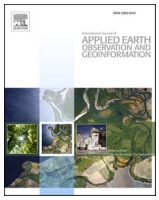
International Journal of Applied Earth Observations and Geoinformation 104 (2021) 102550

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Convolutional neural network model for soil moisture prediction and its transferability analysis based on laboratory Vis-NIR spectral data

Yu Chen [a], Lin Li [b,*], Michael Whiting [c], Fang Chen [a], Zhongchang Sun [a,d], Kaishan Song [e], Qinjun Wang [a,d]

[a] CAS Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, PR China
[b] Department of Earth Sciences, Indiana University-Purdue University Indianapolis (IUPUI), IN 46202, USA
[c] Center for Spatial Technology and Remote Sensing, Department of Land, Air and Water Resources, University of California, Davis 95616, USA
[d] Key Laboratory of the Earth Observation of Hainan Province, Aerospace Information Research Institute, CAS, Hainan Research Institute, Sanya 572029, PR China
[e] Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, PR China

ARTICLE INFO

Keywords:
Convolutional neural network (CNN)
Soil
Spectral analysis
Knowledge-based transfer learning method
Geological environment

ABSTRACT

Laboratory visible near infrared reflectance (Vis-NIR, 400–2500 nm) spectroscopy has the advantages of simplicity, fast and non-destructive which was used for SM prediction. However, many previously proposed models are difficult to transfer to unknown target areas without recalibration. In this study, we first developed a suitable Convolutional Neutral Network (CNN) model and transferred the model to other target areas for two situations using different soil sample backgrounds under 1) the same measurement conditions (DSSM), and 2) under different measurement conditions (DSDM). We also developed the CNN models for the target areas based on their own datasets and traditional PLS models was developed to compare their performances. The results show that one dimensional model (1D-CNN) performed strongly for SM prediction with average R^2 up to 0.989 and RPIQ up to 19.59 in the laboratory environment (DSSM). Applying the knowledge-based transfer learning method to an unknown target area improved the R^2 from 0.845 to 0.983 under the DSSM and from 0.298 to 0.620 under the DSDM, which performed better than data-based spiking calibration method for traditional PLS models. The results show that knowledge-based transfer learning was suitable for SM prediction under different soil background and measurement conditions and can be a promising approach for remotely estimating SM with the increasing amount of soil dataset in the future.

1. Introduction

Soil moisture (SM) is an important environmental variable for understanding the energy exchange between the atmosphere and the underlying land surface and the interactions among the climatic, hydrological, and biological subsystems (Carrao et al., 2016; Lobell and Asner, 2002). Traditional methods for measuring SM content including thermogravimetric, neutron scattering, gamma ray attenuation, soil electrical conductivity, tensiometry, hygrometry and soil dielectric are time consuming or have some other disadvantages (Birdal et al., 2007). Moreover, these in situ point-based methods are not suitable for mapping SM in a large scale. Laboratory visible near infrared reflectance (Vis-NIR, 400–2500 nm) spectroscopy has great advantage in terms of saving cost of SM analysis (Bao et al., 2018).

The methods used for SM prediction from Vis-NIR spectral data could be grouped into spectral indices, radiative transfer, and statistical regression. The spectral index methods use only a few bands and can only be applied to areas with less variation in soil types (Amani et al., 2016; Fabre et al., 2015). Radiative transfer models can achieve a strong performance for some datasets but sometime require soil information that is somewhat difficult to access (Sadeghi et al., 2017). Statistical regression methods build a linear or non-linear relationship between the spectrum and SM by use of principal components regression (PCR) (Chang et al., 2001), partial least squares regression (PLS) (Araujo et al., 2014), inverted Gaussian spectrum fitting (Whiting et al., 2004), support vector machines (SVM) (González Costa et al., 2017) or other machine learning approaches. These methods are the most favorable to the literature especially for those utilizing laboratory measured soil spectra.

* Corresponding author.
E-mail addresses: chenyu@radi.ac.cn (Y. Chen), ll3@iupui.edu (L. Li), chenfang@radi.ac.cn (F. Chen), sunzc@radi.ac.cn (Z. Sun), Songks@iga.ac.cn (K. Song), wangqj@radi.ac.cn (Q. Wang).

https://doi.org/10.1016/j.jag.2021.102550
Received 22 May 2021; Received in revised form 20 August 2021; Accepted 14 September 2021
Available online 4 October 2021

In the recent years, the deep learning method regarded as an effective regression tool in the context of quantifying soil properties has witnessed a great success (Tsakiridis et al., 2020; Zhang et al., 2019). Some studies also analyzed the transferability of CNN model between global model and localization (Padarian et al., 2019). However, these studies did not fully consider the different spectral measurement conditions in transfer learning. Therefore, a high-precision CNN model might perform poorly when the model is transferred to other regions under different spectral measurement conditions (Romero et al., 2018).

As a deep learning model, a trained Convolutional Neural Network (CNN) is often assumed to be adoptable for other unknown areas via transfer learning with the expectation that the knowledge learned by the trained model is applicable to a related problem for which the underlying knowledge is not fully understood. The primary goal of this study was to investigate the extent to which a trained CNN is transferable for estimating SM for different areas or different measurement conditions. We first developed a suitable CNN model based on the 1-dimensional input Vis-NIR spectroscopy data (namely 1D-CNN model) for SM prediction. Then, based on transfer learning method, we calibrated the model to another area in two situations: 1) one has different soil sample background but the same measurement conditions (DSSM), 2) the other has different soil sample background and different measurement conditions (DSDM). We also tested PLS models under the same conditions to compare their performances with the CNN transfer learning method.

2. Materials and methods

2.1. Soil samples

2.1.1. Sample distribution

This study used soil samples collected from three different study sites (see Table 1). The first dataset contains 20 samples collected from the vineyards and grain lands near Tomelloso, Castilla-La Mancha, Spain (Site 1) (Whiting et al., 2004). The parent materials of these soils are characterized by high calcium carbonate concentrations and described as the limestone and marls mixed with conglomerates of clays and gravels. The older surfaces are Petric Calcisols and the younger alluvium Haplic Calcisols based on the UN Food and Agriculture Organization soil classification system (Sanchez et al., 1996). The second dataset contains 20 samples collected from Tulare Lake playa in the southern San Joaquin Valley, near Lemoore, California, United States (Site 2) (Whiting et al., 2004) where the principal crops are cotton, tomatoes, and grain and the soils contain illite and montmorillonite of fine-loamy, mixed (calcareous), thermic Typic Torriorthents, and fine, montmorillonitic, thermic Typic Natrargrids within the US Department of Agriculture soil classification system (USDA, 1978). Soil samples from Sites 1 and 2 were sieved to smaller than 2 mm after a light grinding with mortar and pestle. The fraction passing through the sieve included coarse and finer sands, silt, clay, and small aggregates of clay and sand. Each well-mixed sample was divided into two replicates for spectral measurement.

The third soil sample dataset contains 211 samples collected from north of Indianapolis, Indiana, United States (Site 3), where the principal crops are corn, soybean, and wheat and the dominant soil types are Alfisols, Inceptisols, and Mollisols. During the sample collection, soil samples were kept fresh in Zip-loc bags (17 cm × 20 cm), transported over ice in a cooler to the laboratory and stored in refrigerators (4 °C) before soil property analyses could be performed. Additionally, crop residues and large size gravels were manually removed before sealing the bags. Unlike the samples from sites 1 and 2, these soil samples were not sieved.

2.1.2. Moisture measurement

Replicate samples from site 1 and site 2 were measured at 15 different moisture contents. The water contents were increased in 0.05 g water/g soil increments by adding water, allow equilibration, then weighing just before spectral measurement. The gravimetric water contents were determined using the replicate's oven dried weight at 80 °C for 36 or more hours s (Whiting et al., 2004). The soil moisture for the 211 soil samples from site 3 were measured by weighting the soil samples before and after oven dry at 105 °C for more than 24 h. SM was calculated as:

$$\theta = \frac{W - W'}{W'}$$

where W and W' are the weight (g) of the sample before and after oven dry, respectively.

The SM statistics for the three sites is presented in Table 2. It is expected that the distributions of the SM values of site 1 and site 2 are somewhat different from site 3, which is skewed toward the lower water contents of actual field moisture.

2.1.3. Spectral measurement and preprocessing

Spectral measurement for the samples collected at site 1 and site 2 were made using a Cary 5E spectrophotometer (Varian Inc., San Jose, CA) (Whiting et al., 2004) with the spectral range 400–2500 nm and 1 nm sampling interval. Each of two soil replicates were measured three times at approximately 15 moisture levels. The holders were rotated approximately 60° for each measurement to minimize the influence due to surface geometry. Of the total spectra collected, 121 were discarded because the replicates second oven dry soil weight was not within 0.1% of the initial oven dry weight. The total number of spectra for the combined replicated samples and moisture levels was 1779 and 1700 spectra for site 1 and site 2, respectively.

Spectral measurement for site 3 was made using an ASD Field Spec Pro FR spectrometer (Analytical Spectral Devices, Inc. Boulder, USA) with the spectral range 350–2500 nm. The spectrum acquisition software was used to interpolate reflectance data to a sampling interval of 1 nm (Araujo et al., 2014). To ensure the accuracy of the data, the spectral were measured in parallel with the gravimetric moisture measurement of the soil samples. Each soil sample was measured once (smoothing 5 automatic measurements by the spectrometer).

To maintain the consistency between the spectral data measured by the two different instruments and eliminate the instrument noise at both ends of the spectrum, the reflectance data was resampled to 400 to 2429 nm with 1 nm interval. Then, the spectral data were downsampled by 1:10 (one average value every 10 nm) to reduce the redundancy of spectrum dimension.

Table 1
Soil information from 3 sites.

Location	Soil samples	Numbers of SM data	Clay range (%)	$CaCO_3$ range (%)	Soil textures	Mineral composition	Land use
Site 1: Tomelloso, Castilla-La Mancha, Spain	20	1779	22–57	23–66	Loam sandy loam silt loam	limestone and marls mixed with conglomerates of clays and gravels	Vineyards and grain lands
Site 2: Lemoore, CA, USA	20	1700	22–43	0.5–2.4	Clay loam sandy clay loam silty clay loam	Mixtures of illite and montmorillonite	Cotton, tomatoes, and grain
Site 3: North of Indianapolis, IN, USA	211	211	16–39	No data	Clay loam	Alfisols, Inceptisols, and Mollisols	Corn, soybean, and wheat

Table 2
Soil moisture information for 3 sites.

Location	Numbers of SM data	Min*	Max*	Mean*	SD*	Skewness	Kurtosis
Site 1	1779	0	0.506	0.189	0.156	0.211	−1.374
Site 2	1700	0	0.458	0.192	0.152	0.193	−1.396
Site 3	211	0.011	0.646	0.134	0.095	1.184	3.326

* g water/ g soil.

2.2. Modeling methods

2.2.1. Convolutional neural networks

The earliest Convolutional Neural Networks (CNN) was developed by LeCun for classifying images of handwritten digits in 1989 (LeCun et al., 1990) and then developed to different architectures by many scholars (Krizhevsky et al., 2017; Simonyan and Zisserman, 2015). It usually consists of one or more convolution, pooling, or fully connected operation layers. The convolution operation is to extract different features from the input layer. The pooling operation is to find the macro scale features and reduce the dimensionality of feature maps (Chen et al., 2020; Wang et al., 2020). The fully connected layer reorganizes extracted features to map to the final output.

In this study, a 1D-CNN architecture was built similar to those described in the literature (Liu et al., 2018; Tsakiridis et al., 2020). Fig. 1 shows the overall architecture of the proposed CNN model. First, the input layer is the spectral data which has 203 bands ranging from 400 to 2429 nm with 10 nm intervals. Here, we resampled the spectral data from its original recording resolution 1 nm to 10 nm and down sampling the spectral resolution can reduce the spectral noise and the redundancy of spectral data for prediction of soil properties (Chen et al., 2021; Tsakiridis et al., 2020). The input layer could be spectral reflectance R or its transformation such as logarithmic (log(1/R)), continuum removal (CR) (Clark and Roush, 1984), standard normal variate (SNV) (Barnes et al., 1989), first derivative (FD) (Norris and Williams, 1984) or second derivative (SD) (Rinnan et al., 2009) etc. We finally used the SNV transformed data as model inputs. The reason was that soil moisture lowers spectral reflectance at longer wavelengths more than shorter wavelengths (Lobell and Asner, 2002; Yue et al., 2019). This inharmonic reflectance reduction originates from varying water absorption coefficients to which SNV is very suitable for minimizing this viability. The second layer is a 1-D convolutional layer with 16 filters and the kernel size is 3*1. The Batch Normalization function is used to normalize the layer by shifting and scaling the activations, and Leaky ReLU ($\alpha = 0.01$) is used as an activation function. The third layer is a max pooling layer with size 2*1, which down-samples the filtered signals by halving the dimensionality. The fourth layer is another 1-D convolutional layer with the same parameters as the second layer but has 32 filters. The fifth layer is a flatten layer which contains 3232 nodes in this study. The sixth and seventh layer are both fully connected (Dense) layer with sizes of 10 and 6 respectively and a Leaky ReLU activation ($\alpha = 0.01$) function. All the convolutional layers and fully connected layers use L2 kernel regularizer ($\lambda_2 = 0.0004$) to penalize the weights and reduce over-fitting. The last layer uses the ReLU activation function to determine the soil moisture value. The layer characteristics are summarized in Table 3.

2.2.2. Calibration for model transfer

In this study, two different situations were considered to evaluate the performance of a CNN model for transfer learning calibration as shown in Fig. 2.

Situation 1- different soil sample background and same measurement (DSSM) conditions: Firstly, model was built on site 1 (mark the model as CNN-m1) and tested to site 2 without calibration (mark the model as CNN-m2′$_1$ with the subscript 1 representing that this model is built from site 1 and the label apostrophe "′" representing that the model was built without calibration). Secondly, the model CNN-m1 was recalibrated using the soil samples from site 2 to test its transferability (mark the model as CNN-m2$_1$). Lastly, model CNN-m2 was developed using the samples from site 2 itself to compare the performances.

Situation 2- different soil sample background and different measurement (DSDM) condition: Firstly, model was built on site 1 and site 2 (mark the model as CNN-m12) and tested to site 3 without calibration (mark the model as CNN-m3′$_{12}$). Secondly, the model CNN-m12 was recalibrated using the soil samples from site 3 to test its transferability (mark the model as CNN-m3$_{12}$). Lastly, model CNN-m3 was developed using the samples from site 3 itself to compare the performances.

To compare the accuracy of the proposed CNN model, we also tested the performance of Partial Least Squares (PLS) regression under the same conditions. PLS performs linear regression in a transformed input

Table 3
Sequence and description of the layers used in the CNN architecture.

Layer	Type	Kernel size	Filters	width	Activation
1	Input + batch norm.	–	–	203	–
2	Convolutional + batch norm.	3*1	16	203	Leaky ReLU (alpha = 0.01)
3	Maxpooling	2*1	–	101	–
4	Convolutional + batch norm.	3*1	32	101	Leaky ReLU (alpha = 0.01)
5	Flatten	–	–	3232	–
6	Dense (Fully connected)	–	–	10	Leaky ReLU (alpha = 0.01)
7	Dense (Fully connected)	–	–	6	Leaky ReLU (alpha = 0.01)
8	Output	–	–	1	ReLU

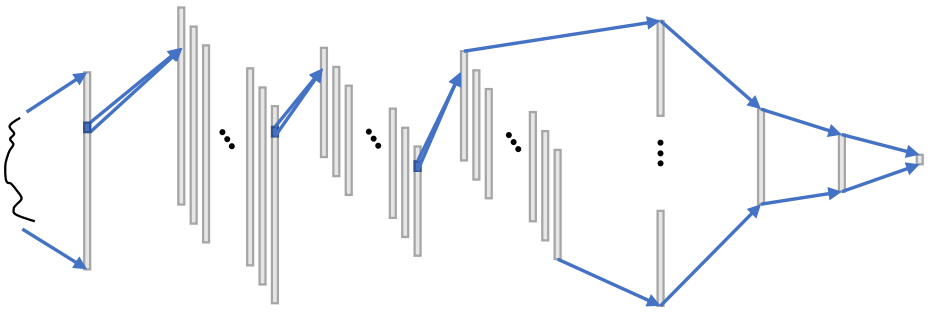


Fig. 1. Architecture of proposed CNN model for SM prediction.

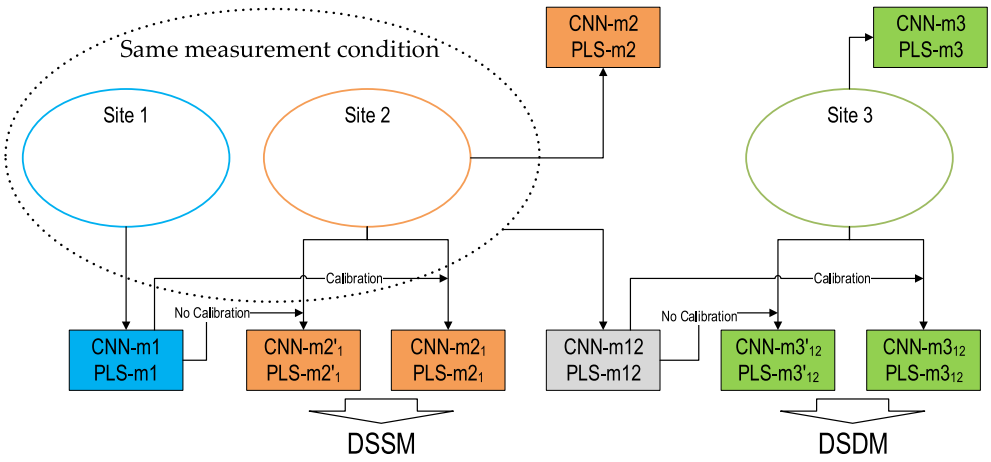


Fig. 2. Two different conditions for transfer learning.

space, which is formed by successively selecting orthogonal factors (latent variables) to maximize the covariance between predictors and the response variable (Wold et al., 1983). It has been commonly used in estimating and mapping soil properties (Araujo et al., 2014; Goge et al., 2014; Guerrero et al., 2016). The PLS model used in this study was the SIMPLS algorithm (Dejong, 1993) implemented using MATLAB 2020a and calibrated with the spiking method by adding a series of spiking subsets from the target area (Hong et al., 2018).

2.2.3. Model assessment

The model performance was assessed in terms of coefficient of determination (R^2), root-mean-square error (RMSE) and the ratio of performance to interquartile range *(RPIQ)*, which were calculated by the following equations:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \quad (1)$$

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

$$RPIQ = \frac{IQR}{RMSE} = \frac{Q3 - Q1}{RMSE} \quad (3)$$

where n stands for the number of samples, y is the measured value, $\bar{y}$ is the mean of the measured value of SM, and $\hat{y}$ is the predicted value of SM. *RPIQ* takes both the prediction error and the variation of observed values into account, without making assumptions about the distribution of the observed values (Tsakiridis et al., 2020). It is defined as the interquartile range (IQR) of the observed values divided by the RMSE of prediction. Q3 and Q1 are the 25th percentile and 75th percentile of the observed values, respectively.

2.3. CNN experiment environment

All the CNN models for estimating SM were developed using the TensorFlow 2.0 software with python. TensorFlow is a flexible software library which enables users to efficiently program and train neural network and deploy them to production (Pang et al., 2020). The hardware environment of this study was NVIDIA Quadro P2000 graphics card, Intel i7-8850H processor, and the memory was 32 GB. The Adam optimizer with a default learning rate of 0.001 and mean square error (MSE) loss with an early stop (patience = 10) was used for training these models.

3. Results

3.1. SM prediction results

A total of 3479 SM samples including 1779 from site 1 and 1700 from site 2 were combined to build the models CNN/PLS-m12. We randomly selected 75% (2609 samples) of the total for model training and 25% (653 samples) from the 2609 subsamples for validation. A subsample of 25% (870 samples) from the total of 3479 samples was used for independent testing. This training, validation and testing subsampling was repeated for 10 rounds. An example for training and validation loss curves for the CNN model is shown in Fig. 3. It took approximately 3 to 5 min and 300 to 400 epochs to process each round for the CNN-m12 model. For the 10 rounds of independent testing, the scatter plots for measured and predicted SM results are shown in Fig. 4a and Fig. 4b for CNN-m12 and PLS-m12, respectively. An R^2 value boxplot for 10 round independent tests of CNN-m12 and PLS-m12 is shown in Fig. 5. Compared with the PLS-m12, the CNN-m12 model performed better. Its average R^2, RMSE and RPIQ values were 0.989, 0.016 and 19.59, respectively, while the PLS-m12 model corresponded to 0.976, 0.024 and 12.99, respectively.

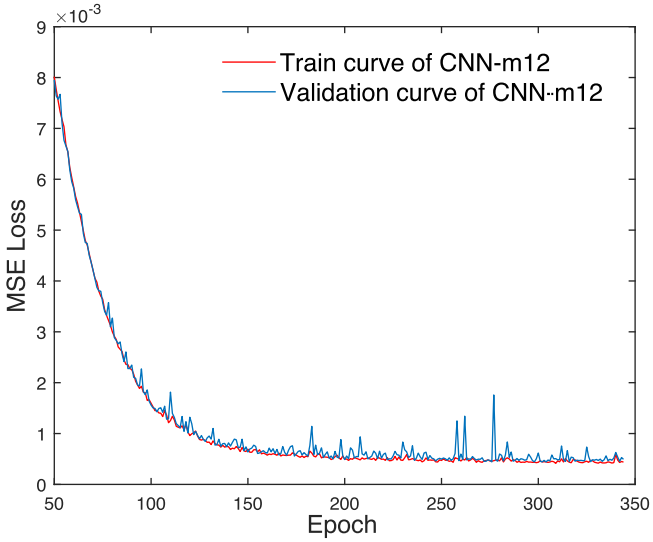


Fig. 3. An example of training and validation loss curves for CNN-m12.

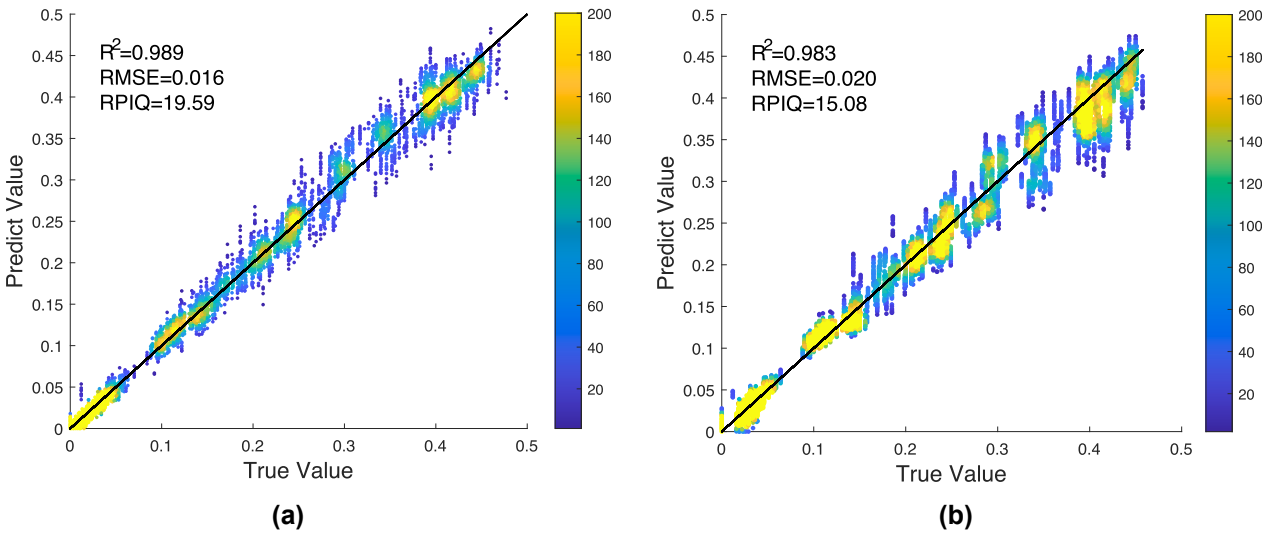


Fig. 4. Scatter plots for measured and model predicted SM for 10 round independent tests (a) CNN-m12; (b) PLS-m12. The values listed on the upper left corner are the average accuracy of 10 round tests.

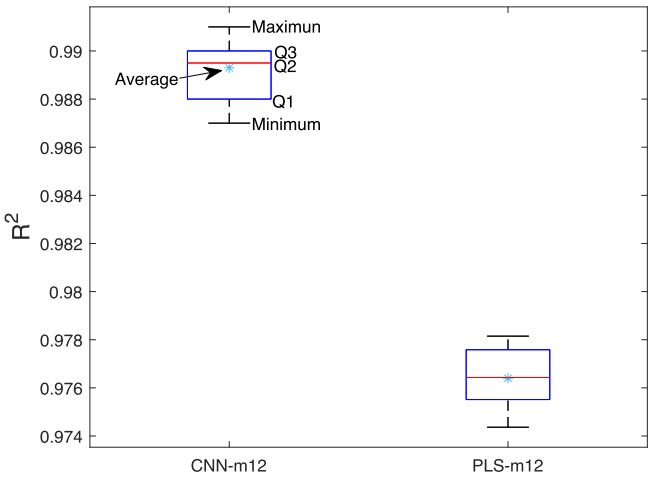


Fig. 5. Comparison of R^2 values for 10 round independent tests of CNN-m12 and PLS-m12. The three lines in the boxplot represent the 25th percentile (Q1), median (Q2), and 75th percentile (Q3), respectively.

3.2. Calibration in different conditions

3.2.1. Different soil sample background and same measurement (DSSM) conditions

The soil samples from site 1 are much greater in $CaCO_3$ and coarser texture than those from site 2. The CNN/PLS-m1 model was built using the 1779 samples from site 1. It was tested using 90% random sampling of 1700 samples from site 2 without calibration namely CNN/PLS- $m2'_1$. Next, the CNN/PLS-m1 model was recalibrated using 10% random sampling of 1700 samples from site 2. This model was named as CNN/PLS-$m2_1$ and tested on the other 90% samples of site 2. The third model namely CNN/PLS-m2 used 10% random sampling of 1700 samples from site 2 and was tested on the remaining 90% of the samples. The results from 10 round independent testing of the three procedures were assessed based on the corresponding scatter plots between measured and predicted SM together (Fig. 6). The R^2 value boxplot of independent testing for the three CNN models for 10 rounds was shown in Fig. 7.

The following observations can be made from the results described above: a) regardless of whether it is recalibrated or not, the performance of the CNN model is better than the PLS model when transferred to a new target area (Fig. 6a vs 6b, Fig. 6c vs 6d); b) even if the same measurement conditions are maintained, transferring a trained model to another area without calibration may also cause large deviation. The extent of deviation may depend on the difference between the two sites. In this study, soil samples with low SM were estimated at a higher accuracy than the samples with high SM (Fig. 6a); c) using even a small dataset to recalibrate a trained model based on the transfer learning method can lead to a stable and highly accurate result when generalizing a model to a new area (Fig. 6a vs 6c, Fig. 7); d) using a small dataset to train a new CNN model was unstable and even performed worse than the traditional method. In this study, 10% of 1700 samples from site 2 were randomly selected for training the model CNN-m2 and PLS-m2. The testing result showed a difficulty to train the model CNN-m2 because the model was prone to overfitting during training and thus the performance on the testing dataset was unstable (sometimes has high accuracy and sometimes has very low accuracy) (Fig. 6e vs 6f and Fig. 7).

3.2.2. Different soil sample backgrounds and different measurement (DSDM) conditions

The soil samples from site 1 and site 2 were measured under the same laboratory conditions, while the samples from site 3 were measured under different conditions. We tested the model CNN/PLS-m12 developed in section 3.2 in two testing scenarios, i.e., using 50% (random sampling) of the samples from site 3 without calibration as well as recalibrating the CNN/PLS-m12 model with 50% (random sampling) of 211 samples from site 3. In the two scenarios, the model was tested using the other 50% of the sample from site 3, and thus called CNN/PLS-$m3'_{12}$ and CNN/PLS-$m3_{12}$, respectively. The third model CNN/PLS-m3 was also built using 50% (random sampling) of 211 samples from site 3 and tested it on the other 50% samples. The results from 10 round independent testing of the three procedures were assessed based on the corresponding scatter plots between measured and predicted SM together (Fig. 8). The R^2 value boxplot of independent testing for the three CNN models for 10 rounds is shown in Fig. 9.

The results suggested that a) transferring trained models to another target area without calibration may not be practical as shown by the large negative R^2 (CNN-$m3'_{12}$ was −0.224 and PLS-$m3'_{12}$ was −0.346) and very low RPIQ value (CNN-$m3'_{12}$ was 1.26 and PLS-$m3'_{12}$ was 1.11) in the case that the measurement conditions differ a lot (Fig. 8a and 8b); b) knowledge-based transfer learning for recalibrating the CNN model could help improve the model performance as indicated by the average R^2 values improved from 0.298 (Fig. 8e) to 0.620 (Fig. 8c); c) data-based spiking calibration of the PLS model might even introduce greater deviations because of the different measurement conditions (Fig. 8d vs

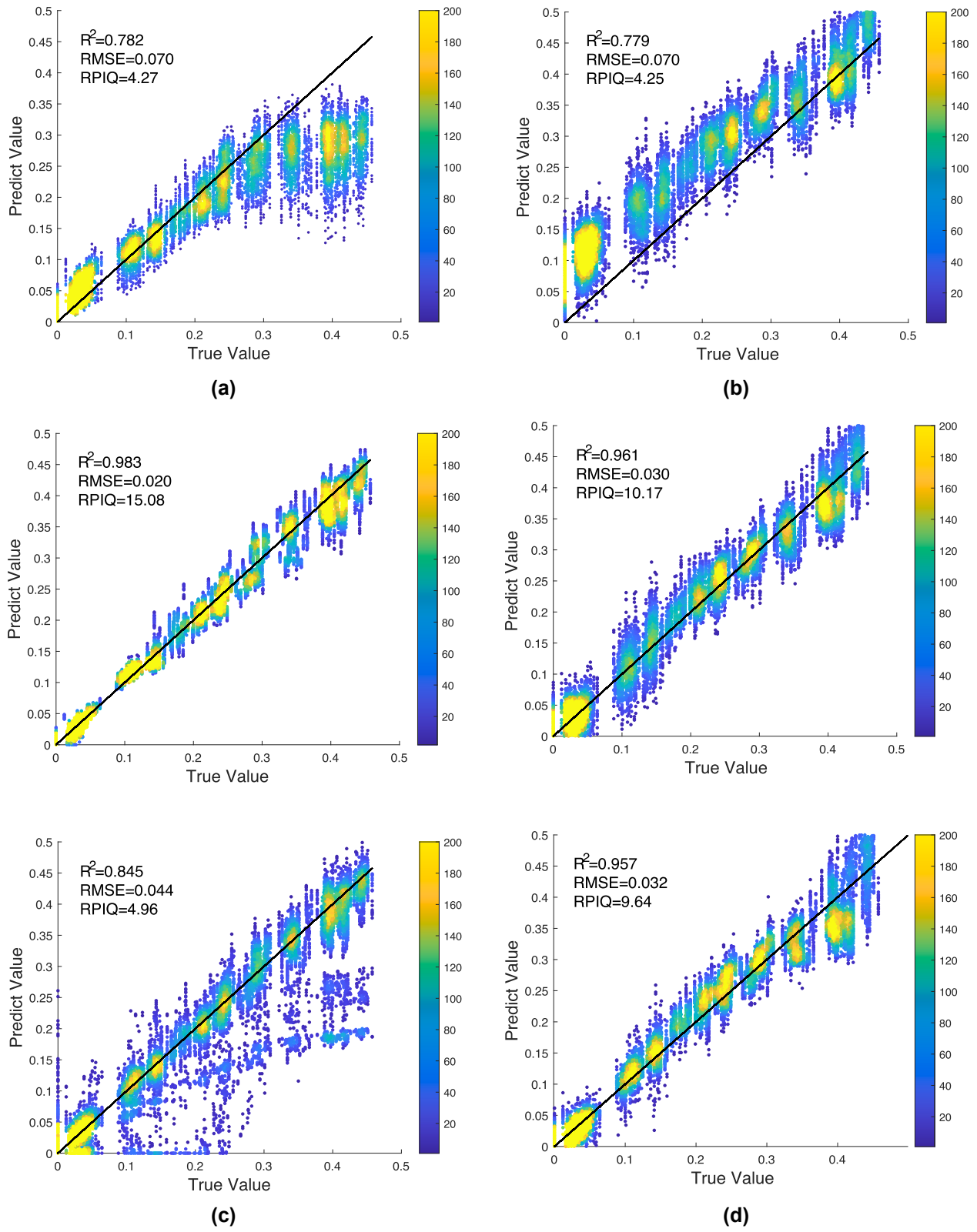


Fig. 6. Scatter plots for measured and predicted SM resulting from 10 round independent model tests (a) CNN-m2′$_1$; (b) PLS-m2′$_1$; (c) CNN-m2$_1$; (d) PLS-m2$_1$; (e) CNN-m2; (f) PLS-m2. The values listed at the upper left corner are the average accuracy for the 10 round tests.

Fig. 8f); d) the same with CNN-m2, a new CNN-m3 model trained using a small dataset was unstable and performed worse than PLS model (Fig. 8e vs 8f and Fig. 9).

4. Discussion

4.1. Model architecture for SM prediction

The CNN model has demonstrated its strong capacity of predicting

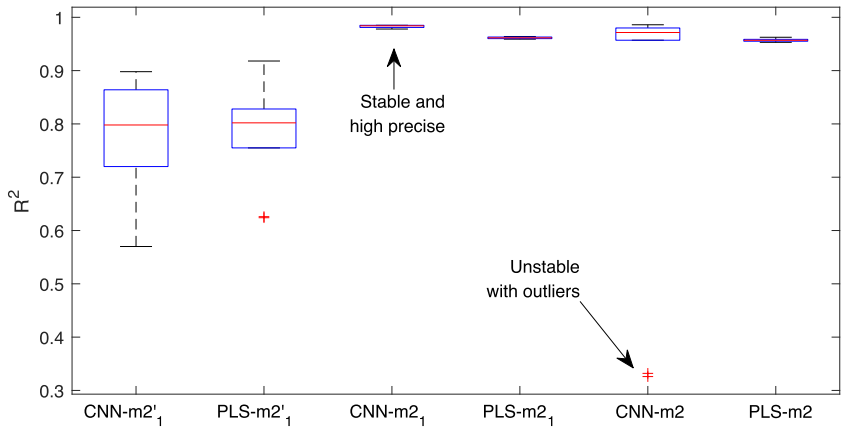


Fig. 7. R^2 value boxplots of independent testing for CNN and PLS for 10 rounds under the DSSM conditions.

soil properties with Vis-NIR spectral data in many recent literatures (Ng et al., 2020; Padarian et al., 2018; Tsakiridis et al., 2020; Wang et al., 2018). The model architectures in these studies were similar, mostly including several convolution layers and max pooling layers in conjunction with several full dense layers and an output layer. The convolution operation is very suitable for extracting valuable absorption peaks/valleys from the spectral data by adjusting the parameters of the convolutional kernels through. However, it is difficult to predetermine the optimal parameters (kernel size, quantity of filters, etc.) in the CNN architecture which can vary among soil samples and depend on soil properties. In this study, a strong performing CNN model became unstable when using a small dataset (170 samples for CNN-m2 and 106 samples for CNN-m3) to train - there were overfitting outliers when training the two models (Figs. 7 and 9). The two models performed even worse than traditional PLS models. We believe that the CNN model architecture could be revised to perform better for the scenarios CNN-m2 and CNN-m3, but that is not conducive to analyzing the performance of transfer learning. The CNN model performed better when trained using a large dataset (2609 samples for CNN-m12). It indicates that a large number of data is required to obtain a strong performing CNN model. In other words, the complexity of the CNN model architecture should be coordinated with the number of samples under a certain soil background. In this study, the convolutional kernel size was set to 3*1 and the first convolutional layers had 16 filters (Fig. 1), resulting in a model having 34,131 trainable parameters and 96 non-trainable parameters. The model performed well for 2,609 soil samples. The model architecture developed by Tsakiridis (2020) has a kernel size of 7*1, and a total number of trainable parameters of 85,373 and performed well for predicting soil properties such as clay, silt, organic carbon of about 20,000 soil samples. This model architecture was also prone to overfitting and poor fitting when trained with the soil samples used in the present study (the results are not shown here). Batch normalization and L2 kernel regularizer ($\lambda 2 = 0.0004$) methods have been used for reducing overfitting, one may rely on this experience and develop a suitable CNN architecture for specific applications.

4.2. Reduced accuracy from DSSM to DSDM

We should realize that the number of soil samples used in this study is limited. A total 40 soil samples from site 1 and site 2 were expanded to 3479 samples for modeling through dividing the samples to two replicates, adding water to produce 15 moisture levels, and recording 3 spectral measurements for each soil sample. Replication upon replication purposely included the variation attributed to incident-exit angle spectral measurements and subsampling variation. In addition, all these soil samples were sieved to grains smaller than 2 mm after a light grinding with mortar and pestle to minimize variation due to grain size. All the soil spectral and moisture measurement were carried in laboratory without in the field. The sample preparation steps for site 1 and site 2 tend to eliminate other interferences such as different soil types or different spectral measurement conditions. This explains why that the accuracy of the model CNN-m12 was as high as 0.989 (average R^2 value), and the R^2 value could still be as high as 0.782 when the model built from site 1 was independently tested on site 2 even without calibration. However, different soil types (composition, particle size, etc.) make CNN model difficult to estimate SM accurately without a sufficient learning. In this study, a 10 round CNN-m3 model was built on 50% samples from a small dataset site 3 (211 samples) and then tested on another 50% samples resulting in a low average R^2 value of 0.298. Therefore, a large soil library with SM content and soil spectra is necessary for knowledge learning to be suitable for more different soil types. Considering that there will be more and more data in the future and the strong knowledge-based transferability of CNN model, it will have broad application prospects.

5. Conclusions

CNN models have advantages for mining the relationship between spectral data and soil properties especially using a large number of samples. More importantly, the CNN model can be potentially generalized to other unknown target areas with a small amount of calibration data. In this study, a 1D-CNN model was built based on soil samples including different compositions and types. The results showed that the knowledge-based transfer learning method could improve the R^2 from 0.845 to 0.983 under the DSSM conditions and the R^2 from 0.298 to 0.620 under the DSDM conditions when transferring a trained CNN model to an unknown target area. Comparing with traditional PLS method, the results showed that the knowledge-based transfer learning method for CNN models has stronger potential for various soil backgrounds or different measurement conditions.

Funding

This research was funded by the National Natural Science Foundation of China (grant No. 41871257 42071312), the Key R&D Program Projects in Hainan Province (grant no. ZDYF2020192) and the Strategic Priority Research Program of the Chinese Academy of Sciences (grant no. XDA19030104; XDA19090121)

CRediT authorship contribution statement

Yu Chen: Conceptualization, Data curation, Methodology, Writing – original draft. **Lin Li:** Methodology, Writing – original draft, Writing – review & editing. **Michael Whiting:** Data curation, Writing – review &

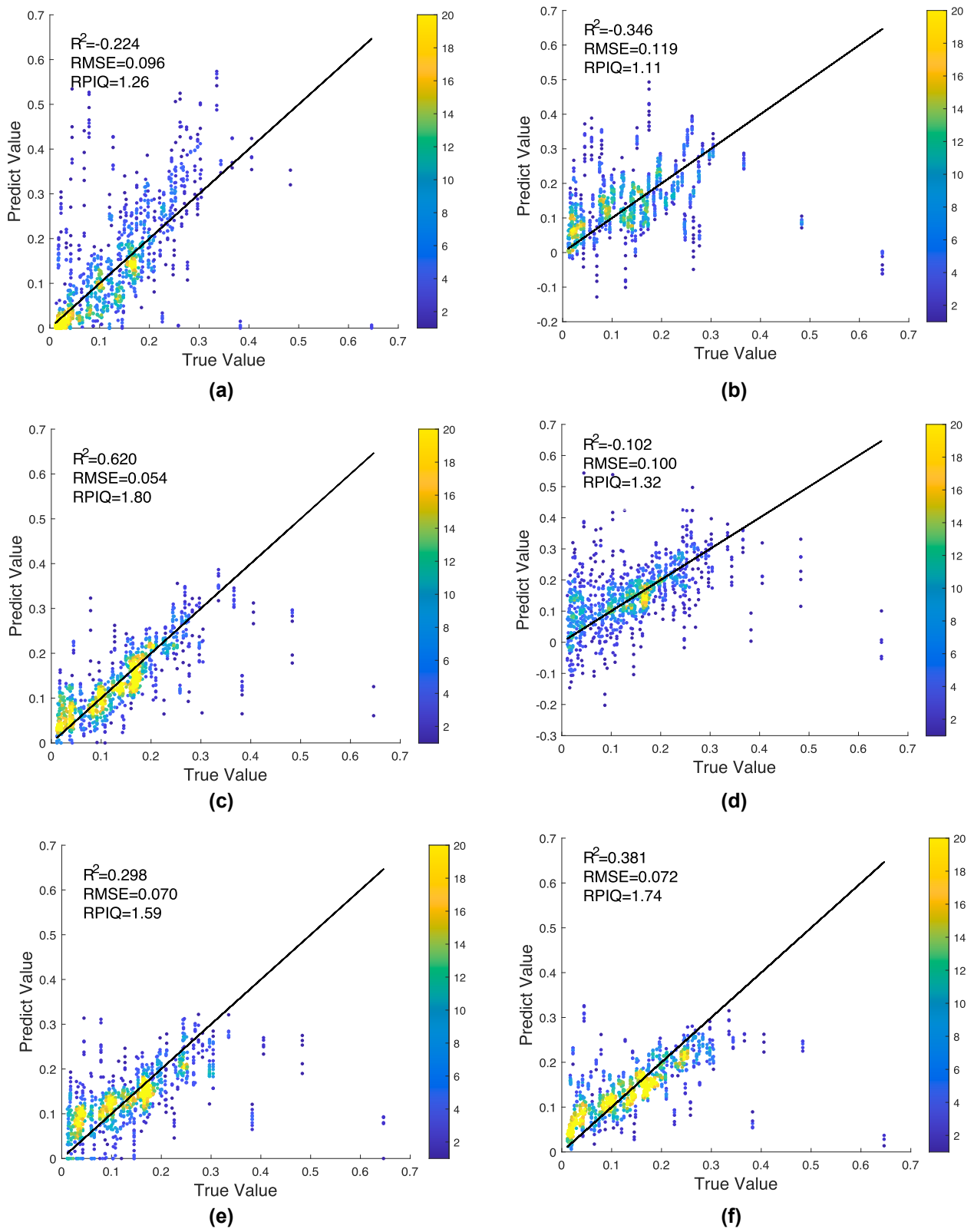


Fig. 8. Scatter plots for measured and predicted SM from 10 round independent model tests: (a) CNN-m3′$_{12}$; (b) PLS-m3′$_{12}$; (c) CNN-m3$_{12}$; (d) PLS-m3$_{12}$; (e) CNN-m3; (f) PLS-m3. The values listed at the upper left corner are the average accuracy for the 10 round tests.

editing. **Fang Chen:** Investigation. **Zhongchang Sun:** Investigation. **Kaishan Song:** Data curation. **Qinjun Wang:** Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

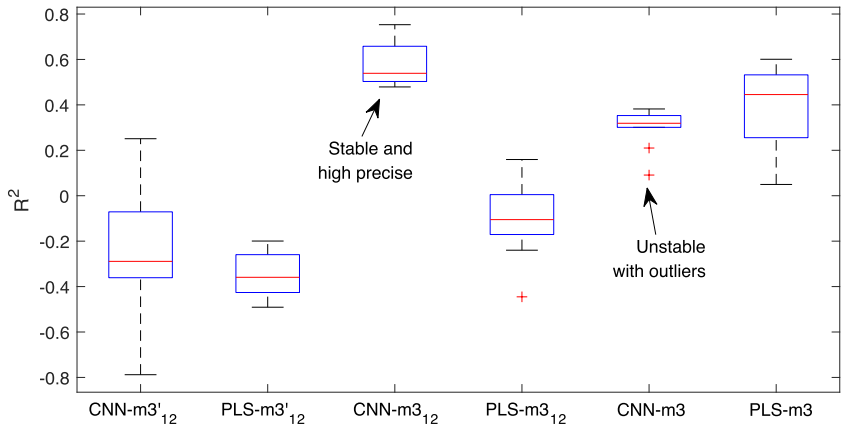


Fig. 9. R^2 value boxplots of independent testing for CNN and PLS for 10 rounds under the DSDM conditions.

the work reported in this paper.

References

Amani, M., Parsian, S., MirMazloumi, S.M., Aieneh, O., 2016. Two new soil moisture indices based on the NIR-red triangle space of Landsat-8 data. Int. J. Appl. Earth Obs. Geoinf. 50, 176–186. https://doi.org/10.1016/j.jag.2016.03.018.

Araujo, S.R., Wetterlind, J., Demattе, J.A.M., Stenberg, B., 2014. Improving the prediction performance of a large tropical vis-NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. Eur. J. Soil Sci. 65 (5), 718–729. https://doi.org/10.1111/ejss.12165.

Bao, Y., Lin, L., Wu, S., Deng, K.A.K., Petropoulos, G.P., 2018. Surface soil moisture retrievals over partially vegetated areas from the synergy of Sentinel-1 and Landsat 8 data using a modified water-cloud model. Int. J. Appl. Earth Obs. Geoinf. 72, 76–85. https://doi.org/10.1016/j.jag.2018.05.026.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. 43 (5), 772–777. https://doi.org/10.1366/0003702894202201.

Birdal, G., Akin, A., and Sarikaya, I., 2007. Soil moisture measurement tools production and development. In S. A. Cetin & I. Hikmet (Eds.), Six International Conference of the Balkan Physical Union (Vol. 899, pp. 764-764).

Carrao, H., Russo, S., Sepulcre-Canto, G., Barbosa, P., 2016. An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data. Int. J. Appl. Earth Obs. Geoinf. 48, 74–84. https://doi.org/10.1016/j.jag.2015.06.011.

Chang, C.W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. Soil Sci. Soc. Am. J. 65 (2), 480–490. https://doi.org/10.2136/sssaj2001.652480x.

Chen, Y., Wei, Y.-M., Wang, Q.-J., Li, L., Lei, S.-H., Lu, C.-Y., 2021. Effects of different spectral resolutions on modeling soil components. Spectroscopy and Spectral Analysis 41 (3), 865–871. https://doi.org/10.3964/j.issn.1000-0593(2021)02-00-06.

Chen, Y., Wei, Y., Wang, Q., Chen, F., Lu, C., Lei, S., 2020. Mapping post-earthquake landslide susceptibility: a u-net like approach. Remote Sens. 12 (17) https://doi.org/10.3390/rs12172767.

Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy - quantitative-analysis techniques for remote-sensing applications. J. Geophys. Res. 89 (NB7), 6329–6340. https://doi.org/10.1029/JB089iB07p06329.

Dejong, S., 1993. Simpls - an alternative approach to partial least-squares regression. Chemometrics and Intelligent Laboratory Systems 18 (3), 251–263. https://doi.org/10.1016/0169-7439(93)85002-x.

Fabre, S., Briottet, X., Lesaignoux, A., 2015. Estimation of soil moisture content from the spectral reflectance of bare soils in the 0.4-2.5 mu m domain. Sensors 15 (2), 3262–3281. https://doi.org/10.3390/s150203262.

Goge, F., Gomez, C., Jolivet, C., Joffre, R., 2014. Which strategy is best to predict soil properties of a local site from a national Vis-NIR database? Geoderma 213, 1–9. https://doi.org/10.1016/j.geoderma.2013.07.016.

González Costa, J.J., Reigosa, M.J., Matías, J.M., Covelo, E.F., 2017. Soil Cd, Cr, Cu, Ni, Pb and Zn sorption and retention models using SVM: Variable selection and competitive model. Sci. Total Environ. 593–594, 508–522. https://doi.org/10.1016/j.scitotenv.2017.03.195.

Guerrero, C., Wetterlind, J., Stenberg, B., Mouazen, A.M., Gabarron-Galeote, M.A., Ruiz-Sinoga, J.D., Rossel, R.A.V., 2016. Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy? Soil Tillage Res. 155, 501–509. https://doi.org/10.1016/j.still.2015.07.008.

Hong, Y., Chen, Y., Zhang, Y., Liu, Y., Liu, Y., Yu, L., Cheng, H., 2018. Transferability of Vis-NIR models for soil organic carbon estimation between two study areas by using spiking. Soil Sci. Soc. Am. J. 82 (5), 1231–1242. https://doi.org/10.2136/sssaj2018.03.0099.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L. D., 1990. Handwritten digit recognition with a back-propagation network. Adv. Neural Inf. Process. Syst. 396–404.

Liu, L.F., Ji, M., Buchroithner, M., 2018. Transfer learning for soil spectroscopy based on convolutional neural networks and its application in soil clay content mapping using hyperspectral imagery. Sensors 18 (9). https://doi.org/10.3390/s18093169.

Lobell, D.B., Asner, G.P., 2002. Moisture effects on soil reflectance. Soil Sci. Soc. Am. J. 66 (3), 722–727.

Norris, K.H., Williams, P.C., 1984. Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat.1. influence of particle-size. Cereal Chem. 61 (2), 158–165.

Padarian, J., Minasny, B., McBratney, A.B., 2019. Transfer learning to localise a continental soil vis-NIR calibration model. Geoderma 340, 279–288. https://doi.org/10.1016/j.geoderma.2019.01.009.

Pang, B., Nijkamp, E., Wu, Y.N., 2020. Deep learning with TensorFlow: A review. J. Educ. Behav. Statistics 45 (2), 227–248. https://doi.org/10.3102/1076998619872761.

Rinnan, A., van den Berg, F., Engelsen, S.B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. Trac-Trends in Analytical Chemistry 28 (10), 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007.

Romero, D.J., Ben-Dor, E., Demattе, J.A.M., Barros e Souza, A., Vicente, L.E., Tavares, T. R., ... Eitelwein, M.T., 2018. Internal soil standard method for the Brazilian soil spectral library: Performance and proximate analysis. Geoderma, 312, 95-103. doi: 10.1016/j.geoderma.2017.09.014.

Sadeghi, M., Babaeian, E., Tuller, M., Jones, S.B., 2017. The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. Remote Sens. Environ. 198, 52–68. https://doi.org/10.1016/j.rse.2017.05.041.

Sanchez, J., Boluda, R., Morell, C., Colomer, J. C., Artigao, A., and Tebar, J. I., 1996. Assessment of soil degradation in desertification threatened areas: A case study in Castilla-La Mancha (Spain). Retrieved from Albacete, Spain: Universidad de Castilla-La Mancha.

Simonyan, K., Zisserman, A., 2015. Very Deep Convolutional Networks for Large-scale Image Recognition. Paper presented at the ICLR.

Tsakiridis, N.L., Keramaris, K.D., Theocharis, J.B., Zalidis, G.C., 2020. Simultaneous prediction of soil properties from VNIR-SWIR spectra using a localized multi-channel 1-D convolutional neural network. Geoderma 367. https://doi.org/10.1016/j.geoderma.2020.114208.

USDA, N., 1978. Soil Survey Kings County, California. Retrieved from Hanford, CA.

Wang, N., Chen, F., Yu, B., Qin, Y., 2020. Segmentation of large-scale remotely sensed images on a Spark platform: A strategy for handling massive image tiles with the MapReduce model. ISPRS J. Photogramm. Remote Sens. 162, 137–147. https://doi.org/10.1016/j.isprsjprs.2020.02.012.

Whiting, M.L., Li, L., Ustin, S.L., 2004. Predicting water content using Gaussian model on soil spectra. Remote Sens. Environ. 89 (4), 535–552. https://doi.org/10.1016/j.rse.2003.11.009.

Wold, S., Martens, H., Wold, H., 1983. The multivariate calibration-problem in chemistry solved by the PLS method. Lect. Notes Math. 973, 286–293.

Yue, J.B., Tian, J., Tian, Q.J., Xu, K.J., Xu, N.X., 2019. Development of soil moisture indices from differences in water absorption between shortwave-infrared bands. ISPRS J. Photogramm. Remote Sens. 154, 216–230. https://doi.org/10.1016/j.isprsjprs.2019.06.012.

Zhang, X.L., Lin, T., Xu, J.F., Luo, X., Ying, Y.B., 2019. DeepSpectra: An end-to-end deep learning approach for quantitative spectral analysis. Anal. Chim. Acta 1058, 48–57. https://doi.org/10.1016/j.aca.2019.01.002.