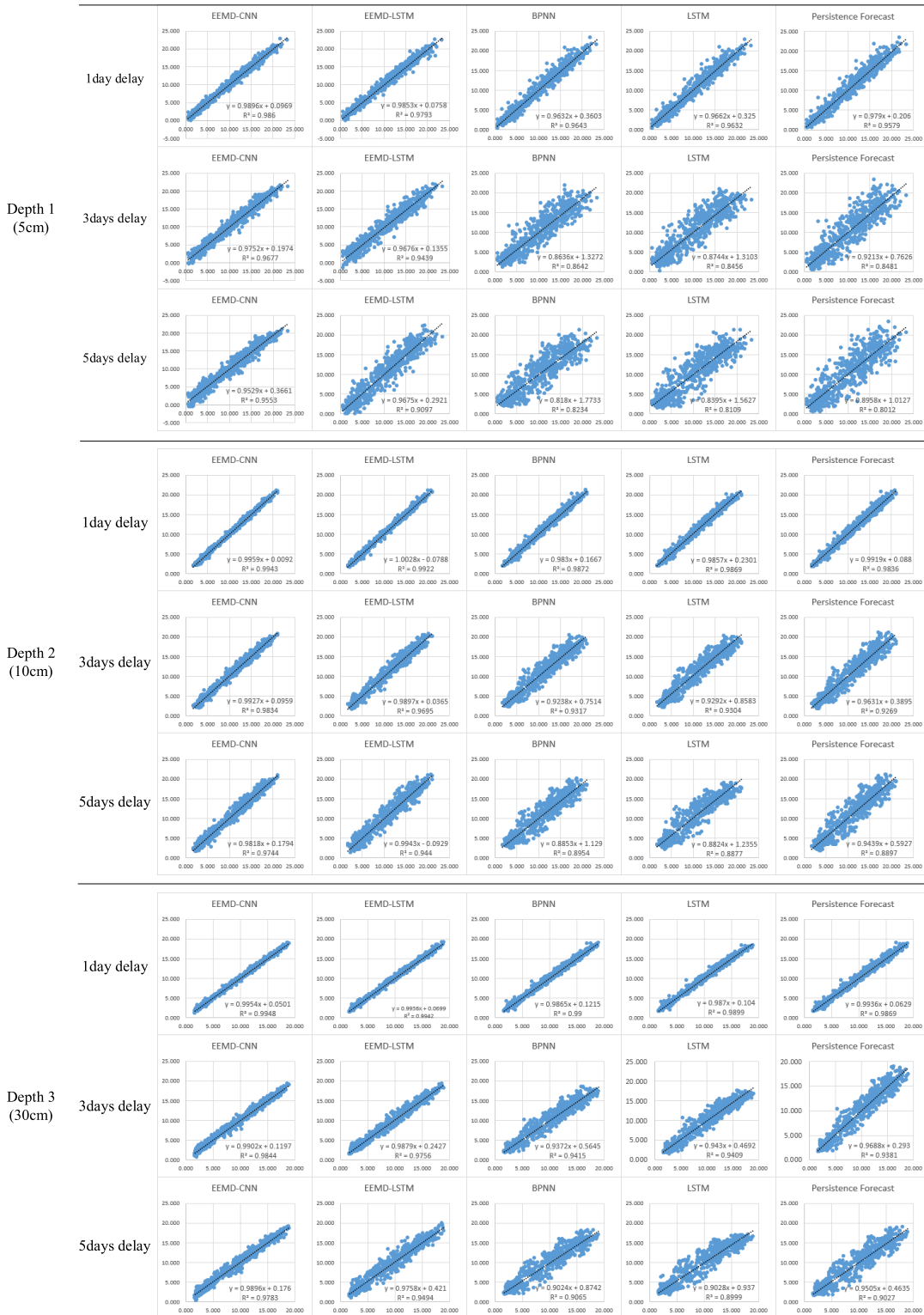


FIGURE 8. The scatterplots of the estimated and observed values of temperature (°C) using the EEMD-CNN and comparison models at Lägern.

kind of neuron has connections not only between cells, but also along the timeline. Recurrent cell enables the model to inherit some of the previous states during the training process, so that the continuous sequence can perform better

on the model. LSTM is an improvement over RNN. LSTM added a record of the state of the entire training process to the traditional RNN, so that LSTM can be better applied to continuous time series. Both BPNN and LSTM performed

reasonably well at predicting the temperature of one day after. The reason for that is the last input node in the BPNN is close to the predicted label temperature and assigned a higher weight, with a little correction and then the model could get a prediction with a small error. The same as LSTM, where the input of the last cell accounts for a more important effect, and the state which was produced by the previous cells adjust the final input to produce the final prediction. Then it is an acceptable performance in the case of one day delay. But in the case of three or even five days delays, BPNN still gives high weight to the last input node. However, with delays of three or five days, the input of the last input node has a greater error from the label. That makes the fitting effect of the model decreases with the increase of the delay. As the results shown in the Figure 7, the model tries to fit the results to the label, but the supplied sequence can only be predicted as far as possible based on the input value of the last node. Therefore, BPNN's prediction results in the case of three or five days delay are more similar with the line graph of persistence forecast results. LSTM has the same problem, each input into the recurrent cell is the temperature of 1 day in the temperature series, the input of the last cell influences the predicted of the output more, so when the delay increases, LSTM's result is similar with BPNN's.

With the addition of EEMD, the predicted result was improved. In EEMD-LSTM, it also uses recurrent cells, but the input of cells is not just the temperature of one day. The input also has the IMFs which are decomposed by EEMD from temperature series. IMFs can enrich the features input into the network and help the model to fit better. With the help of IMFs, EEMD-LSTM not only performs better than the LSTM in the case of a one-day delay. Even when predicting the temperature of a few days later, EEMD-LSTM fits better than LSTM. CNN is usually used for image data and voice data. These types of data are characterized by more than one dimension, which is just suitable for convolution kernel to extract local features within a range. And the way to extend the temperature series with IMFs is not just to insert an IMFs sequence between the temperatures of two adjacent days. IMFs can also be used to extend the data in longitudinal dimensions, make temperature series extend into two dimensions. And the expanded data just fits the data dimension which is suitable for convolution operation. Any block data in two dimensions is related to adjacent blocks. Therefore, different from those models using one-dimensional data such as BPNN and LSTM, the features extracted by convolution operation of EEMD-CNN can better capture the features that can affect the temperature change between the data. These features become clearer after being extracted by several layers of convolution operation. As a result, the EEMD-CNN model fits best, on both the data of different depths and the predicted temperature after a few days.

IV. CONCLUSION

In this article, EEMD-CNN is proposed for daily soil temperature forecasting. The daily temperature series is one

dimensional data. Therefore, EEMD is used to split the temperature series to make the single temperature features into multiple features. The features are combined with the original temperature series to obtain a new two-dimensional feature. Then the CNN is trained with the two-dimensional feature to forecast temperature. During the test, the EEMD-CNN model was tested under three areas, three depths and three different delays. And after comparing the performance with other models, EEMD-CNN has the best performance in all these cases. The tests show that EEMD-CNN may be useful for soil temperature prediction and will be a start for more fields of temperature prediction systems.

REFERENCES

- [1] A. Araghi, M. Mousavi-Baygi, J. Adamowski, C. Martinez, and M. van der Ploeg, "Forecasting soil temperature based on surface air temperature using a wavelet artificial neural network," *Meteorol. Appl.*, vol. 24, no. 4, pp. 603–611, Oct. 2017.
- [2] O. Kisi, H. Sanikhani, and M. Cobaner, "Soil temperature modeling at different depths using neuro-fuzzy, neural network, and genetic programming techniques," *Theor. Appl. Climatol.*, vol. 129, nos. 3–4, pp. 833–848, Aug. 2017.
- [3] A. Keshavarzi, F. Sarmadian, E.-S.-E. Omran, and M. Iqbal, "A neural network model for estimating soil phosphorus using terrain analysis," *Egyptian J. Remote Sens. Space Sci.*, vol. 18, no. 2, pp. 127–135, Dec. 2015.
- [4] L. Lai, X. Zhao, L. Jiang, Y. Wang, L. Luo, Y. Zheng, X. Chen, and G. M. Rimmington, "Soil respiration in different agricultural and natural ecosystems in an arid region," *PLoS ONE*, vol. 7, no. 10, Oct. 2012, Art. no. e48011.
- [5] Y. Zhang, "Soil temperature in Canada during the twentieth century: Complex responses to atmospheric climate change," *J. Geophys. Res.*, vol. 110, no. D3, 2005, doi: [10.1029/2004JD004910](https://doi.org/10.1029/2004JD004910).
- [6] M. Zeynoddin, H. Bonakdari, I. Ebtehaj, F. Esmailbeiki, B. Gharabaghi, and D. Z. Haghi, "A reliable linear stochastic daily soil temperature forecast model," *Soil Tillage Res.*, vol. 189, pp. 73–87, Jun. 2019.
- [7] L. L. Liang, D. A. Riveros-Iregui, R. E. Emanuel, and B. L. McGlynn, "A simple framework to estimate distributed soil temperature from discrete air temperature measurements in data-scarce regions," *J. Geophys. Res., Atmos.*, vol. 119, no. 2, pp. 407–417, Jan. 2014.
- [8] H. Tabari, A.-A. Sabziparvar, and M. Ahmadi, "Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region," *Meteorol. Atmos. Phys.*, vol. 110, nos. 3–4, pp. 135–142, Jan. 2011.
- [9] A. C. Harvey and P. H. J. Todd, "Forecasting economic time series with structural and box-jenkins models: A case study," *J. Bus. Econ. Statist.*, vol. 1, no. 4, pp. 299–307, Oct. 1983.
- [10] J.-F. Chen, W.-M. Wang, and C.-M. Huang, "Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting," *Electric Power Syst. Res.*, vol. 34, no. 3, pp. 187–196, Sep. 1995.
- [11] J. Guo, Z. Xie, Y. Qin, L. Jia, and Y. Wang, "Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM," *IEEE Access*, vol. 7, pp. 42946–42955, 2019.
- [12] M. Bermúdez, L. Cea, and J. Puertas, "A rapid flood inundation model for hazard mapping based on least squares support vector machine regression," *J. Flood Risk Manage.*, vol. 12, no. S1, Oct. 2019, Art. no. e12522.
- [13] M. Delbari, S. Sharifzari, and E. Mohammadi, "Modeling daily soil temperature over diverse climate conditions in Iran—A comparison of multiple linear regression and support vector regression techniques," *Theor. Appl. Climatol.*, vol. 135, nos. 3–4, pp. 991–1001, Feb. 2019.
- [14] C. Hutengs and M. Vohland, "Downscaling land surface temperatures at regional scales with random forest regression," *Remote Sens. Environ.*, vol. 178, pp. 127–141, Jun. 2016.
- [15] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [16] M. Bilgili, "Prediction of soil temperature using regression and artificial neural network models," *Meteorol. Atmos. Phys.*, vol. 110, nos. 1–2, pp. 59–70, Dec. 2010.

- [17] M. Bilgili, "The use of artificial neural networks for forecasting the monthly mean soil temperatures in Adana, Turkey," *Turkish J. Agricult. Forestry*, vol. 35, no. 1, pp. 83–93, 2011.
- [18] M. Hayati and Z. Mohebi, "Application of artificial neural networks for temperature forecasting," *Eng. Technol.*, vol. 28, no. 2, pp. 275–279, 2007.
- [19] H. Tabari, P. Hosseinzadeh Talaei, and P. Willems, "Short-term forecasting of soil temperature using artificial neural network," *Meteorol. Appl.*, vol. 22, no. 3, pp. 576–585, Jul. 2015.
- [20] J. Shiri, S. Kim, and O. Kisi, "Estimation of daily dew point temperature using genetic programming and neural networks approaches," *Hydrol. Res.*, vol. 45, no. 2, pp. 165–181, Apr. 2014.
- [21] S. Salcedo-Sanz, R. C. Deo, L. Carro-Calvo, and B. Saavedra-Moreno, "Monthly prediction of air temperature in australia and new zealand with machine learning algorithms," *Theor. Appl. Climatol.*, vol. 125, nos. 1–2, pp. 13–25, Jul. 2016.
- [22] K. Rasouli, W. W. Hsieh, and A. J. Cannon, "Daily streamflow forecasting by machine learning methods with weather and climate inputs," *J. Hydrol.*, vols. 414–415, pp. 284–293, Jan. 2012.
- [23] G. Mihalakakou, "On estimating soil surface temperature profiles," *Energy Buildings*, vol. 34, no. 3, pp. 251–259, Mar. 2002.
- [24] Y. LeCun, Y. Bengio, and G. J. N. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [25] J. H. Friedman "Greedy function approximation: A gradient boosting machine," in *Proc. Ann. Statist.*, 2001, pp. 1189–1232.
- [26] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.
- [27] N. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Proc. Int. Joint Conf. Neural Netw.*, 1989, pp. 65–93.
- [28] Y. Feng, N. Cui, W. Hao, L. Gao, and D. Gong, "Estimation of soil temperature from meteorological data using different machine learning models," *Geoderma*, vol. 338, pp. 67–77, Mar. 2019.
- [29] W. Zaremba, I. Sutskever, and O. Vinyals, "Recurrent neural network regularization," 2014, *arXiv:1409.2329*. [Online]. Available: <https://arxiv.org/abs/1409.2329>
- [30] R. E. Abdel-Aal, "Hourly temperature forecasting using abductive networks," *Eng. Appl. Artif. Intell.*, vol. 17, no. 5, pp. 543–556, Aug. 2004.
- [31] A. Shirvani, F. Moradi, and A. A. Moosavi, "Time series modelling of increased soil temperature anomalies during long period," *Int. Agrophysics*, vol. 29, no. 4, pp. 509–515, Oct. 2015.
- [32] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks," *Neurocomputing*, vol. 323, pp. 203–213, Jan. 2019.
- [33] J. Zhang, Y. Zhu, X. Zhang, M. Ye, and J. Yang, "Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas," *J. Hydrol.*, vol. 561, pp. 918–929, Jun. 2018.
- [34] P. Hewage, A. Behera, M. Trovati, and E. Pereira, "Long-short term memory for an effective short-term weather forecasting model using surface weather data," in *Artificial Intelligence Applications and Innovations. AIAI (IFIP Advances in Information and Communication Technology)*, vol. 559. Cham, Switzerland: Springer, 2019, pp. 382–390, doi: [10.1007/978-3-030-19823-7_32](https://doi.org/10.1007/978-3-030-19823-7_32).
- [35] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [36] F. Wu, S. Lu, L.-A. Armando, and J. She, "Temperature prediction based on long short term memory networks," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2019, pp. 312–317.
- [37] S. Berkahn, L. Fuchs, and I. Neuweiler, "An ensemble neural network model for real-time prediction of urban floods," *J. Hydrol.*, vol. 575, pp. 743–754, Aug. 2019.
- [38] Y. Li, R. W. Liu, Z. Liu, and J. Liu, "Similarity grouping-guided neural network modeling for maritime time series prediction," *IEEE Access*, vol. 7, pp. 72647–72659, 2019.
- [39] R. W. Liu, J. Chen, Z. Liu, Y. Li, Y. Liu, and J. Liu, "Vessel traffic flow separation-prediction using low-rank and sparse decomposition," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 1–6.
- [40] X. Zhang, Q. Zhang, G. Zhang, Z. Nie, Z. Gui, and H. Que, "A novel hybrid data-driven model for daily Land Surface Temperature forecasting using long short-term memory neural network based on ensemble empirical mode decomposition," *Int. J. Environ. Res. Public Health*, vol. 15, no. 5, p. 1032, May 2018.
- [41] S. Rasp, P. D. Dueben, S. Scher, J. A. Weyn, S. Mouatadid, and N. Thuerey, "WeatherBench: A benchmark data set for data-driven weather forecasting," *J. Adv. Model. Earth Syst.*, vol. 12, no. 11, Nov. 2020, doi: [10.1029/2020MS002203](https://doi.org/10.1029/2020MS002203).
- [42] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London. Ser. A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
- [43] N. E. Huang, Z. Shen, and S. R. Long, "A new view of nonlinear water waves: The Hilbert spectrum," *Annu. Rev. Fluid Mech.*, vol. 31, no. 1, pp. 417–457, Jan. 1999.
- [44] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.
- [45] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," 2014, *arXiv:1404.2188*. [Online]. Available: <https://arxiv.org/abs/1404.2188>
- [46] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.
- [47] S. Lawrence, C. L. Giles, A. Chung Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, Jan. 1997.
- [48] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1790–1798.



HUIBOWEN HAO received the bachelor's degree in computer science and technology from Jilin University, Jilin, China, in 2017. He is currently pursuing the master's degree with the College of Computer Science and Technology, Changchun Normal University, Changchun, China. His current research interests include deep learning, image understanding, and atmospheric physics.



FANHUA YU received the B.S. degree in mathematics and applied mathematics from Changchun Normal University, Changchun, China, in 1992, and the M.S. degree in computer software and theory and the Ph.D. degree in road and railway engineering from Jilin University, Changchun, in 2004 and 2008, respectively. He is currently a Professor with the College of Computer Science and Technology, Changchun Normal University. His research interests include image processing and artificial intelligence.



QINGLIANG LI received the Ph.D. degree from Jilin University, Changchun, China, in 2016. He is currently with the College of Computer Science and Technology, Changchun Normal University. His current research interests include deep learning, image understanding, and atmospheric physics.