

Article

# Linear Multi-Task Learning for Predicting Soil Properties Using Field Spectroscopy

Haijun Qi <sup>1,2</sup> , Tarin Paz-Kagan <sup>3</sup>, Arnon Karnieli <sup>2,\*</sup> and Shaowen Li <sup>1,\*</sup>

<sup>1</sup> School of Information and Computer Science, Anhui Agricultural University, Hefei 230036, China; haijun.qi@ahau.edu.cn

<sup>2</sup> The Remote Sensing Laboratory, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boqer Campus 84990, Israel

<sup>3</sup> The Department of Sensing, Information and Mechanization Engineering, Institute of Agricultural Engineering, Agricultural Research Organization (ARO), Volcani Center, Rishon LeZion 7528809, Israel; tarin@volcani.agri.gov.il

\* Correspondence: karnieli@bgu.ac.il (A.K.); shwli@ahau.edu.cn (S.L.); Tel.: +972-8-659-6855 (A.K.); +86-0551-6578-6146 (S.L.)

Received: 18 September 2017; Accepted: 23 October 2017; Published: 30 October 2017

**Abstract:** Field spectroscopy has been suggested to be an efficient method for predicting soil properties using quantitative mathematical models in a rapid and non-destructive manner. Traditional multivariate regression algorithms usually regard the modeling of each soil property as a single task, which means only one response variable is considered as the output during modeling. Therefore, these algorithms are less suitable for the prediction of several key soil properties with low concentrations or unobvious spectral absorption signals. In the current study, we investigated the performance of a linear multi-task learning (LMTL) algorithm based on a regularized dirty model for modeling and predicting several key soil properties using field spectroscopy (350–2500 nm) as an integrated approach. We tested seven key soil properties including available nitrogen (N), phosphorus (P) and potassium (K), pH, water content (WC), organic matter (OM), and electrical conductivity (EC) in drylands. The model performances of LMTL models were compared with the commonly used single-task algorithm of the partial least squares regression (PLS-R). Our results show that the LMTL models outperformed the PLS-R models with the advantage of shared features; the ratio of performance to deviation (RPD) values in the validation set improved by 10.24%, 4.93%, 25.77%, 11.76%, 6.74%, 53.13%, and 3.15% for N, P, K, pH, WC, OM, and EC, respectively. The best prediction was obtained for OM with RPD = 2.29, indicating high accuracy (RPD > 2). The prediction results of N, P, WC, and pH were categorized as of moderate accuracy (1.4 < RPD < 2), while K and EC were categorized as of poor accuracy (RPD < 1.4). However, the explanatory power of the LMTL models was moderate due to fewer features being selected by the regularization algorithm of the LMTL approach, which should be further studied in the soil spectral analysis. Our results highlight the use of LMTL in field spectroscopy analysis that can improve the generalization performance of regression models for predicting soil properties.

**Keywords:** visible-near infrared and shortwave infrared (VNIR/SWIR); soil measurement; dirty model; partial least squares regression (PLS-R); regularization; shared features

## 1. Introduction

The assessment and monitoring key soil properties are important processes for quantifying soil quality and developing tools for soil management in general and precision agriculture in particular. Conventional laboratory methods for detecting soil properties and quality are expensive and time-consuming. An alternative approach, namely reflectance spectroscopy, has been proposed as

a rapid, non-destructive, reproducible, and cost-effective analytical method for assessing soil properties and quality [1]. Field spectroscopy conducts in-situ soil spectral measurements directly and omits the steps of soil sampling and soil pretreatment, making it much faster and more effective than laboratory-based spectroscopy. It is also much more suitable for soil mapping, large-scale monitoring and making real-time predictions [2]. Stenberg et al. [3] stated in their review that visible-near infrared and shortwave infrared (VNIR/SWIR, 350–2500 nm) spectroscopy is a feasible technique for predicting several key soil properties in field conditions with various levels of prediction accuracy. However, under the influence of external factors, such as ambient light, temperature, water content and atmospheric distortions, field soil spectral reflectance would be greatly affected [4]. Moreover, the multi-collinearity and redundancy in the spectral data also limit the model prediction accuracy [5]. One way to improve field spectroscopy accuracy in predicting soil properties is by applying advanced multivariate regression algorithms [3].

The commonly used linear regressions such as principal component regression (PCR) [6] and partial least squares regression (PLS-R) (e.g., [7–14]) can decompose the original spectral matrix through linear combinations to extract useful components and overcome the problems of collinearity with high interpretable ability [15]. In addition to linear algorithms, there is a growing use of non-linear methods by several machine learning algorithms such as the least square support vector machine (LS-SVM) (e.g., [16]), artificial neural networks (ANNs) (e.g., [17–19]), multivariate adaptive regression splines (MARS) (e.g., [20]), the random forest regression (e.g., [21]) and more, which have been proven to enhance the prediction performance based on their excellent non-linear learning ability.

The multivariate regression algorithms mentioned above usually use all the predictor variables (in spectroscopy, we refer to wavebands) to enhance learning ability. However, it is well known that a large number of wavebands are not necessary since some of them are not correlated to the predicted property, and may even contribute to interference during modeling, which can result in overfitting [22]. Feature selection can extract useful information from the spectral signal to reduce model complexity and improve robustness [8,23,24]. However, soil VNIR/SWIR spectroscopy is largely nonspecific due to the overlapping absorption of different soil constituents; thus, extracting specific features for single soil constituents is quite problematic [19,25]. For example, soil water content (WC) (e.g., [10,20]), organic matter (OM) (e.g., [26,27]) and some other high content soil properties could be relatively well predicted using field VNIR/SWIR spectroscopy benefiting from well-recognized spectral absorption signals [19,25]. However, several important soil nutrient properties, such as available nitrogen (N), phosphorus (P), and potassium (K), do not have any obvious specific spectral feature signal, and they usually exist in low concentrations in the soil, especially in dryland soils [16]. Chemical properties, including pH and EC, also have no direct spectral absorption features. Previous studies attributed the occasionally successful predictions of N, P, K, pH, and EC to chemical relations with WC, OM, iron oxides, clay minerals, or other soil properties with sensitive absorption features (e.g., [25,28]).

The current multivariate regression algorithms in soil spectroscopy are usually applied as single-task modeling, for each soil property, although some algorithms such as PLS-R, ANN, etc., could also provide a mode with multiple response variables. This means that only one dependent variable is considered as the output during the development of a regression model with a soil spectral matrix as the input independent variables. Since there are relations between the modeling of several soil properties, it can be advantageous to build all of the regression models simultaneously using the approach of multi-task learning instead of following the more traditional approach of learning each single task independently from the others [29]. Considering the underlying cross-relatedness between different dependent variables, multi-task learning aims to improve the generalization by using the shared information contained in the training data of related tasks [30]. This approach is particularly efficient for high-dimensional data [31], such as high-resolution spectroscopy. It has recently received increasing attention in machine learning, artificial intelligence, and computer vision [32,33]. Daniel et al. [19] proposed an ANN model with OM, P, and K as simultaneous outputs and attained satisfactory results; however, in their research, the underlying correlation was not fully interpreted.

The overarching aim of the study was to explore the performance of (linear multi-task learning) LMTL algorithms based on a regularized dirty model (composed of shared and non-shared features) for modeling and predicting several key soil properties using field spectroscopy (350–2500 nm). The specific objectives are three-fold: (1) investigate the predictive ability of field spectroscopy to predict soil N, P, K, WC, pH, EC, and OM in dryland regions; (2) compare the prediction performance of LMTL via a regularized dirty model with traditional single-task learning via the PLS-R; and (3) study the shared features between different soil properties in LMTL as compared to single-task learning.

## 2. Materials and Methods

### 2.1. Study Area

The soil samples were collected from four small watersheds located in the central Negev Highlands of Israel (30°54'N, 34°49'E). The mean annual rainfall in this area is 95 mm and is limited to the winter season, with a high annual variability (ranging from 20–180 mm). The study site included watersheds containing runoff harvesting systems (RHSs), which are used for increasing agricultural production or for developing afforestation systems in drylands. RHSs are designed to collect runoff water and nutrients from small rocky watersheds into ponds bounded by soil dikes (termed *limans*) that are used as afforestation groves. Geologically, the area is composed of limestone and chalk of the Turonian age. The hillslopes are relatively steep (up to 29 degrees) and subdivided into two distinct sections: (1) the upper parts are mainly barren, with steep limestone rocky outcrops and shallow patches of soil cover; and (2) the lower parts consist of colluvium embedded with unconsolidated rocks [34,35]. A similar subdivision is also observed along the channels. The upper part of the channel is rocky while the lower part is covered with an alluvial fill [34]. Lithology across the study site consists of limestone dominantly covering the area, frequently mixed with dolomite, chalk and marl. The stream channels characterized by loessial soils, and are composed of clay, silt, and gravel alluvial soil. In general, the RHSs are located in the downstream area of the watershed where there is a relatively high volume of alluvial soil. In the current study, the differences in soil properties at different locations resulted from the proximity to the RHSs (upstream, downstream, and RHS).

### 2.2. Soil Field Spectroscopy Measurements

The soil spectra were acquired under field conditions (undisturbed samples) before soil sampling with the portable analytical spectral device (ASD) Field Spec<sup>®</sup> Pro spectrometer, with a 25° field of view. During the field campaigns, the skies were clear, and the soil spectral samples were taken on exposed bare soil. The field spectral measurement was vertical in relation to the soil surface using a bare fiber with a height interval of about 1 m, and exposed to the sun as illumination. The instrument was repeatedly calibrated to spectral reflectance using a standard white reference panel (Spectralon Labsphere Inc., North Sutton, NH, USA). To reduce spectral noise and the effects of micro-topography shadowing, four spectral readings for each soil sample were measured and averaged to a final value representing the field sample. The ASD covers a spectral range of 350–2500 nm with a spectral resolution that varies from 3 nm in the VIS-NIR range to 8–10 nm in the SWIR range. We resampled the ASD's spectral band to 5 nm uniformly along the entire spectral region.

### 2.3. Soil Sampling and Physicochemical Lab Analysis

The soil sampling included 10 replicates in each sampling location in the watershed in the upstream, downstream and RHS sites ( $n = 30$  replicates in a watershed) with a total of 120 soil samples collected in September 2015, which is at the end of the dry season, at a depth of 0–0.15 m. All soil samples were transferred to the laboratory and were stored unopened at room temperature until they were analyzed. Soil was air dried, passed through a 2-mm sieve, and analyzed for soil physicochemical properties. The soil properties included the following: N was measured by potassium chloride extractions [36]; P was measured by the Olsen method [37]; K was measured by a flame

photometer [38]; WC was measured by drying the soil in 105 °C; pH was measured in a saturation paste using a handheld portable probe; EC was measured in the extracts from the saturation paste by a handheld portable probe; and OM was measured by drying the soil for two hours at 500 °C. The results of the soil physiochemical properties were tested for their statistical variation. The outliers were determined by boxplot [39], and 12 samples with extremely large N or K concentrations beyond the upper whisker were removed. Thus, the remaining 108 samples were used for further study in this work.

#### 2.4. Spectral Preprocessing and Transformations

We removed the low signal-to-noise ratio wavebands at both ends and the atmospheric water absorption wavebands ranging from 1350–1420 and 1800–1960 nm. The resultant reflectance spectrum of 355 wavebands (400–2400 nm) was henceforth used. In addition, several commonly used preprocessing methods and transformations were sequentially applied to the soil spectral reflectance in this study, including: the Savitzky–Golay filtering algorithm (SG) [40] with a second-order polynomial that was selected to smooth spectral reflectance; a standard normal variate (SNV) [41,42] that was performed to correct additive and multiplicative effects; and a first derivative (FD) that was conducted to remove the baseline and improve the linear trend [43,44].

#### 2.5. Learning Algorithms

To compare the performance of multi-task and single-task algorithms, the regression models were built upon the same dataset. As a commonly used single-task algorithm, the PLS-R is particularly useful for predicting a set of dependent variables from a large set of independent variables [45]. To overcome the problem of collinearity between predictors, the PLS-R decomposed independent variables and dependent variables by linear combinations to extract latent variables (LVs, or components) and built the regression model based on the LVs instead of the original training variables [15]. To avoid overfitting or underfitting, a leave-one-out cross-validation was used to determine the number of LVs with the smallest mean squared error in calibration. The variable importance for projection (VIP) scores [46] obtained by PLS-R has been recognized as a useful measure to identify important wavelengths when the score is more than 1 [47,48].

Regularization is commonly used as a shrinkage method in least square linear regression modeling to avoid overfitting [49]. The main idea of multi-task learning in multiple linear regressions is to take advantage of the shared feature structure (block-sparse) between each task and model all the tasks simultaneously with  $L_1/L_q$  norm block-sparse regularization, particularly with  $L_1/L_\infty$  [50,51]. Since non-shared features might have existed for several specific tasks, the regression coefficient matrix  $W$  (features  $\times$  tasks) cannot fall cleanly into any one structural bracket. To overcome this, a dirty model was proposed to decompose the matrix  $W$  into a block-sparse matrix  $W_b$  (corresponding to the shared features) and an elementwise sparse matrix  $W_e$  (corresponding to the non-shared features); details can be found in [31]. The object function is:

$$\min_W \sum_{i=1}^t \| W_i^T X_i - Y_i \|_2^2 + \lambda_b \| W_b \|_{1,\infty} + \lambda_e \| W_e \|_1 \quad (1)$$

$$\text{subject to : } W = W_b + W_e \quad (2)$$

where  $X_i$  is the spectral matrix of task  $i$  that is the same for each task,  $Y_i$  is the predicted variable of task  $i$ , and  $\lambda_b$  and  $\lambda_e$  are regularization parameters to control the degree of penalty on  $W_b$  and  $W_e$ , respectively.

For the calibration and validation of the regression models (multi-task learning and single-task learning), 108 soil samples were divided into two parts with a split ratio of 0.7 to 0.3, respectively, based on the Kennard–Stone algorithm [52], conducted on the preprocessed spectral matrix. Thus, 76 samples were selected as the calibration set (also used for cross-validation during training) and the

remaining 32 as the validation set (independent testing for the established model). To avoid attributes in the higher numerical ranges dominating those in the lower numerical ranges, both the calibration and validation sets were standardized by mapping their mean and standard deviations to 0 and 1, respectively, before calibration and validation [53].

### 2.6. Accuracy Comparison

In this study, the prediction accuracy of the regression models was validated and compared with the ratio of performance to deviation (RPD) of the validation set that was calculated as:

$$\text{RPD} = \text{SD}/\text{RMSE} = \sqrt{m \sum_{i=1}^m (y_i - \bar{y})^2 / \sum_{i=1}^m (f(X_i) - y_i)^2} / (m - 1) \quad (3)$$

where  $m$  is the number of testing samples in the validation set,  $y_i$  is the real value of sample  $i$ ,  $f(X_i)$  is the predicted value of sample  $i$ , and  $\bar{y}$  is the average value of  $y$ . According to [54], the following three categories of predictability were adopted: Category A (RPD > 2.0) with good accuracy; Category B (1.4 < RPD < 2.0) with moderate accuracy; and Category C (RPD < 1.4) with poor accuracy. The ratio between the interpretable sum squared deviation and the real sum squared deviation (SSR/SST), which was calculated as:

$$\text{SSR}/\text{SST} = \sum_{i=1}^m (f(X_i) - \bar{y})^2 / \sum_{i=1}^m (y_i - \bar{y})^2 \quad (4)$$

was recognized as the proportion of the variability of the dependent variable explained by the regression model [55–59]. A good model should have both high RPD and SSR/SST [13]. Usually, the SSR/SST should be greater than 0.5 to ensure the model stability. All mathematical analysis methods mentioned above were conducted in MATLAB (MathWorks, Natick, MA, USA). The process of single-task learning and multi-task learning were carried out with MALSAR Version 1.1 [60].

## 3. Results

### 3.1. Soil Properties and Spectral Response

Table 1 shows the descriptive statistics of the measured concentrations of the seven soil properties in the 108 samples used for calibrating and validating the regression models. The results showed high variation in the soil properties, indicating that the data could be used for the regression analysis. Table 2 shows the Pearson correlation coefficients (R) between different soil properties. The highest positive correlation was found between OM and WC (R = 0.74). Moderately positive correlations were found between N and P (R = 0.69), N and K (R = 0.58), P and K (R = 0.57), and OM and K (R = 0.51). In addition, it was found that pH has a low negative correlation with every other soil property.

**Table 1.** Descriptive statistics of the seven studied soil properties.

Soil Properties	Units	Mean	STD	Min	Median	Max
N	mg/Kg	27.42	15.57	5.73	22.66	72.56
P	mg/Kg	19.36	9.70	7.00	16.70	55.20
K	mg/Kg	25.44	13.49	7.30	23.45	64.80
WC	%	5.61	1.81	3.01	4.92	10.77
pH		7.20	0.21	6.78	7.14	7.87
EC	μS/cm	0.50	0.22	0.18	0.46	1.77
OM	%	4.29	1.92	1.73	3.68	9.58

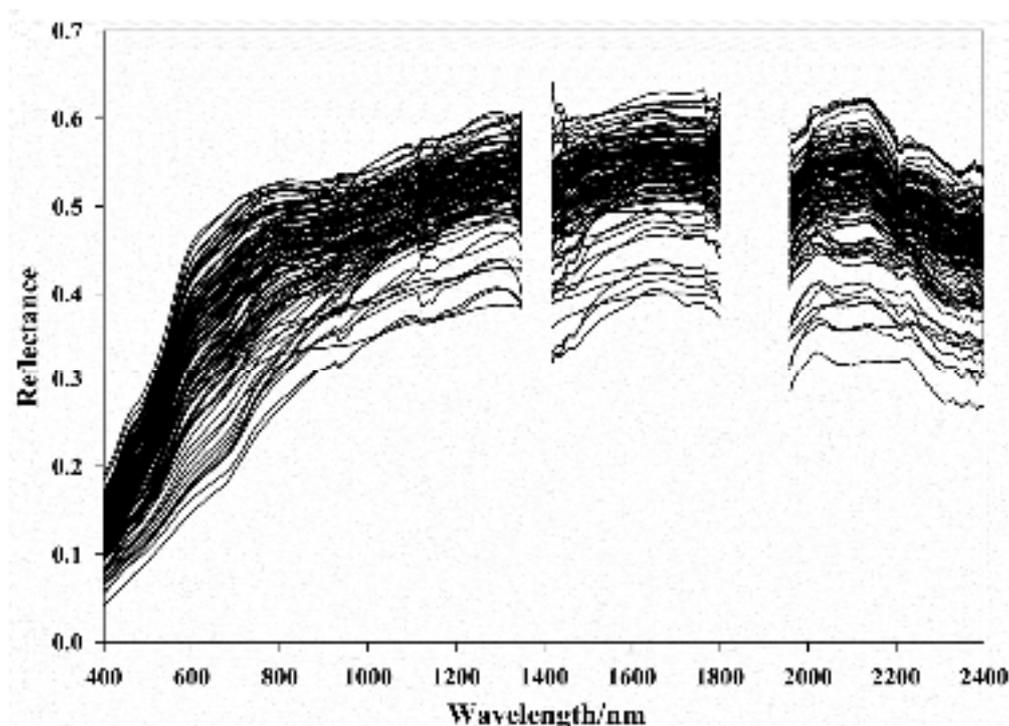
Abbreviations used: available nitrogen (N), available phosphorous (P), available potassium (K), water content (WC), pH, electrical conductivity (EC), and organic matter (OM).

**Table 2.** Pearson correlation coefficients between the seven soil properties.

Soil Properties	N	P	K	WC	pH	EC	OM
N	1.00						
P	<b>0.69</b>	1.00					
K	<b>0.58</b>	<b>0.57</b>	1.00				
WC	0.24	0.15	0.22	1.00			
pH	−0.25	−0.27	−0.24	−0.30	1.00		
EC	0.30	0.19	0.28	0.25	−0.38	1.00	
OM	0.45	0.39	<b>0.51</b>	<b>0.74</b>	−0.26	0.31	1.00

Abbreviations used: available nitrogen (N), available phosphorous (P), available potassium (K), water content (WC), pH, electrical conductivity (EC), and organic matter (OM).

Figure 1 shows the field spectral reflectance of 108 soil samples. The spectra show large variations between different samples that are caused by soil color, soil composition, water content, particle size, and the coverings on the soil surface, such as residual dry vegetation, rock particles, and mineral deposits [61]. The obvious spectral signatures near 930 nm may be related to the absorption of hydroxyl in Fe oxides [25], the ones near 940 nm, 1150 nm and 1450 nm may be influenced by the absorption of the atmospheric water content [62]; the ones near 1765 nm may be related to the signal jump of the spectral instrument; the ones near 2205 nm may be related to the absorption of Al-OH [25]. The spectral signatures ranging from 2300–2400 nm may be related to several other clay minerals and soil organics [63].



**Figure 1.** Field spectra excluding the atmospheric water absorption wavebands and the low signal-to-noise ratio wavebands.

### 3.2. Model Performance of PLS-R

#### 3.2.1. Prediction Results

Table 3 shows the prediction results of the PLS-R models for the seven soil properties. The numbers of LVs were determined by leave-one-out cross-validations, which are also shown in Table 3. Among all

soil properties, OM was the most accurately predicted with a RPD = 2.22, and with a model prediction accuracy categorized as A (RPD > 2). The prediction accuracies of P, WC, and pH were categorized as B ( $1.4 < \text{RPD} < 2$ ) with RPD = 1.42, 1.53, and 1.78, respectively, indicating moderate prediction ability. All properties in Category A or B achieved SSR/SST values that were more than 0.5, confirming the model stability for predicting P, WC, pH, and OM. However, the prediction accuracies of N, K, and EC were categorized as C (RPD < 1.4) with RPD = 1.27, 0.97, and 0.64, respectively, demonstrating poor prediction ability. In addition, it is worth mentioning that the SSR/SST of N was 0.66, but the SSR/SSTs of K and EC were over 1 and related to low prediction accuracy and overfitting of the models.

**Table 3.** The used parameters and prediction results of the partial least squares regression (PLS-R) and linear multi-task learning (LMTL) models for predicting seven soil properties.

Algorithm	Property	Parameter <sup>1</sup>		<i>n</i> <sup>2</sup>	Calibration		Validation		Accuracy Category
					RPD	SSR/SST	RPD	SSR/SST	
PLS-R	N	5		355	2.15	0.78	1.27	0.66	C
	P	5		355	1.85	0.71	1.42	0.80	B
	K	6		355	2.71	0.86	0.97	-	C
	WC	4		355	1.99	0.75	1.53	0.72	B
	pH	6		355	2.78	0.87	1.78	0.83	B
	EC	6		355	2.33	0.81	0.64	-	C
	OM	5		355	2.58	0.85	2.22	0.82	A
LMTL	N	40	20	81	1.94	0.53	1.40	0.58	B
	P	20	21	114	2.18	0.54	1.49	0.64	B
	K	160	26	11	1.56	0.29	1.22	0.52	C
	WC	30	7	75	2.30	0.61	1.71	0.55	B
	pH	20	3	79	3.45	0.76	1.90	0.92	B
	EC	60	20	66	1.42	0.21	0.98	-	C
	OM	40	25	75	2.31	0.68	2.29	0.70	A

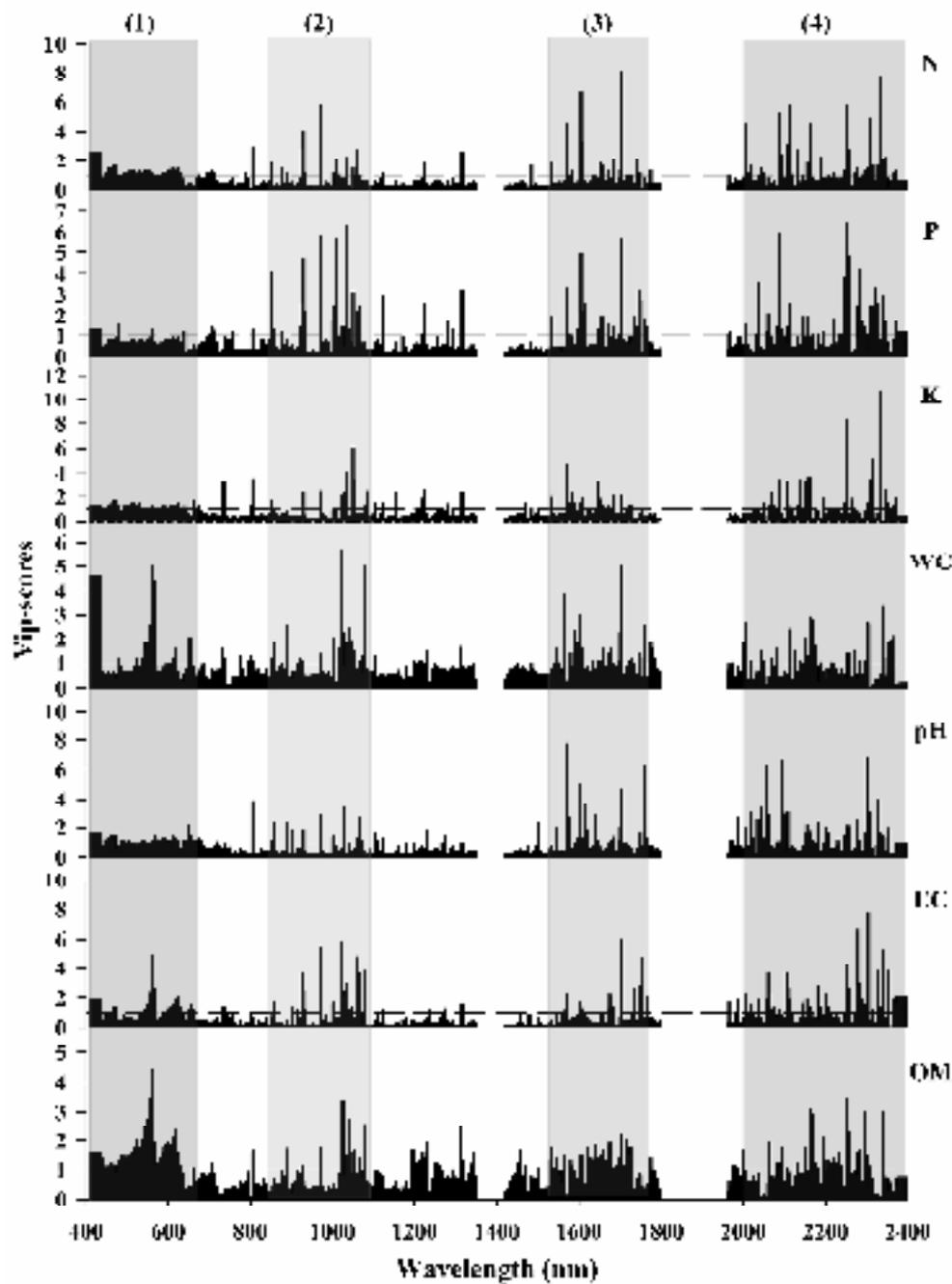
<sup>1</sup> Note: The parameter for PLS-R is the number of latent variables; for LMTL, there are two regularization parameters,  $\lambda_b$  and  $\lambda_e$ . <sup>2</sup> Note: *n* is the number of features used in the model. Category A: RPD > 2.0, Category B:  $1.4 < \text{RPD} < 2.0$ , Category C: RPD < 1.4. Abbreviations used: available nitrogen (N), available phosphorous (P), available potassium (K), water content (WC), pH, electrical conductivity (EC), and organic matter (OM), the ratio of performance to deviation (RPD), the ratio between the interpretable sum squared deviation and the real sum squared deviation (SSR/SST).

### 3.2.2. Feature Importance in PLS-R

Figure 2 shows the distribution of the VIP scores of different soil properties over the entire wavelength range (400–2400 nm). We used all of the 355 wavebands in the VNIR/SWIR region in the PLS-R models, which was necessary to evaluate the feature importance in the prediction of each soil property. We identified four feature-block regions with high importance existing in the entire VNIR/SWIR region whose properties were associated with soil OM, WC, clay minerals and Fe oxides [64]. The first region, ranging from 410 to 650 nm, is mostly related to the Fe oxides [19,25], with the 560 nm waveband related to OM [24,65]. The second region ranges from 850 to 1075 nm; of this, the 850–930 nm range is related to hydroxyl in Fe oxides [19,25], the 970 nm waveband is related to the soil water absorption waveband [66], the 1010 nm waveband is a hydrate-related absorption feature [67], and the 1025–1075 nm range is mostly related to the electronic transition bands of Fe<sup>2+</sup> or Fe<sup>3+</sup> [19]. The third region, ranging from 1530 to 1770 nm, is almost entirely related to soil organics [19,25]. The fourth region, ranging from 2005 to 2400 nm, may be connected to water, organics and clay minerals [19,25,68].

Besides the four feature-block regions, several important individual wavebands with significantly high VIP scores still could be seen in Figure 2. The 730 nm waveband in the prediction of K, WC and EC may be related to the sensitive absorption of soil salinity [67]. The 805 nm waveband is related to the red-edge, which is known to be sensitive to biomass [69], and the predictions of P, K, pH and OM were found to be correlated. The 1120 and 1155 nm wavebands may be related to the *v* S–O stretching bands of sulfate [70], which can be used to predict P and K. The 1225 nm [71] and 1315 nm [72] wavebands

were found to be important bands for predicting soil organics; here, several properties are correlated. The 1450 nm waveband in the prediction of OM may be related to the carboxylic acids in organics [25].

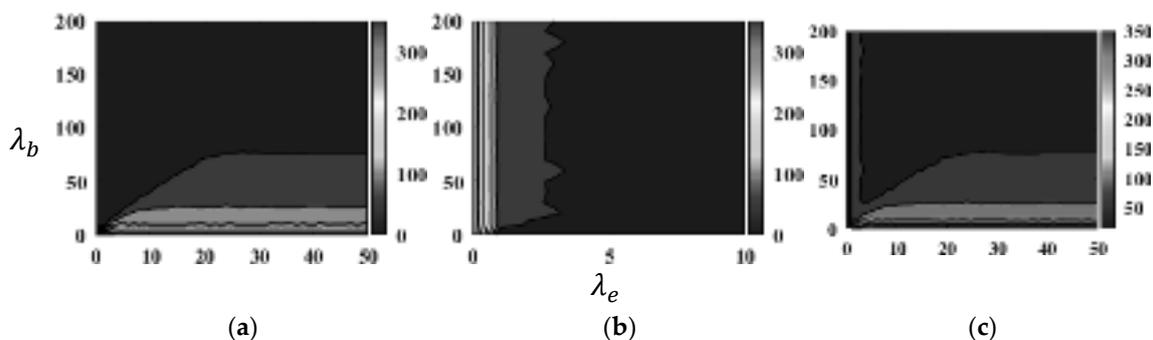


**Figure 2.** Variable importance for projection (VIP) scores of the PLS-R models for predicting available nitrogen (N), available phosphorous (P), available potassium (K), water content (WC), pH, electrical conductivity (EC), and organic matter (OM). Dashed line:  $y = 1$ . (1), (2), (3), and (4) represent four feature-block regions.

### 3.3. Model Performance of LMTL

#### 3.3.1. Effects of Regularization Parameters on Modeling

The range of  $\lambda_b$  was set to 0–200 with a gradient of 10, and the range of  $\lambda_e$  was set to 0–50 with a gradient of 1. Figure 3 shows the sparsity (the number of non-zero elements in the regression coefficients) of the block-sparse matrix ( $W_b$ ), the elementwise sparse matrix ( $W_e$ ), and the combined regression coefficients matrix ( $W$ ) of the regression model generated from LMTL for predicting N when changing  $\lambda_b$  and  $\lambda_e$ ; other properties had similar characteristics (Supplementary Material, Figure S1) and can be seen in the Appendix. According to Figure 3, the sparsity of  $W_b$  decreased significantly with increasing  $\lambda_b$ , especially in the low  $\lambda_b$  values ranging from 0 to 75. The sparsity of  $W_e$  decreased significantly with increasing  $\lambda_e$ , especially in the low  $\lambda_e$  values ranging from 0 to 3.5. However, the sparsity of  $W_b$  was also affected by  $\lambda_e$  in the range of 0–20, and the sparsity of  $W_e$  was also affected by  $\lambda_b$  in the range of 0–10. Therefore,  $\lambda_b$  and  $\lambda_e$  controlled the degree of penalty on  $W_b$  and  $W_e$ , respectively, but also interacted with each other. Both regularization parameters greatly affected the sparsity of  $W$ . The degree of effects gradually weakened with the increase of  $\lambda_b$  or  $\lambda_e$ .



**Figure 3.** Sparsity (the number of non-zero elements) of the block-sparse matrix  $W_b$  (a); the elementwise sparse matrix  $W_e$  (b); and the combined regression coefficients matrix  $W$  (c) of the model generated from linear multi-task learning for predicting available nitrogen.

Figure 4 shows the RPD and SSR/SST performance of the LMTL prediction models for predicting the seven soil properties of the validation set when changing  $\lambda_b$  and  $\lambda_e$ . According to Figure 4, the SSR/SSTs of all the properties showed an obviously decreasing trend with increasing  $\lambda_b$  and  $\lambda_e$ , which is similar to the changing characteristics of model sparsity shown in Figure 3c. The RPD performances generally increased with  $\lambda_b$  and  $\lambda_e$  in the low-value range and decreased in the high-value range. The aim of the object function of LMTL is to minimize the overall squared error of the seven soil properties (see Equation (2)). Thus, obtaining the individual maximum RPD for every soil property simultaneously is a difficult task. The locations of the high RPDs and the decreasing trends of N, P, and OM are similar and are dependent on  $\lambda_b$ , representing the shared features. The locations of the high RPDs of WC, pH and EC are determined by both  $\lambda_b$  and  $\lambda_e$ , representing both the shared and non-shared features. The location of the high RPD of K is almost on the border of the regularization parameters, indicating that few features are used. This means that the best prediction of different properties may depend on the special combination of shared and non-shared features; thus, different combinations of  $\lambda_b$  and  $\lambda_e$  were selected to obtain the best RPDs for different properties under the condition of SSR/SST values higher than 0.5 (Table 3).

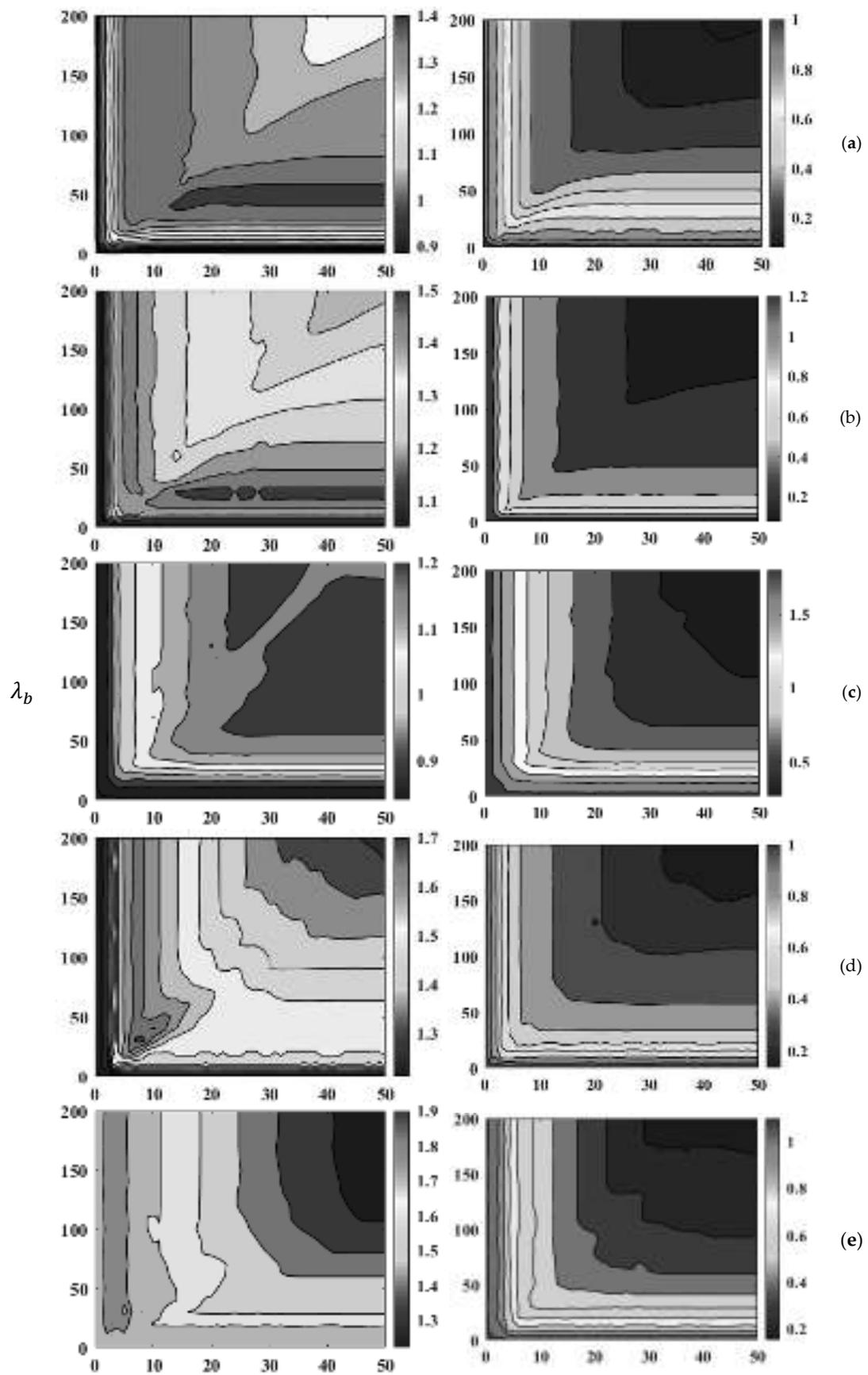
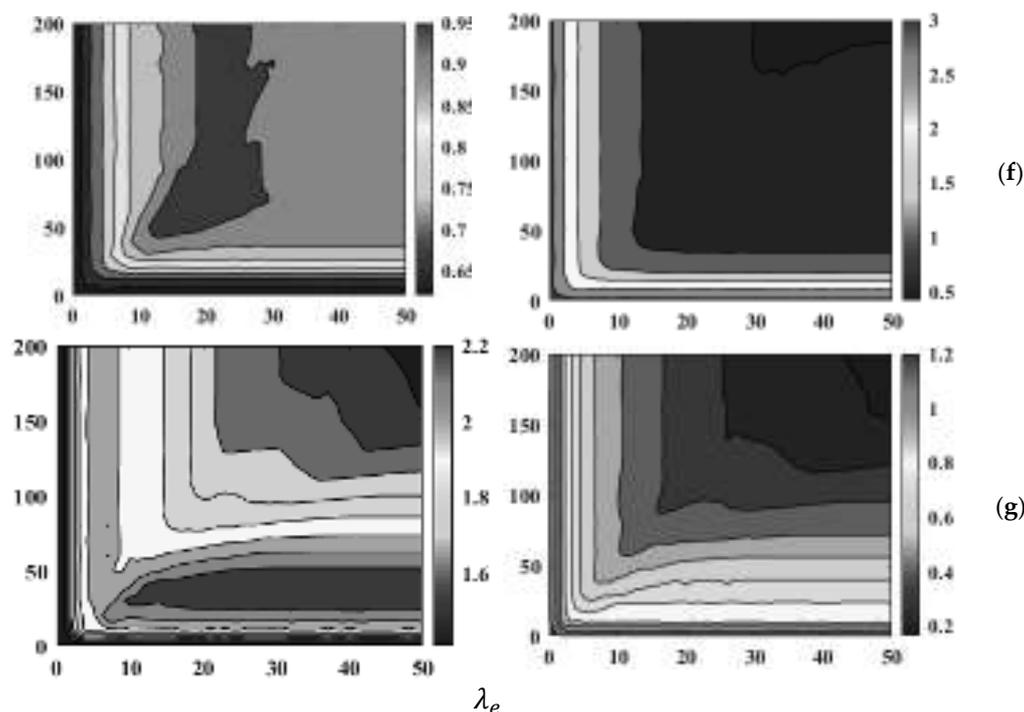


Figure 4. Cont.



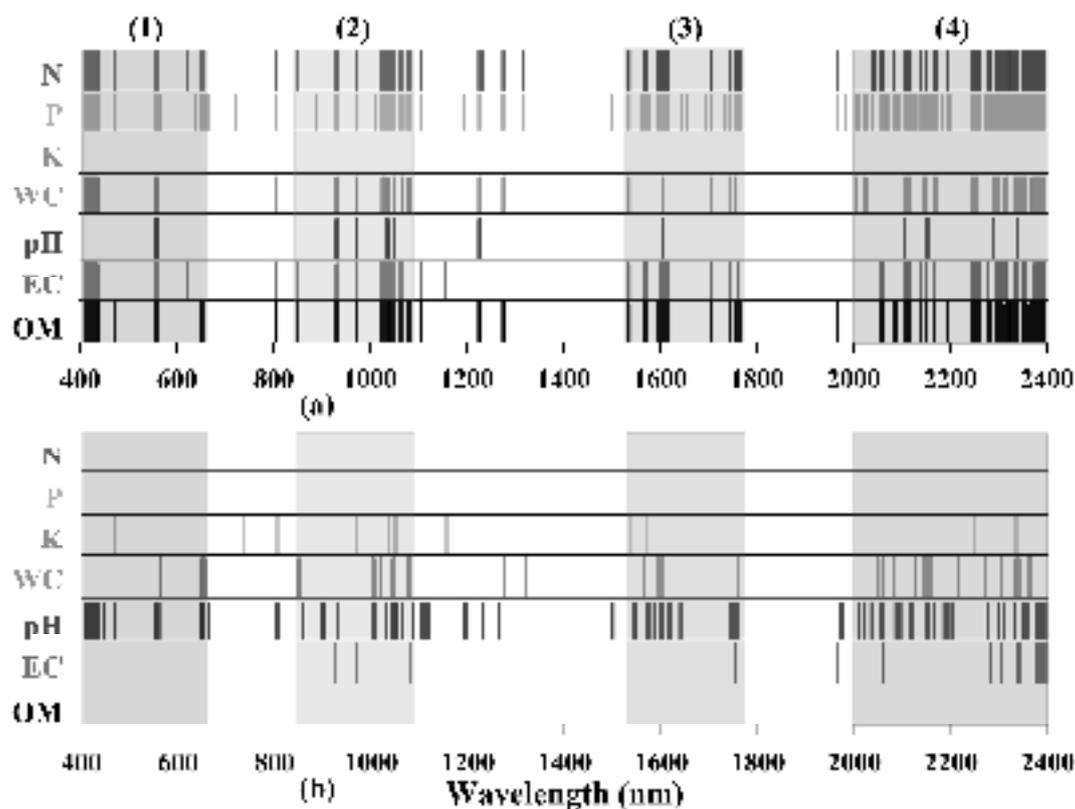
**Figure 4.** The ratio of performance to deviation (RPD) (left) and the ratio between the interpretable sum squared deviation and the real sum squared deviation (SSR/SST) (right) performance of linear multi-task learning prediction models for predicting: available nitrogen (a); available phosphorous (b); available potassium (c); water content (d); pH (e); electrical conductivity (f); and organic matter (g) in a validation set when changing the regularization parameters ( $\lambda_b$  and  $\lambda_e$ ).

### 3.3.2. Prediction Results and Used Features

Table 3 shows the selected regularization parameters, the numbers of used features and the prediction results of LMTL models for predicting seven soil properties. Figure 5 shows the distributions of the features that were used in the models with different combinations of  $\lambda_b$  and  $\lambda_e$  (The distributions of the features with specific  $\lambda_b = 40$  and  $\lambda_e = 10$  could be seen in Supplementary Material, Figure S2). The regularization algorithm reduced the numbers of used features to a large extent and selected a few wavebands as optimal features, which made the distributions of the used features sparse. Nevertheless, the RPD performances of all the soil properties increased compared to the single-task models built by PLS-R (Table 3): N increased to 1.40 from 1.27; P increased to 1.49 from 1.42; K increased to 1.22 from 0.97; WC increased to 1.71 from 1.53; pH increased to 1.90 from 1.78; EC increased to 0.98 from 0.64; OM increased to 2.29 from 2.22. Consequently, the prediction accuracy of OM was categorized as A; Category B included pH, WC, P, and N; and K and EC were categorized as C. However, the SSR/SSTs of all properties except P, decreased due to regularization (Table 3). Despite this, the SSR/SST results were still more than 0.5, which can confirm the model stability, except EC which had an inappropriately large value.

According to Figure 5, the used features of N, P, and OM were all in the block-sparse matrix, indicating shared features. Some of the used features of WC, pH, and EC were in the block-sparse matrix, and others were in the elementwise sparse matrix, indicating that both shared features and non-shared features were used for modeling. The used features of K were all in the elementwise sparse matrix, indicating non-shared features. These results agree with the analysis results in Figure 4. Shared features can be seen in Figure 5a, with most of them in the four feature-block regions. Out of these four regions, important features, such as 805, 1225, and 1315 nm, with large VIP scores that were used in the PLS-R were also recognized as shared features. Some less important features with small VIP scores, such as 1105, 1275, and 1965 nm (Figure 2), were also shared to provide more information.

The distribution positions of the non-shared features used by K, WC, pH, and EC are shown in Figure 5b and are similar to the shared features in Figure 5a. These indicated that the predictions of the seven soil properties were correlated. However, because different regularization degrees of  $W_b$  and  $W_e$  were implemented for different models, the features were divided into either the block-sparse matrix or the elementwise sparse matrix.



**Figure 5.** Used features (non-zero items in the transpose of block-sparse matrix  $W_b$  (a); and elementwise sparse matrix  $W_e$  (b)) of the linear multi-task learning models for predicting available nitrogen (N), available phosphorous (P), available potassium (K), water content (WC), pH, electrical conductivity (EC), and organic matter (OM), respectively. (1), (2), (3), and (4) represent four feature-block regions.

#### 4. Discussion

We investigated the performances of the LMTL algorithm for modeling several soil properties simultaneously using field VNIR/SWIR spectroscopy. To the best of our knowledge, most studies dealing with quantitative soil spectroscopy analysis focus on building regression models via individual single-task learning algorithms with only one response variable as the output. PLS-R could also work in a multiple-response mode, but features could not be shared during modeling. In this study, we used PLS-R to represent single-task learning algorithms for comparing the performance of LMTL. Our results show that the use of LMTL algorithms with multiple response variables as output improves the RPD in all of the seven tested soil properties. We found that shared features can be used to improve the generalization performance of regression models for predicting these seven key soil properties. In addition, we found that low concentration soil properties such as K and EC, with few spectral absorption signals are usually difficult to predict with current optical methods.

##### 4.1. Comparison of Two Algorithms

The PLS-R models explained most of the variance in the dependent variable with several latent variables obtained from the spectral matrix, and they recognized the important wavebands for each

soil property efficiently. Therefore, the SSR/SST values of the PLS-R models were very high, indicating strong explanatory power (Table 2). However, the low RPD values in the validation set indicating prediction accuracy were not satisfactory due to the large number of used features and the high complexity of the PLS-R models. In comparison, the LMTL based on the dirty model algorithm regularized the overall wavebands and built the regression models with selected shared and non-shared features (Figure 5), which led to certain degrees of improvement of the RPD values in the validation set, indicating higher prediction accuracy and stronger generalization performance. We found that the RPD values improved by 10.24%, 4.93%, 25.77%, 11.76%, 6.74%, 53.13%, and 3.15% for N, P, K, pH, WC, OM, and EC, respectively. It is worth mentioning that the prediction accuracy category of N improved from C to B, and the prediction ability of K was highly improved. Regrettably, the explanatory power of the LMTL models correspondingly decreased as fewer features were used, increasing the model sparsity (Table 2). The SSR/SST of pH increased with the LMTL model, this result might be because the redundant information was successfully removed by the regularization and the useful information was kept.

#### 4.2. The Shared Features

Correlations between the modeling of different soil properties have been suggested by various studies (e.g., [25,28]). Our study also showed that the distributions of the important wavebands with high VIP scores for the seven key soil properties were quite similar, especially in the four feature-block regions (Figure 2). In addition, the features used in the LMTL models illustrated the existence of correlations between the different soil properties (Figure 5). The correlations were mostly attributed to the soil Fe oxides, water content, organics and clay minerals, which constitute the basis of the feature-block regions and shared features in soil spectroscopy. With the advantage of shared features, the prediction accuracies of several soil properties with low concentrations or unobvious spectral absorption signals improved. For example, the 1105, 1275, and 1965 nm wavebands obtained low VIP scores in the prediction of N with the PLS-R model, but were useful with the LMTL model as they were shared by P, WC and OM. The 650 nm waveband in the prediction of N, P and OM was shared by WC, the 1195 nm waveband in the prediction of P was shared by OM, the 1500 nm and 1985 nm wavebands in the prediction of P were shared by pH, and so on (Figure 5).

#### 4.3. Assessing the Performance of Field VNIR/SWIR Spectroscopy

The application of field VNIR/SWIR spectroscopy for quantitative soil property prediction is not a new idea, and great effort has been made toward the goal of improving the prediction accuracy of regression models, especially during the past two decades [3]. Table 4 summarizes the results of past studies for predicting seven key soil properties using the field spectroscopy analysis approach. Many studies conducted field spectral measurements using either a contact probe to touch the soil surface (or soil profile) (e.g., [16]) or a mobile subsoiler to penetrate the soil (e.g., [7]), both with a built-in light source. The performance results of non-contact field VNIR/SWIR spectroscopy were usually not as accurate as those obtained using contact spectroscopy due to the effects of atmospheric water absorption, residual coverings on the soil surface, and scattering [73].

**Table 4.** Summary of previous research results for predicting the soil properties used in this study.

Property	Range (nm)	Measurement Method	Regression Algorithm	RPD	R <sup>2</sup>	Accuracy Category	Literature
N	500–1600	Mobile	PLS-R	1.60	0.69	B	[8]
	350–2500	Contact probe	LS-SVM	1.91	0.76	B	[16]
P	920–1718	Mobile	PCR	-	0.65	B	[6]
	306.5–1710.9	Mobile	PLS-R	1.80	0.69	B	[7]
	500–1600	Mobile	PLS-R	1.80	0.72	B	[8]
	350–2500	Contact probe	PLS	1.33	0.43	C	[16]
	400–1050	Non-contact	ANN	-	0.87	A	[17]
	350–2500	Contact probe	MPLS-R	1.70	0.65	B	[26]
	1100–2300	Mobile	PLS-R	1.27	0.41	C	[23]
K	920–1718	Mobile	PCR	-	0.26	C	[6]
	350–2500	Contact probe	LS-SVM	0.91	0.14	C	[16]
	400–1050	Non-contact	ANN	-	0.85	A	[17]
	350–2500	Contact probe	MPLS-R	2.90	0.88	A	[26]
	1100–2300	Mobile	PLS-R	1.08	0.19	C	[23]
WC	920–1718	Mobile	PCR	-	0.40	C	[6]
	306.5–1710.9	Mobile	PLS-R	3.00	0.89	A	[7]
	500–1600	Mobile	PLS-R	3.60	0.93	A	[8]
	305–2200	Mobile	PLS-R	3.54	-	A	[10]
	305–2200	Mobile	MARS	3.25	0.72	A	[20]
pH	920–1718	Mobile	PCR	-	0.43	C	[6]
	306.5–1710.9	Mobile	PLS-R	2.14	0.71	A	[7]
	500–1600	Mobile	PLS-R	1.6	0.69	B	[8]
	350–2500	Contact probe	LS-SVM	2.23	0.80	A	[16]
	1100–2300	Mobile	PLS-R	1.88	0.71	B	[23]
EC	500–1600	Mobile	PLS-R	1.30	0.60	C	[8]
OM	920–1718	Mobile	PCR	-	0.67	B	[6]
	500–1600	Mobile	PLS-R	2.90	0.90	A	[8]
	350–2500	Contact probe	LS-SVM	2.18	0.81	A	[16]
	400–1050	Non-contact	ANN	-	0.84	A	[17]
	350–2500	Contact probe	MPLS-R	2.80	0.86	A	[26]
	350–2500	Contact probe	PLS-R	1.94	0.79	B	[74]
	1100–2300	Mobile	PLS-R	1.59	0.61	B	[23]

Category A: RPD > 2.0, Category B: 1.4 < RPD < 2.0, Category C: RPD < 1.4. Abbreviations used: available nitrogen (N), available phosphorous (P), available potassium (K), water content (WC), pH, electrical conductivity (EC), and organic matter (OM), the ratio of performance to deviation (RPD), coefficient of determination (R<sup>2</sup>), partial least squares regression (PLS-R), least squares support vector machine (LS-SVM), principal components regression (PCR), artificial neural network (ANN), modified partial least squares regression (MPLS-R).

It is well-known that soil OM can be successfully predicted by VNIR/SWIR spectroscopy because of its sensitivity to broad overtones and combination absorptions, such as O–H, C–H, and N–H [75]. Our results show that the prediction performance of OM had the highest accuracy among all of the seven soil properties, with RPD = 2.29, which is categorized as good accuracy (RPD > 2) and is comparable to previous studies that used contact spectroscopy. The prediction accuracy of pH varies among different studies, possibly due to the fact that pH is related to many soil properties but has no direct spectral absorptions [25]. The RPD of pH was 1.90, indicating moderate accuracy (1.4 < RPD < 2). The removal of the atmospheric water absorption wavebands (350–1420 and 1800–1960 nm) caused the loss of several useful features, especially for soil WC, which has significant absorption bands at around 1400 and 1900 nm [76]. Therefore, the prediction performance of WC in this study is not high with an RPD = 1.71, which can only be categorized as moderate accuracy. We found that among the three soil available nutrients, P has been relatively well studied with different levels of prediction accuracy; this could be attributed to the chemical measurement method of P, which can greatly affect its prediction [3]. The prediction accuracies of N, P and K in this study (RPD = 1.40, 1.49, and 1.22, respectively) were low, perhaps because of their extremely low concentrations in dryland soils. The prediction of EC in this study was unreliable possibly due to the wavebands sensitive to several key soil salinities, such as

NaCl at 1930 nm and Na<sub>2</sub>SO<sub>4</sub> at 1825 nm [77], which were removed along with the atmospheric water absorption wavebands.

#### 4.4. Next Steps

We have shown that using a LMTL algorithm based on a regularized dirty model can improve the prediction accuracies of seven key soil properties with the advantage of shared features and regularization. A larger dataset of soil samples may improve the performances of LMTL algorithms and enhance the shared features. The concentrations of soil Fe oxides and clay minerals should also be considered as outputs to enhance the learning ability of the LMTL models. Previous studies have argued that a nonlinear correlation exists between soil properties and spectral features (e.g., [16–20]). A future study should be conducted to apply nonlinear multi-task learning algorithms, such as a deep neural network [78,79], focusing on optimizing these algorithms to improve both the prediction accuracy and the explanatory power. In addition, airborne and satellite-based hyperspectral remote sensing should also be an important research area in which LMTL could be used to develop large-scale soil property monitoring and mapping.

## 5. Conclusions

Our study illustrates that LMTL algorithms can improve the prediction accuracies of seven key soil properties by field VNIR/SWIR spectroscopy in the drylands. Our results demonstrate that: (1) The used features for predicting different soil properties are correlated and most of them are attributed to soil Fe oxides, WC, OM and clay minerals. (2) In the current study, OM was predicted with good accuracy (RPD > 2); N, P, WC and pH were predicted with moderate accuracy (1.4 < RPD < 2); K and EC were predicted with poor accuracy (RPD < 1.4). (3) Compared to the PLS-R, LMTL models based on regularization algorithms usually have slightly higher prediction accuracy (with respect to the RPD values) and lower explanatory power (with respect to the SSR/SST values) as the used features are sparse. (4) LMTL could use the advantages of the shared features in the soil spectroscopy of different soil properties and improve the model generalization performance. Our study provides a novel analysis method for in-deep research on the underlying correlations in the soil spectroscopy of different soil properties, which can be used in future studies of soil property prediction and soil quality assessment based on spectroscopy.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2072-4292/9/11/1099/s1](http://www.mdpi.com/2072-4292/9/11/1099/s1). Figure S1: Sparsity (the number of non-zero elements) of: the block-sparse matrix  $W_b$  (1); the elementwise sparse matrix  $W_e$  (2); and the combined regression coefficients matrix  $W$  (3), of the model generated from linear multi-task learning for predicting available nitrogen (a); available phosphorous (b); available potassium (c); water content (d); pH (e); electrical conductivity (f); and organic matter (g). Figure S2: Used features (non-zero items in the transpose of the block-sparse matrix  $W_b$  (a); the elementwise sparse matrix  $W_e$  (b); and the combined regression coefficients matrix  $W$  (c)) of linear multi-task learning models with  $\lambda_b = 40$  and  $\lambda_e = 10$ .

**Acknowledgments:** This study was partially funded by H2020 project number 641762 “Improving Future Ecosystem Benefits through Earth Observations” (ECOPOTENTIAL) and partially funded by the International S&T Cooperation Project of the China Ministry of Agriculture (2015-Z44, 2016-X34). The authors also appreciate the financial support received by Haijun Qi from the China Scholarship Council (201608340066).

**Author Contributions:** Haijun Qi was principal to all phases of the investigation as well as manuscript preparation. He was supervised by Shaowen Li and Arnon Karnieli who coordinated the activities. Tarin Paz-Kagan provided the spectral data and significantly contributed her skills and experience.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ben-Dor, E.; Banin, A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Sci. Soc. Am. J.* **1995**, *59*, 364–372. [CrossRef]
2. Rossel, R.A.V.; McBratney, A.B. Soil chemical analytical accuracy and costs: Implications from precision agriculture. *Aust. J. Exp. Agric.* **1998**, *38*, 765. [CrossRef]

3. Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind, J. Chapter Five—Visible and Near Infrared Spectroscopy in Soil Science. *Adv. Agron.* **2010**, *107*, 163–215. [CrossRef]
4. Jiang, Q.; Chen, Y.; Guo, L.; Fei, T.; Qi, K. Estimating Soil Organic Carbon of Cropland Soil at Different Levels of Soil Moisture Using VIS-NIR Spectroscopy. *Remote Sens.* **2016**, *8*, 755. [CrossRef]
5. Selige, T.; Bohner, J.; Schmidhalter, U. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma* **2006**, *136*, 235–244. [CrossRef]
6. Christy, C.D. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Comput. Electron. Agric.* **2008**, *61*, 10–19. [CrossRef]
7. Mouazen, A.M.; Maleki, M.R.; De Baerdemaeker, J.; Ramon, H. On-line measurement of some selected soil properties using a VIS-NIR sensor. *Soil Tillage Res.* **2007**, *93*, 13–27. [CrossRef]
8. Kodaira, M.; Shibusawa, S. Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma* **2013**, *199*, 64–79. [CrossRef]
9. Viscarra Rossel, R.A.; Cattle, S.R.; Ortega, A.; Fouad, Y. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma* **2009**, *150*, 253–266. [CrossRef]
10. Kuang, B.; Mouazen, A.M. Effect of spiking strategy and ratio on calibration of on-line visible and near infrared soil sensor for measurement in European farms. *Soil Tillage Res.* **2013**, *128*, 125–136. [CrossRef]
11. Maimaitiyiming, M.; Ghulam, A.; Bozzolo, A.; Wilkins, J.L.; Kwasniewski, M.T. Early Detection of Plant Physiological Responses to Different Levels of Water Stress Using Reflectance Spectroscopy. *Remote Sens.* **2017**, *9*, 745. [CrossRef]
12. Sawut, M.; Ghulam, A.; Tiyip, T.; Zhang, Y.; Ding, J.; Zhang, F.; Maimaitiyiming, M. Estimating soil sand content using thermal infrared spectra in arid lands. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *33*, 203–210. [CrossRef]
13. Qi, H.; Paz-Kagan, T.; Karnieli, A.; Jin, X.; Li, S. Evaluating calibration methods for predicting soil available nutrients using hyperspectral VNIR data. *Soil Tillage Res.* **2018**, *175*, 267–275. [CrossRef]
14. Qi, H.; Jin, X.; Zhao, L.; Dedo, S.I.M.M.; Li, S. Predicting sandy soil moisture content with hyperspectral imaging. *Int. J. Agric. Biol. Eng.* **2017**, *10*. [CrossRef]
15. Wold, S.; Ruhe, A.; Wold, H.; Dunn, W.J., III. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735–743. [CrossRef]
16. Ji, W.; Shi, Z.; Huang, J.; Li, S. In situ measurement of some soil properties in paddy soil using visible and near-infrared spectroscopy. *PLoS ONE* **2014**, *9*, e105708. [CrossRef] [PubMed]
17. Daniel, K.W.; Tripathi, N.K.; Honda, K. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Aust. J. Soil Res.* **2003**, *41*, 47–59. [CrossRef]
18. Kuang, B.; Tekin, Y.; Mouazen, A.M. Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content. *Soil Tillage Res.* **2015**, *146*, 243–252. [CrossRef]
19. Bayer, A.; Bachmann, M.; Müller, A.; Kaufmann, H. A Comparison of feature-based MLR and PLS regression techniques for the prediction of three soil constituents in a degraded South African Ecosystem. *Appl. Environ. Soil Sci.* **2012**, *2012*, 1–20. [CrossRef]
20. Nawar, S.; Mouazen, A.M. Predictive performance of mobile vis-near infrared spectroscopy for key soil properties at different geographical scales by using spiking and data mining techniques. *CATENA* **2017**, *151*, 118–129. [CrossRef]
21. Wijewardane, N.K.; Ge, Y.; Morgan, C.L.S. Moisture insensitive prediction of soil properties from VNIR reflectance spectra based on external parameter orthogonalization. *Geoderma* **2016**, *267*, 92–101. [CrossRef]
22. Dyar, M.D.; Carmosino, M.L.; Breves, E.A.; Ozanne, M.V.; Clegg, S.M.; Wiens, R.C. Comparison of partial least squares and lasso regression techniques as applied to laser-induced breakdown spectroscopy of geological samples. *Spectrochim. Acta Part B At. Spectrosc.* **2012**, *70*, 51–67. [CrossRef]
23. Schirrmann, M.; Gebbers, R.; Kramer, E. Performance of Automated Near-Infrared Reflectance Spectrometry for Continuous in Situ Mapping of Soil Fertility at Field Scale. *Vadose Zone J.* **2013**, *12*, 1–14. [CrossRef]
24. Melendez-Pastor, I.; Navarro-Pedreño, J.; Gómez, I.; Koch, M. Identifying optimal spectral bands to assess soil properties with VNIR radiometry in semi-arid soils. *Geoderma* **2008**, *147*, 126–132. [CrossRef]
25. Rossel, R.A.V.; Behrens, T. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **2010**, *158*, 46–54. [CrossRef]

26. Gras, J.P.; Barthès, B.G.; Mahaut, B.; Trupin, S. Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils. *Geoderma* **2014**, *214*–215, 126–134. [CrossRef]
27. Ji, W.; Li, S.; Chen, S.; Shi, Z.; Viscarra Rossel, R.A.; Mouazen, A.M. Prediction of soil attributes using the Chinese soil spectral library and standardized spectra recorded at field conditions. *Soil Tillage Res.* **2016**, *155*, 492–500. [CrossRef]
28. Soriano-Disla, J.M.; Janik, L.J.; Viscarra Rossel, R.A.; MacDonald, L.M.; McLaughlin, M.J. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Appl. Spectrosc. Rev.* **2014**, *49*, 139–186. [CrossRef]
29. Evgeniou, T.; Pontil, M. Regularized Multi-Task Learning. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004; ACM: New York, NY, USA, 2004; pp. 109–117.
30. Caruana, R. Multitask Learning. *Mach. Learn.* **1997**, *28*, 41–75. [CrossRef]
31. Jalali, A.; Ravikumar, P.; Sanghavi, S.; Ruan, C. A Dirty Model for Multi-task Learning. In Proceedings of the Advances in Neural Information Processing Systems 23: 24th Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 6–9 December 2010; pp. 964–972.
32. Liu, J.; Ji, S.; Ye, J. Multi-task feature learning via efficient  $l_2, 1$ -norm minimization. In Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, 18–21 June 2009.
33. Zhang, M.L.; Zhou, Z.H. A Review On Multi-Label Learning Algorithms. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1819–1837. [CrossRef]
34. Yair, A.; Danin, A. Spatial variations in vegetation as related to the soil moisture regime over an arid limestone hillside, northern Negev, Israel. *Oecologia* **1980**, *47*, 83–88. [CrossRef] [PubMed]
35. Olsvig-Whittaker, L.; Shachak, M.; Yair, A. Vegetation patterns related to environmental factors in a Negev Desert watershed\*. *Plant Ecol.* **1983**, *54*, 153–165. [CrossRef]
36. Norman, R.; Stucki, J. The determination of nitrate and nitrite in soil extracts by ultraviolet spectrophotometry. *Soil Sci. Soc. Am.* **1981**, *45*, 347–353. [CrossRef]
37. Olsen, S. *Estimation of Available Phosphorus in Soils by Extraction with Sodium Bicarbonate*; United States Department of Agriculture: Washington, DC, USA, 1954.
38. Chen, S.; Peng, S.; Chen, B.; Chen, D.; Cheng, J. Effects of fire disturbance on the soil physical and chemical properties and vegetation of *Pinus massoniana* forest in south subtropical area. *Acta Ecol. Sin.* **2010**, *30*, 184–189. [CrossRef]
39. McGill, R.; Tukey, J.W.; Larsen, W.A. Variations of Box Plots. *Am. Stat.* **1978**, *32*, 12. [CrossRef]
40. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [CrossRef]
41. Helland, I.S.; Næs, T.; Isaksson, T. Related versions of the multiplicative scatter correction method for preprocessing spectroscopic data. *Chemom. Intell. Lab. Syst.* **1995**, *29*, 233–241. [CrossRef]
42. Gholizadeh, A.; Boruvka, L.; Saberioon, M.M.; Kozák, J.; Vašát, R.; Nemecek, K. Comparing different data preprocessing methods for monitoring soil heavy metals based on soil spectral features. *Soil Water Res.* **2015**, *10*, 218–227. [CrossRef]
43. Roberts, C.A.; Workman, J., Jr.; Reeves, J.B., III; Duckworth, J. *Mathematical Data Preprocessing*. In *Near-Infrared Spectroscopy in Agriculture*; American Society of Agronomy, Crop Science Society of America, Soil Science Society of America: Madison, WI, USA, 2004; pp. 115–132.
44. Rinnan, Å.; van den Berg, F.; Engelsen, S.B. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal. Chem.* **2009**, *28*, 1201–1222. [CrossRef]
45. Wold, S.; Martens, H.; Wold, H. The multivariate calibration problem in chemistry solved by the PLS method. In *Matrix Pencils. Lecture Notes in Mathematics*; Kågström, B., Ruhe, A., Eds.; Springer: Berlin/Heidelberg, Germany, 1983; pp. 286–293.
46. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [CrossRef]
47. Rossel, R.A.V.; Jeon, Y.S.; Odeh, I.O.A.; McBratney, A.B. Using a legacy soil sample to develop a mid-IR spectral library. *Aust. J. Soil Res.* **2008**, *46*, 1–16. [CrossRef]
48. Paz-Kagan, T.; Shachak, M.; Zaady, E.; Karnieli, A. A spectral soil quality index (SSQI) for characterizing soil function in areas of changed land use. *Geoderma* **2014**, *230*–231, 171–184. [CrossRef]

49. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning; Data Mining, Inference and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2008; ISBN 9780387848570.
50. Zhang, C.H.; Huang, J. The sparsity and bias of the lasso selection in high-dimensional linear regression. *Ann. Stat.* **2008**, *36*, 1567–1594. [CrossRef]
51. Negahban, S.; Wainwright, M.J. Joint support recovery under high-dimensional scaling: Benefits and perils of  $1, \infty$ -regularization. In Proceedings of the 21st International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–10 December 2008; pp. 1161–1168.
52. Kennard, R.W.; Stone, L.A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137–148. [CrossRef]
53. Hsu, C.-W.; Chang, C.-C.; Lin, C.-J. A Practical Guide to Support Vector Classification. Available online: <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf> (accessed on 12 October 2016).
54. Chang, C.-W.; Laird, D.A.; Mausbach, M.J.; Hurburgh, C.R. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [CrossRef]
55. Weisberg, S. *Applied Linear Regression*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2005; ISBN 9780471704096.
56. Wen, W.; Hao, Z.; Yang, X. A heuristic weight-setting strategy and iteratively updating algorithm for weighted least-squares support vector regression. *Neurocomputing* **2008**, *71*, 3096–3103. [CrossRef]
57. Peng, X. TSVR: An efficient Twin Support Vector Machine for regression. *Neural Netw.* **2010**, *23*, 365–372. [CrossRef] [PubMed]
58. Shmueli, G. To explain or to predict? *Stat. Sci.* **2010**, *25*, 289–310. [CrossRef]
59. Cope, M.; van der Zee, P.; Essenpreis, M.; Arridge, S.R.; Delpy, D.T. Data analysis methods for near-infrared spectroscopy of tissue: Problems in determining the relative cytochrome aa3 concentration. In Proceedings of the Time-Resolved Spectroscopy and Imaging of Tissues, Los Angeles, CA, USA, 23–24 January 1991; Volume 1431, pp. 251–262.
60. Zhou, J.; Chen, J.; Ye, J. MALSAR: Multi-tAsk Learning via StructurAl Regularization. Arizona State University, 2012. Available online: <http://www.yelab.net/software/MALSAR/> (accessed on 25 October 2016).
61. Paz-Kagan, T.; Zaady, E.; Salbach, C.; Schmidt, A.; Lausch, A.; Zacharias, S.; Notesco, G.; Ben-Dor, E.; Karnieli, A. Mapping the spectral soil quality index (SSQI) using airborne imaging spectroscopy. *Remote Sens.* **2015**, *7*, 15748–15781. [CrossRef]
62. Rollin, E.M.; Milton, E.J. Processing of high spectral resolution reflectance data for the retrieval of canopy water content information. *Remote Sens. Environ.* **1998**, *65*, 86–92. [CrossRef]
63. Post, J.L.; Noble, P.N. The near-infrared combination band frequencies of dioctahedral smectites, micas, and illites. *Clays Clay Miner.* **1993**, *41*, 639–644. [CrossRef]
64. Mouazen, A.M.; Kuang, B.; De Baerdemaeker, J.; Ramon, H. Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma* **2010**, *158*, 23–31. [CrossRef]
65. Wang, J.; He, T.; Lv, C.; Chen, Y.; Jian, W. Mapping soil organic matter based on land degradation spectral response units using Hyperion images. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, S171–S180. [CrossRef]
66. Rodriguez, J.M.; Ustin, S.L.; Riaño, D. Contributions of imaging spectroscopy to improve estimates of evapotranspiration. *Hydrol. Process.* **2011**, *25*, 4069–4081. [CrossRef]
67. Dehaan, R.; Taylor, G.R. Image-derived spectral endmembers as indicators of salinisation. *Int. J. Remote Sens.* **2003**, *24*, 775–794. [CrossRef]
68. Weidong, L.; Baret, F.; Xingfa, G.; Qingxi, T.; Lanfen, Z.; Bing, Z. Relating soil surface moisture to reflectance. *Remote Sens. Environ.* **2002**, *81*, 238–246. [CrossRef]
69. Van Evert, F.K.; Van Der Schans, D.A.; Van Geel, W.C.A.; Malda, J.T.; Vona, V. From theory to practice: Using canopy reflectance to determine sidedress N rate in potatoes. In *Precision Agriculture '13*; Stafford, J.V., Ed.; Wageningen Academic Publishers: Wageningen, The Netherlands, 2013; pp. 119–127.
70. Parfitt, R.L.; Smart, R.S.C. The mechanism of sulfate adsorption on iron oxides. *Soil Sci. Soc. Am. J.* **1978**, *42*, 48–50. [CrossRef]
71. Kusumo, B.H.; Hedley, C.B.; Hedley, M.J.; Hueni, A.; Tuohy, M.P.; Arnold, G.C. The use of diffuse reflectance spectroscopy for in situ carbon and nitrogen analysis of pastoral soils. *Aust. J. Soil Res.* **2008**, *46*, 623–635. [CrossRef]
72. Summers, D.; Lewis, M.; Ostendorf, B.; Chittleborough, D. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecol. Indic.* **2011**, *11*, 123–131. [CrossRef]
73. Baumgardner, M.; Silva, L.; Biehl, L.; Stoner, E. Reflectance properties of soils. *Adv. Agron.* **1986**, *38*, 1–44.

74. Ji, W.; Viscarra Rossel, R.A.; Shi, Z. Accounting for the effects of water and the environment on proximally sensed vis-NIR soil spectra and their calibrations. *Eur. J. Soil Sci.* **2015**, *66*, 555–565. [CrossRef]
75. Clark, R.N.; King, T.V.V.; Klejwa, M.; Swayze, G.; Vergo, N. High Spectral Resolution Reflectance Spectroscopy of Minerals. *J. Geophys. Res.* **1990**, 12653–12680. [CrossRef]
76. Stoner, E.R.; Baumgardner, M.F. *Physicochemical, Site, and Bidirectional Reflectance Factor Characteristics of Uniformly Moist Soils*; Laboratory for Applications of Remote Sensing, Purdue University: West Lafayette, IN, USA, 1980; Volume 1.
77. Farifteh, J.; van der Meer, F.; van der Meijde, M.; Atzberger, C. Spectral characteristics of salt-affected soils: A laboratory experiment. *Geoderma* **2008**, *145*, 196–206. [CrossRef]
78. Collobert, R.; Weston, J. A unified architecture for natural language processing: Deep neural networks with multitask learning. In Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, 5–9 July 2008; ACM: New York, NY, USA, 2008; pp. 160–167.
79. Seltzer, M.L.; Droppo, J. Multi-Task Learning in Deep Neural Networks for Improved Phoneme Recognition. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 26–31 May 2013; pp. 6965–6969.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).