Using hyperspectral remote sensing to monitor the properties of salt-affected soils

Gopal Mahajan[1], Bappa Das[2], Bhaskar Gaikwad[3], Ashwini Desai[2], Shaiesh Morajkar[2], Kiran Patel[2], Rahul Kulkarni[4], and Dayesh Murgaonkar[2]

[1]ICAR-Central Coastal Agricultural Research Institute
[2]ICAR Central Coastal Agricultural Research Institute
[3]National Institute of Abiotic Stress Management
[4]ICAR - Central Coatal Agricultural Research Institute

May 5, 2020

Abstract

The aim of the study was to estimate the properties of the salt-affected soils (SAS) using hyperspectral remote sensing. The study was carried out on typical SAS from 372 locations covering 17 coastal districts from west coast region of India. The spectral reflectance of processed soil samples was recorded in the wavelength range of 350-2500 nm. The full data set (n=372) was split into two as calibration dataset (n=260, 70%) to develop the model and validation dataset (n=112, 30%) to evaluate the performance of the model independently. The spectral data were calibrated using the laboratory estimated soil properties with five different multivariate techniques: (a) linear – partial component regression (PCR) and partial least square regression (PLSR) and (b) non-linear– multivariate adaptive regression spline (MARS), random forest (RF) and support vector regression (SVR). In general, the spectral reflectance from the soils decreased with increasing levels of salinity (electrical conductivity, EC). The wavelengths, 494, 673, 800, 1415, 1748, 1915, 2207 and 2385 nm showed peculiar absorption characteristics. The study showed significant achievement in predicting soil properties like soil pH, salinity (EC), bulk density (BD), soil available nitrogen (N), exchangeable magnesium (Mg), soil available zinc (Zn) and boron (B) with acceptable to excellent predictions (ratio of performance to deviation (RPD) ranged 1.48-2.06). Amongst predicted models, SVR, PLSR and PCR were found to be more robust than MARS and RF. The results of the study indicated that the visible near-infrared spectroscopy has the potential predict properties of the SAS.

1. INTRODUCTION

Soil salinization is one of the major land degradation processes in almost all parts of the world. About 836 million hectares (Mha) area of the world (3% of total geographical area) is affected with soil salinization, of which, 48% is saline and 52% is sodic (Singh, 2016). The extent of the salt-affected soils (SAS) in India is 6.73 Mha, of which 1.28 Mha (19%) is in the coastal region (CSSRI, 2018). Unlike the SAS formed by secondary salinization, the SAS of the coastal region possesses unique physical and chemical properties. The co-existence of the soil acidity (low soil pH) with high salinity level (high electrical conductivity, EC) is a unique feature of these soils. The ingression of saline water from sea and estuaries either naturally or through anthropogenic activities is the major reason for salinity in these soils. The excess amount of the salts affects the physical, chemical and biological properties of the soils and ultimately the plant growth (Mahajan *et al.* , 2016; Yuan *et al.* , 2007). The productivity of these soils is often very low. Quantitative assessment of the soil properties is very crucial to understand, maintain and improve the soil quality and enhance crop yield (Askari *et al.* , 2015; Xu *et al.* , 2018). But, the soil is a heterogeneous resource due to complex processes and mechanisms involved in soil formation (Morellos *et al.* , 2016). Real-time assessment of soil salinity at high temporal and spatial resolution and other properties are essential to manage these soils efficiently for crop

production. Soil sampling and laboratory analysis to adequately assess the spatial and temporal variability of soil properties are time-consuming and expensive (Xu *et al.* , 2018) and often limited to smaller areas. Such a challenge has attracted researchers worldwide in recent years to find alternate ways to overcome it. Compared to the conventional laboratory analysis methods, hyperspectral remote sensing (HRS) has been proposed as one of the modern, valid and alternate techniques for monitoring the soil properties (Stenberg *et al.* , 2010). Moreover, multiple soil properties can be estimated using a single representative spectral scan of each sample (Vohland *et al.* , 2011).

Based on the spectral absorption and reflection features in visible (VIS) and near-infrared (NIR) region, remote sensing technology could be employed for the estimation of soil properties and salt content (Islam *et al.* , 2003). Researchers have investigated the use of VIS-NIR based spectroscopy for estimating different soil attributes in different soil types. Multivariate statistics is required to build the relationship between the complex features or patterns of soil spectral data and the soil properties (Araújo *et al.* , 2014; Stenberg *et al.* , 2010). Among different multivariate statistical techniques, commonly used linear techniques are stepwise multiple linear regression (SMLR), principal component regression (PCR) and partial least square regression (PLSR), however, use of non-linear techniques like multivariate adaptive regression splines (MARS), random forest (RF), support vector machine regression (SVMR) are also becoming popular as relationships between spectral data and soil attributes are rarely linear (Araújo *et al.* , 2014; Xu *et al.* , 2018). Wang *et al.* (2018) found that the RF based model with 1.5 order derivative of absorbance was most effective, stable and accurate to quantify the soil salinity with coefficient of determination (R^2) = 0.93, root means square error (RMSE) = 4.57 dS m^{-1} and ratio of performance to deviation (RPD) = 2.78. Cécillon *et al.* (2009) and Stenberg *et al.* (2010) investigated the applicability of the VIS-NIR remote sensing and observed absorption bands for NaCl, KCl and MgSO$_4$ at 1930 nm, 1430 nm and 1480 nm, respectively. The use of PLSR and artificial neural network (ANN) to predict the salt concentration (NaCl, KCl, MgCl$_2$ and MgSO$_4$) has been successfully demonstrated by Farifteh *et al.* (2007). Nawar *et al.* (2014) recorded good prediction of soil EC based on the MARS model using soil spectral reflectance (R^2 = 0.73, RMSE = 6.53 dS m^{-1}, and RPD = 1.96) compared to PLSR. The use of VIS-NIR remote sensing for predicting the soil properties like soil organic carbon (SOC), soil pH, EC, total nitrogen (N), available N, total phosphorus (P), Mehlich 1 extractable P, total potassium (K), cation exchange capacity (CEC), moisture, soil texture, clay content, etc. has been studied widely (Araújo *et al.* , 2015; Cécillon *et al.* , 2009; Chang *et al.* , 2001; Christy, 2008; Morellos *et al.* , 2016; Schirrmann *et al.* , 2013; Vasques *et al.* , 2008, 2009; Viscarra Rossel *et al.* , 2006; Wenjun *et al.* , 2014; Xu *et al.* , 2018). The non-linear multivariate techniques (SVMR and Back Propagation Neural Network (BPNN)) outperformed the linear techniques (PCR and PLSR) to predict the soil organic matter, total N, total P and total K in soil cores of paddy fields (Xu *et al.* , 2018). The use of the VIS-NIR spectra for characterizing soils gives a large number of predictor variables but using the full spectra at high-resolution compromises with the multi-collinearity and noise (Vohland *et al.* , 2011). Thus, the selection of a proper multivariate techniques for calibration and prediction of a variable is an important factor (Mouazen *et al.* , 2010; Nawar *et al.* , 2016; Xu *et al.* , 2018).

To the best of our knowledge, very few studies have been conducted on SAS and almost none on SAS having acidic soil reaction. The objectives of the study were to (1) study effect of the soil salinity on the VIS-NIR spectral reflectance pattern, (2) investigate potential of VIS-NIR spectroscopy to estimate the various properties of SAS and (3) compare the predictive ability of the linear and non-linear multivariate techniques for estimation of soil properties of SAS.

2. MATERIALS AND METHOD

2.1 Description of sampling sites

The study was carried out on SAS of the west coast region of India. The study sites were selected from 17 coastal districts of the states of Maharashtra, Goa, Karnataka and Kerala. The sampling locations of the present study are presented in Figure 1. The locations were confirmed for the source of salinity before collecting the soil samples. The source of salinity was either the ingression of the saline water from sea and estuaries or saline groundwater due to natural or anthropogenic activities. The ten years average annual

rainfall and temperature of the study region were 2139 mm and 27.4 °C and the climate is warm and humid. Rice cultivation is prevalent in such soils due to reduction in salinity during monsoon season. However, the crop experiences salinity stress during the initial crop establishment and maturity stages and this, in turn, reduces the crop yield. These lands are mostly kept fallow by farmers during the rest period of the year due to the rise in soil salinity.

2.2 Soil sampling

Three replicate soil samples representing one location were collected during the pre-monsoon season of the year 2016 and 2017. The surface soil sample was drawn up to 0.15 m depth. Each replicate sample was a composite sample made from three subsamples to ensure homogeneity. For each location, an approximate distance of 200–300 m between replicate samples was maintained. The soil samples collected were air-dried, grounded, sieved through 2 mm sieve and stored in polythene bags for further analysis. Core samples were also drawn for the determination of soil bulk density (BD). A total of 372 soil samples, representing 372 sampling locations, were considered for the study.

2.3 Analysis of soil properties

The soil pH and EC were estimated in 1:2.5 soil to water suspension (Jackson, 1973). The BD was estimated using a soil core method (Blake, 1965). Ammonium acetate extractable cations like sodium (Na), potassium (K), calcium (Ca) and magnesium (Mg) (Hanway & Heidel, 1952) were estimated using atomic absorption spectrophotometer (AAS) (nova400P, Analytik Jena, Germany). The SOC was analyzed by Walkley & Black (1934) wet oxidation method. The soil available nutrients were estimated as - soil available nitrogen (N) by alkaline potassium permanganate method (Asija & Subbiah, 1956) and soil available phosphorus (P) by a colorimetric method (Bray & Kurtz, 1945). Diethylenetriaminepentaacetic acid extractable micronutrients like iron (Fe), manganese (Mn), copper (Cu) and zinc (Zn) (Lindsay & Norvell, 1978) were determined using AAS. The 0.15% $CaCl_2$ extractable sulfur (S) (Williams & Steinbergs, 1959) and hot water-soluble boron (B) (Gupta, 1967) were estimated using a spectrophotometer. The soil properties were tested for normality and under the situation of non-linearity Box-Cox transformation was used to make it normal for model development and the model estimation were back-transformed to original units for those soil properties for evaluation (Table 1).

2.4 Spectroscopic measurements

The spectral reflectance measurement of soils at 350-2500 nm wavelength were recorded using Fieldspec®4 ASD Spectroradiometer (Analytic Spectral Devices, CO USA). The bandwidth or sampling intervals of the Spectroradiometer was 1.4 nm and 2 nm in the 350-1000 nm and 1000-2500 nm wavelength range, respectively. The measurements were acquired using a contact probe attached to Spectroradiometer with a fiber-optic cable having a built-in light source and a 20 mm viewing window. The Spectroradiometer was calibrated with the Spectralan® panel (100% spectral reflectance) before recording the spectral measurement of the soil samples. For each sample, 50 spectral observations were captured and averaged to get the final spectral signature to improve the signal-to-noise ratio. Only every 5th wavelength was retained between 400 to 2500 nm and this led to 420 wavebands for further analysis.

2.5 Model calibration and validation

The full dataset (n=372) was split into two as calibration dataset with 70% (n=260) samples for model development and validation dataset with 30% (n=112) samples to assess the performance of the model independently. The spectral data were calibrated using the laboratory estimated soil properties by five multivariate techniques: (a) linear - PCR and PLSR and (b) non-linear - MARS, RF and support vector regression (SVR). The models were evaluated for prediction accuracy using model evaluation parameters like R^2, mean bias error (MBE), RMSE and RPD. The prediction accuracy of different models was categorized based on RPD as excellent (RPD > 2), acceptable (2 [?] RPD [?] 1.4) and non-reliable (RPD > 1.40) (Chang et al. , 2001). Generally, high R^2 and RPD and low RMSE indicate a model with good predictive ability. We, in the present study, calculated ranks for different models considering the R^2, MBE, RMSE and

3

RPD. Based on the ranking for calibration and validation, an average rank was calculated to interpret the prediction performance of the model. Lower the value of the rank, better was the prediction performance.

2.6 Statistical analysis

The descriptive statistics of the soil properties for calibration, validation and full dataset were carried out using SAS (SAS Institute, 2012). The calibration and validation of multivariate statistical models and estimation of the model evaluation parameters was done using R software version 3.5.2 (R Core Team, 2018).

RESULTS AND DISCUSSION

3.1 Descriptive statistics of the soil properties

Descriptive statistics of the properties of SAS are presented in Table 1. For the full dataset, the coefficient of variation varied from 16.75-118.30% with the highest observed for Fe (118.30%) and lowest for soil pH (16.75%). For the calibration and validation dataset coefficient of variation varied from 18.46 to 122.91% and 15.44 to 110.85%, respectively. The highest coefficient of variation (%) was recorded for Fe (122.91%) and lowest for soil pH (18.46%) for the calibration dataset whereas in the case of validation dataset it was highest for Mn (110.85%) and lowest for soil pH (15.44%). The soil pH and EC of the full dataset varied from 3.47-8.04 (mean 5.42) and 0.10-24.11 (mean 6.91) dS m^{-1}, respectively. It indicates co-existence of the salinity and acidity which is the unique property of these soil . The SOC ranged from 0.02% to 3.20% with a CV of 56%. A physical property, BD, varied between 0.89-2.17 Mg m^{-3} with an average of 1.44 Mg m^{-3}and CV of 18.56%. The variability with respect to the macronutrient elements N and P was 42.76% and 93.00%. The cations, K, Na, Ca and Mg had coefficient of variation ranging from 65.97% to 83.44%. The variability of the micronutrients (Fe, Mn, Cu, Zn and B) varied from 63.93% to 118.30%. The results of the descriptive statistics revealed presence of adequate variability in most of the parameters measured. All the soil properties were positively skewed for full, calibration (except pH) and validation dataset. To ensure the random selection of calibration and validation dataset, test of equal means, variance, distribution and coefficient of variation was carried out and the results are presented in Table 2. The results of all the test were non-significant (p [?] 0.06) except F-test (equal variance) for S (p=0.05) and Zn (p=0.00004). This indicates that the calibration and validation dataset effectively represent the variability existed in the full dataset.

3.2 Relationship between different soil properties

The soil pH and EC correlated significantly (r=0.37, p<0.01) with each other (Table 3). The Fe had significant and negative (r=-0.46, p<0.01) correlation with soil pH whereas it was significant and positive with Mn (r=0.13, p<0.05) and Cu (r=0.23, p<0.01) and non-significant with Zn. The EC strongly and positively correlated with K (r=0.37, p<0.01) and Mg (r=0.29, p<0.01) whereas it was negatively correlated with Na (r=-0.17, p<0.01). The SOC and BD had a significant and negative correlation (r=-0.32, p<0.01). The K content correlated significantly with Na (r=0.30, p<0.01) and Mg (r=0.37, p<0.01). It was interesting to note the highest, significant and positive correlation (r=0.60, p<0.01) between the Fe and S content.

3.3 Spectral reflectance characteristics of the SAS

The spectral reflectance of the calibration and validation dataset exhibited variations which is evident from Figure 2. In order to understand the effect of soil salinity on spectral reflectance, the samples were classified in five different salinity classes (Nawar*et al.* , 2011, 2015) and their class-wise average spectral reflectance was calculated. The classes were calculated as non-saline (1.05 dS m^{-1}, 0-2 dS m^{-1}), slightly saline (3.17 dS m^{-1}, 2-4 dS m^{-1}), moderately saline (5.72 dS m^{-1}, 4-8 dS m^{-1}), strongly saline (11.58 dS m^{-1}, 8-16 dS m^{-1}) and very strongly saline (22.15 dS m^{-1}, >16 dS m^{-1}). The spectral reflectance characteristics by salinity class are presented in Figure 3. There were conspicuous absorption deeps in the region close to wavelength 427, 487 and 1917 nm. Besides, weak absorption features were also observed at regions near to 950, 1414, 2206, 2380 and 2460 nm. It is noticeable that the absorption at these wavelengths varies with the salinity class and generally increased with increasing salinity level. The absorption features close to 1414 and 1917

nm represent the stretching of oxygen (O)-hydrogen (H) and bending of H-O-H of the free water and its overtones. As the salinity increased the features became more asymmetrical and extended (Nawar *et al.*, 2014). The shape of the spectral signature in the VIS-NIR region was more or less similar in all five salinity classes. Overall, the reflectance spectra showed a tendency to decrease with the increase in EC. Similar results were reported by Nawar *et al.* (2014). . The SAS of the coastal region primarily has chlorides and sulphates of Na, Mg, Ca and K (Mahajan*et al.*, 2016), of which few are hygroscopic in nature (e.g. $MgCl_2$). These can absorb water vapor and increase soil moisture content. So, the soil moisture content increases with an increase in soil salinity, resulting in reduced reflectance or higher absorption. These results are concordant with those by Wang *et al.* (2018), Nawar *et al.* (2014) and Sidike *et al.* (2014).

3.4 Prediction accuracy of the multivariate models

The results pertaining to the prediction accuracy of multivariate models for both calibration and validation dataset has been presented in Table 4.

3.4.1 Soil pH prediction

The soil pH was most accurately predicted by MARS (R^2_c=0.90; MBE=-0.003; $RMSE_c$=0.34; RPD_c=2.98) and SVR (R^2_p=0.66; MBE_p=0.04; $RMSE_p$=0.51; RPD_p=1.62) during calibration and validation, respectively (Figure 4a). The calibration accuracy of the SVR (R^2_c=0.89; MBE_c=-0.008; $RMSE_c$=0.35; RPD_c=2.89) was comparable to that of MARS. The validation prediction accuracy of SVR was classified as acceptable. The overall rank based on the model evaluation parameters indicated SVR (1.5) as the best performing model to predict soil pH. Though the calibration ranking of MARS was 1, the rank during validation was 4.25 indicated inefficiency of the MARS model. The cross-validation accuracy to predict soil pH using the PLSR of reflectance in mid-infrared (MIR) and combined VIS-NIR-MIR was recorded by Viscarra Rossel *et al.*, (2006) as R^2_{adj} 0.75 and 0.33, respectively. They recorded the best prediction using MIR and observed RMSE of 0.10 unit. The prediction accuracies reported in the present study for pH were less accurate than literatures as $R^2$of 0.74 using PLSR (Reeves & McCarty, 2001), 0.73 using PLSR (Reeves*et al.*, 1999), 0.70 using MARS (Shepherd & Walsh, 2002), 0.70 using PCR (Islam *et al.*, 2003), 0.56 using PCR (Sun *et al.*, 2003). In the present study, the poor predictions for the soil pH might be attributable to the lower variability in the full (CV=16.75%), calibration (CV=18.46%) and validation (CV=15.44%) dataset (Table 1). Most of the samples of the study had acidic soil reaction.

3.4.2 Salinity (EC) prediction

The soil salinity is usually measured in terms of $EC_{1:2.5}$ and EC of saturation paste extract (EC_e). In the present investigation, $EC_{1:2.5}$ has been considered as a measure of soil salinity. The results of the prediction of $EC_{1:2.5}$ were similar to that of soil $pH_{1:2.5}$. The MARS model performed the best for calibration (R^2_c=0.95, MBE_c=-0.05, $RMSE_c$=1.45 and RPD_c = 3.68) and it was classified as excellent which was followed by the SVR (R^2_c=0.92; MBE_c=-0.03; $RMSE_c$=1.92; RPD_c=2.76). The prediction accuracy of the SVR for validation was found excellent (R^2_p=0.80; MBE_p=0.30; $RMSE_p$=2.64; RPD_p=2.06) (Figure 4b). The SVR had the best validation and an overall rank of 1.00 and 1.38, respectively. This was followed by MARS which had a prediction accuracy of R^2_p=0.66; MBE_p=0.86; $RMSE_p$=3.88; RPD_p=1.40). In the present study, non-linear models predicted the soil salinity ($EC_{1:2.5}$) better than the linear models. These results are in line with the findings of Bilgili*et al.*, (2011), Farifteh *et al.*, (2007), Nawar *et al.*, (2014), Sidike *et al.*, (2014). Nawar *et al.*, (2014) reported that the non-linear multivariate technique (MARS) (R^2=0.73; RMSE=6.53; RPD=1.96) is more suitable to map the soil salinity than the linear model (PLSR) (R^2=0.70; RMSE=6.95; RPD=1.82) and the performance of the MARS model was improved using the continuum-removed spectral data in 400-2500 nm wavelength range. On the contrary to these results, Improvement in prediction using the non-linear model might be due to their capability to fit the complex and non-linear relationships (Friedman, 1991; Nawar *et al.*, 2014; Volkan Bilgili *et al.*, 2010). The studies have demonstrated that the high soil salinity implies a non-linear relationship between the measured salinity and the spectral reflectance (Farifteh *et al.*, 2007; Sidike *et al.*, 2014; Weng *et al.*, 2008). Farifteh *et al.*, (2007) considered the PLSR as advantageous model over ANN as the prediction accuracies were similar and low time requirement for the

establishment of the model and reproducibility of the final model. But, in the present investigation, the predictions using the non-linear models (MARS and SVR) were comparatively better than those by linear.

3.4.3 SOC prediction

The results obtained in the present study for prediction of the SOC were found poor. For both, calibration (R^2_c=0.83, MBE_c=1.39, $RMSE_c$=1.44 and RPD_c=0.43) and validation (R^2_c=0.52; MBE_c=1.40; $RMSE_c$=1.54; RPD_c=0.41), the SVR model was found to have better performance (Figure 4c). In general, the prediction results for SOC using different multivariate models revealed that the predictive ability of these models was poor. If the organic matter content is less than 2%, the spectral reflectance properties of the soil are hardly affected (Baumgardner *et al.* , 1986; Weng*et al.* , 2010; Nawar *et al.* , 2014). In the present study, the average SOC content was 1.11% with a CV of 56.0%. The results are concordant with those reported by Nawar *et al.* (2014).

3.4.4 BD prediction

Although the prediction accuracy during calibration for BD was observed as excellent (R^2_c = 0.99; MBE_c = -0.01; $RMSE_c$ = 0.05; RPD_c = 5.94) using the SVR model, the results during validation were categorized as non-reliable. The highest prediction accuracy (R^2_p = 0.52; MBE_p = 0.03; $RMSE_p$ = 0.22; RPD_p = 1.13) for BD was recorded with PLSR (Figure 4d). The BD is one of the fundamental soil physical properties and is often used as an indicator of soil quality, site productivity and soil compaction (Hakansson, 1990; Suuster *et al.* , 2011). Previously, only a few studies were carried out mainly focusing on the mid-infrared region (MIR) for predicting soil BD (Minasny *et al.* , 2008). They observed unsuccessful prediction as BD is related to a structural pore-space condition that cannot be captured by MIR spectra. Saeki*et al.* (2003) predicted BD using the NIR spectroscopy with a good prediction accuracy using the PLSR technique (R^2 = 0.96). Moreira *et al.* (2009) reported good and satisfactory prediction with R^2_c= 0.34 obtained using modified PLSR technique (mPLSR) and NIR spectral information. According to Minasny *et al.* (2008) under certain conditions soil BD is strongly related to the organic matter content of the soil, of which major mass (approximately 58%) exists as carbon (C). Poor prediction accuracy using VIS-NIR spectroscopy in our study could be attributed indirectly with low prediction accuracy for some soil parameters (e.g. total C and N) which shows relatively a small variation in the soil as suggested by van Groenigen *et al.* (2003).

3.4.5 Soil N prediction

For N, excellent calibration prediction accuracy of R^2_c=0.87; MBE_c=0.10; $RMSE_c$=35.16; RPD_c=2.49 was achieved by using the MARS model. The validation prediction accuracy was non-reliable with R^2_p=0.58; MBE_p=-8.80; $RMSE_p$=62.08; RPD_p=1.34. Among all the tested models, the overall ranking of MARS (2.38) and PCR (2.38) was found the best. However, the validation prediction rank of PCR (1.25) was best to predict the N. Similar to our findings Xu *et al.* (2018) reported better prediction accuracy for total nitrogen (TN) using PCR as one of the calibration models with R^2_p=0.78-0.86 and RPD_p=2.13-2.69. Yet in other studies, support vector machine regression (SVMR) and back propagation neural network (BPNN) showed better performance with R^2_p=0.69-0.88 predicting TN (Aliah Baharom *et al.* , 2015; Cozzolino *et al.* , 2013; Kodaira & Shibusawa, 2013; Kusumo *et al.* , 2008; Wenjun *et al.* , 2014). Bands around 1100, 1600, 1700-1800, 2000 and 2000-2400 nm have been identified as being important for SOC and TN (Stenberg *et al.* , 2010). Martin *et al.* (2002) found a high correlation (r=0.96) for the NIR predicted soil C and N. According to Williams & Norris (2001) prediction of N could be due to the known nitrogen specific absorption bands such as covalent bonds with H, C or O.

3.4.6 Soil P prediction

The relationship between the observed and predicted P for calibration revealed MARS to predict the P with acceptable prediction accuracy (R^2_c=0.80; MBE_c=-1.12; $RMSE_c$=10.75; RPD_c=1.80 and rank=1.00). These predictions were classified as good. The prediction accuracy for the validation using PLSR was observed as non-reliable with R^2_p=0.33; MBE_p=-3.41; $RMSE_p$=19.09; RPD_p=0.89. The best overall rank of 2.25 was recorded for the PLSR (Figure 4f). Xu *et al.*(2018) reported comparably better prediction accuracy based

on SVMR and BPNN model for total phosphorus (TP) with R^2=0.35-0.81. Similarly, Udelhoven et al. , (2003), Cozzolino et al. , (2013), Hu (2013) and Schirrmann et al. (2013) also reported possibilities of P prediction using spectroscopy through multivariate modeling.

3.4.7 Soil K prediction

The prediction of the K during calibration was the best when SVR was used (R^2_c=0.95; MBE_c=-27.59; $RMSE_c$=246.34; RPD_c=3.35) and it was considered as an excellent prediction (rank 1.00). It was followed by MARS which also generated excellent predictions (R^2_c=0.88; MBE_c=-43.62; $RMSE_c$=346.90; RPD_c=2.38) and had a rank of 2.00. For validation, the highest prediction accuracy (R^2_p = 0.46; MBE_p = 35.46; $RMSE_p$=604.31; RPD_p=1.31) was achieved using the PLSR, but was considered as non-reliable (Figure 4g). The validation and an overall ranks of PLSR were 1.25 and 2.25, respectively. Xu et al. (2018) reported better and improved prediction for soil total potassium (TK) with R^2=0.55-0.77 using SVMR and BBPNN models over the PLSR model by Cozzolino et al. , (2013) and Schirrmann et al. , (2013). Better performance of SVMR and BBPNN models for predicting P and K might be attributed to non-linear behavior of the soil variables with spectral reflectance data which was better captured by SVMR and BBPNN compared to the best performing model (PLSR) identified in our study. Poor prediction of soil P and K using the VIS-NIR region have been reported for lab and field-based spectroscopy in numerous studies (He et al. , 2007; Kuang & Mouazen, 2011; Malmir et al. , 2019; Viscarra Rossel et al. , 2006). Wenjun et al. (2014) recorded poor predictions for soil P and K using the lab and field-based VIS-NIR spectroscopy. Malmir et al. (2019) found the inability of PLSR model to predict soil P and K using VIS-NIR (400-1000 nm) spectroscopy. Based on the overall ranking, the PLSR model for predicting soil P and K found to be the better compared to others, which could be due to the high variability of P and K content and broader spectral range (350-2500 nm).

3.4.8 Soil Na prediction

The MARS model gave the highest prediction accuracy (R^2_c=0.87; MBE_c=-0.98; $RMSE_c$=10.65; RPD_c=1.79) during calibration to predict the Na and the predictions were acceptable. The second-best model was SVR with a prediction accuracy of R^2_c=0.74; MBE_c=-0.66; $RMSE_c$=15.67; RPD_c=1.21. The Na prediction using validation dataset had non-reliable prediction accuracy (R^2_p=0.55; MBE_p=4.81; $RMSE_p$=20.63; RPD_p=0.82) with the SVR model. The prediction accuracy was the highest with SVR compared to other models. The best overall rank of 2.38 was recorded by SVR (Figure 4h). Malmir et al. (2019) reported acceptable Na prediction using PLSR and soil spectral data from 400 to 1000 nm. A very good prediction of Na extracted from soil saturation (n=402, R^2=0.88, RMSE=2.45, RPD=2.89, excellent) using laboratory-measured soil spectral reflectance (350-2500 nm) was reported by Das et al. (2015). Islam et al. (2003) reported poor prediction for exchangeable Na using PCR modeling of spectral data in the range of 250-2500 nm. Similarly, Chang et al. (2001) reported inaccuracy in predicting Mehlich-3 extractable exchangeable Na using NIR-PCR model (R^2<0.50) for a spectral range of 400-2498 nm. Poor prediction (R^2=0.33) of exchangeable Na using the MIR spectroscopy was observed by Janik et al. (1998). Better performance of the non-linear model could be due to the fact that the relationship between spectral data and soil characteristic is rarely linear in nature. Variability in prediction accuracy of Na using reflectance spectroscopy could be due to variations in modeling, extraction or spectral data collection methods (field, processed or dried, intact core soil samples).

3.4.9 Soil Ca prediction

Similar to the results of Na prediction, the MARS model had the highest prediction accuracy for Ca using the calibration dataset (R^2_c=0.81; MBE_c=-0.09; $RMSE_c$=1.60; RPD_c=1.78). It was then followed by SVR (R^2_c=0.77; MBE_c=-0.06; $RMSE_c$=1.49; RPD_c=1.91) which was noticeable here that the rank for calibration using SVR (1.25) was better than MARS (1.75) and their corresponding predictions were considered as excellent. The PLSR provided the highest accuracy using the validation dataset (Figure 4i) with a prediction accuracy of R^2_p=0.43; MBE_p=-0.11; $RMSE_p$=2.27; RPD_p=1.20 and rank=2.63. But, this validation prediction was non-reliable. Malmir et al.(2019) also reported excellent prediction using PLSR model of reflectance data (400-1000 nm) with R^2_{cv}=0.81 and $RMSE_{cv}$=260.79 (for grounded sample) and R^2_{cv}=0.81

and $RMSE_{cv}$=260.97 (for sieved sample). With the use of the airborne hyperspectral imaging (350-2400 nm) and VIS-NIR spectroscopy (400-2500nm), the soil Ca concentration could be predicted with acceptable R^2=0.69-0.80 (Dematte *et al.* , 2016; Hively *et al.* , 2011). Janik *et al.* (1998) and Cozzolino & Moron (2003) recorded a very good Ca prediction accuracy (R^2=0.89 and 0.90, respectively) using PLSR modeling of VIS-NIR and MIR spectral data, respectively.

3.4.10 Soil Mg prediction

Similar to that of Na and Ca, calibration prediction accuracy of MARS for predicting the Mg was found the highest and excellent (R^2_c=0.89; MBE_c=-0.09; $RMSE_c$=1.37; RPD_c=2.36, rank=1.25). The acceptable validation was achieved with PCR (R^2_p=0.61; MBE_p=-0.17; $RMSE_p$=2.27; RPD_p=1.60 and rank=1.00). The best overall rank was observed using the PCR (1.75) (Figure 4j). Our results of the best performance of the PCR model are concordant with Islam *et al.* (2003) and Chang *et al.* (2001) who reported R^2=0.68 and 0.63, respectively. Janik *et al.*(1998) reported good prediction (R^2=0.79) for the same linear technique i.e. PCR using the MIR spectral data. Prediction of soil Mg using mPLSR (R^2=0.90) (Cozzolino *et al.* , 2013), airborne hyperspectral imaging sensors (R^2_c=0.69) (Dematte *et al.* , 2016) and NIR spectroscopy (R^2=0.89 and RPD=3.08) (Hively *et al.* , 2011) has been reported in the literature.

3.4.11 Soil S prediction

The results pertaining to the S showed that the calibration prediction accuracy was the highest with SVR (R^2_c=0.96; MBE_c=-0.41; $RMSE_c$=5.79; RPD_c=3.21; rank=1.25) and it was followed by MARS (R^2_c=0.93; MBE_c=-0.08; $RMSE_c$=6.58; RPD_c=2.82; rank=1.75). Among different models validated, PLSR had the highest but non-reliable predictions (R^2_p=0.76; MBE_p=0.39; $RMSE_p$=11.33; RPD_p=1.27; rank=1.00). The best overall rank was obtained by using PLSR to predict the S (Figure 4k). Cozzolino *et al.* (2013) predicted soil total sulfur (TS) using the spectral reflectance data (350-1800 nm) with accuracy of R^2=0.81 using PLSR analysis. Wijewardane *et al.*(2018) reported satisfactory prediction of TS with R^2>0.95 and RPD>5.5 using PLSR and ANN models in the MIR region.

3.4.12 Soil Fe prediction

The unique feature of the salt-affected soils of the study is that they possess an acidic soil reaction. The average soil pH of the full dataset was 5.42. These soils found to have high amounts of micronutrients like Fe, Mn, Cu and Zn under acidic soil reaction. The prediction accuracy for calibration was the highest (R^2_c=0.94; MBE_c=1.88; $RMSE_c$=22.57; RPD_c=3.35) with MARS model (Figure 4l). The prediction accuracy was considered excellent. Excellent, but lesser compared to MARS, prediction accuracy was observed with RF (R^2_c=0.92; MBE_c=-0.65; $RMSE_c$=29.60; RPD_c=2.56) and SVR (R^2_c=0.92; MBE_c=0.18; $RMSE_c$=26.39; RPD_c=2.87). The Fe prediction using the validation dataset was highest but non-reliable using PLSR (R^2_p=0.82; MBE_c=6.16; $RMSE_c$=39.62; RPD_c=1.13 and rank=1.25). Malmir *et al.* (2019) reported poor prediction accuracy for soil Fe and Mn using laboratory-based hyperspectral imaging (HIS) operated in 400-1000 nm with R^2_{cv} value in the range of 0.24-28 and 0.09-0.18 for Fe and Mn, respectively. Researchers have reported the use of spectral data through multivariate models to predict soil Fe with an accuracy R^2 of 0.49-0.90 (Janik*et al.* , 1998; Chang *et al.* , 2001; Islam *et al.* , 2003; Cozzolino & Moron, 2003).

3.4.13 Soil Mn prediction

Excellent predictions (R^2_c=0.99; MBE_c=-0.72; $RMSE_c$=2.16; RPD_c=3.34) were obtained using the SVR model for Mn (Figure 4m). The MARS models also recorded an excellent prediction (R^2_c=0.93; MBE_c=-0.25; $RMSE_c$=3.19; RPD_c=2.25). Though the predictions for calibration using SVR and MARS were excellent, the validation predictions using PCR (R^2_p=0.73; MBE_p=-0.84; $RMSE_p$=6.20; RPD_p=1.24) and PLSR (R^2_p=0.52; MBE_p=-0.89; $RMSE_p$=6.04; RPD_p=1.28) were better than other models tested. These results of prediction were considered as non-reliable. The results pertaining to Mn are in agreement with those reported by Janik *et al.* , (1998), who recorded R^2 of 0.57 and 0.66 to predict DTPA extractable-Mn and exchangeable-Mn using the PLSR analysis of the MIR spectral data (2500-25000 nm). Mehlich-III extractable-Mn could be monitored using PCR of the VIS–NIR (400–2498 nm) with an accuracy of R^2=0.70

(Chang *et al.* , 2001). Similar to Fe, Mn also showed poor prediction accuracy and this is in accordance with Vendrame *et al.* (2012) who reported poor prediction with R^2=0.35 and RPD=1.1 using multivariate analysis of NIR spectral data to predict soil MnO_2 content.

3.4.14 Soil Cu prediction

For Cu, the MARS model had excellent prediction accuracy (R^2_c=0.92; MBE_c=-0.08; $RMSE_c$=1.39; RPD_c=3.02; rank=1.00) for calibration (Figure 4n). The accuracy of the prediction during validation was found the highest and more or less similar for PLSR (R^2_p=0.40; MBE_p=-0.27; $RMSE_p$=3.12; RPD_p=1.17) and RF (R^2_p=0.39; MBE_p=-0.09; $RMSE_p$=3.12; RPD_p=1.17). The results are concordant with those of Cozzolino & Moron (2003) and Chang *et al.* (2001) for soil Cu with the mPLSR model.

3.4.15 Soil Zn prediction

In case of Zn, highest and acceptable calibration predictions were recorded using the MARS (R^2_c=0.79; MBE_c=0.01; $RMSE_c$=0.72; RPD_c=1.83) and SVR (R^2_c=0.78; MBE_c=-0.01; $RMSE_c$=0.69; RPD_c=1.91). The later had highest prediction accuracy (R^2_p=0.63; MBE_p=-0.07; $RMSE_p$=0.73; RPD_p=1.48) for validation which is considered as an acceptable prediction. Prediction accuracy of the R^2 from 0.64 to 0.84 using the airborne hyperspectral imaging and VIS-NIR spectroscopy is reported in literature (Hively *et al.* , 2011; Siebielec *et al.* , 2004; Sun *et al.* , 2017; Wu *et al.* , 2005). Chang *et al.* (2001) predicted Mehlich-III extractable-Zn with an accuracy of R^2=0.44 using PCR model in VIS–NIR region (400–2498 nm).

3.4.16 Soil B prediction

Excellent prediction for calibration was obtained using the MARS model (R^2_c=0.88; MBE_c=0.06; $RMSE_c$=4.03; RPD_c=2.42) (Figure 4p). Whereas, PCR had the highest and acceptable prediction accuracy (R^2_p=0.73; MBE_p=-1.79; $RMSE_p$=6.12; RPD_p=1.74) for validation. Soil B has been predicted with low R^2=0.17-0.30 using airborne hyperspectral imaging and laboratory MIR spectroscopy (2500-25000 nm) (Hively *et al.* , 2011; Janik *et al.* , 1998). Malmir *et al.* (2019) reported acceptable prediction accuracy of R^2=0.71 with PLSR model to predict soil B using hyperspectral image data in the spectral range of 400-1000 nm.

4. Conclusion

This study concludes a decreasing trend of spectral reflectance with increasing levels of salinity for SAS. The best probable wavelengths were identified that showed potential absorption characteristics to predict soil properties. Overall, hyperspectral remote sensing (350-2500 nm) showed acceptable to excellent prediction accuracy for soil pH, EC, BD, N, Mg, Zn and B. For SOC, P, K, Na, Ca, S, Fe, Mn and Cu, the prediction results were non-reliable. The best prediction was obtained for soil salinity (electrical conductivity). Multivariate techniques like SVR, PLSR and PCR were found more robust compared to MARS and RF for prediction of properties in SAS.

Conflict of interest: There is no conflict of interest

References

Aliah Baharom SN, Shibusawa S, Kodaira M, Kanda R. 2015. Multiple-depth mapping of soil properties using a visible and near infrared real-time soil sensor for a paddy field. *Engineering in Agriculture, Environment and Food* **8** : 13–17. DOI: 10.1016/j.eaef.2015.01.002

Araujo SR, Soderstrom M, Eriksson J, Isendahl C, Stenborg P, Dematte JM. 2015. Determining soil properties in Amazonian Dark Earths by reflectance spectroscopy. *Geoderma* **237** –**238** : 308–317. DOI: 10.1016/j.geoderma.2014.09.014

Araujo SR, Wetterlind J, Dematte JAM, 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. *European Journal of Soil Science* **65** : 718–729. DOI: 10.1111/ejss.12165

Asija GL, Subbiah B V. 1956. A rapid procedure for the estimation of available nitrogen in soils. *Curr. Sci.* **25** : 259–260

Askari MS, O'Rourke SM, Holden NM. 2015. Evaluation of soil quality for agricultural production using visible–near-infrared spectroscopy.*Geoderma* **243** –**244** : 80–91. DOI: 10.1016/j.geoderma.2014.12.012

Baumgardner MF, Silva LF, Biehl LL, Stoner ER. 1986. Reflectance Properties of Soils. , 1–44. DOI: 10.1016/S0065-2113(08)60672-0

Bilgili AV, Cullu MA, van Es H, Aydemir A, Aydemir S. 2011. The Use of Hyperspectral Visible and Near Infrared Reflectance Spectroscopy for the Characterization of Salt-Affected Soils in the Harran Plain, Turkey.*Arid Land Research and Management* **25** : 19–37. DOI: 10.1080/15324982.2010.528153

Bray RH, Kurtz LT. 1945. Determination of total, organic, and available forms of phosphorus in soils. *Soil Science* **59** : 39–46. DOI: 10.1097/00010694-194501000-00006

Cecillon L, Barthes BG, Gomez C, Ertlen D, Genot V, Hedde M, Stevens A, Brun JJ. 2009. Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS). *European Journal of Soil Science* **60** : 770–784. DOI: 10.1111/j.1365-2389.2009.01178.x

Chang C-W, Laird DA, Mausbach MJ, Hurburgh CR. 2001. Near-Infrared Reflectance Spectroscopy– Principal Components Regression Analyses of Soil Properties. *Soil Science Society of America Journal***65** : 480. DOI: 10.2136/sssaj2001.652480x

Christy CD. 2008. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. *Computers and Electronics in Agriculture* **61** : 10–19. DOI: 10.1016/j.compag.2007.02.010

Cozzolino D, Cynkar WU, Dambergs RG, Shah N, Smith P. 2013. In Situ Measurement of Soil Chemical Composition by Near-Infrared Spectroscopy: A Tool Toward Sustainable Vineyard Management. *Communications in Soil Science and Plant Analysis* **44** : 1610–1619. DOI: 10.1080/00103624.2013.768263

COZZOLINO D, MORON A. 2003. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics.*The Journal of Agricultural Science* **140** : 65–71. DOI: 10.1017/S0021859602002836

CSSRI. 2018. *Extent and distribution of salt affect soils in India* .

Das BS, Sarathjith MC, Santra P, Sahoo RN, Srivastava R, Routray A, Ray SS. 2015. Hyperspectral remote sensing: opportunities, status and challenges for rapid soil assessment in India. *Current science*860–868

Dematte JAM, Oliveira J de C, Tavares TR, Lopez LR, Terra F da S, Araujo SR, Fongaro CT, Maia SMF, Mello FF de C, Rizzo R, Vicente S, de Melo Bortolleto MA, Cerqueira PHR. 2016. Soil chemical alteration due to slaughterhouse waste application as identified by spectral reflectance in Sao Paulo State, Brazil: an environmental monitoring useful tool.*Environmental Earth Sciences* **75** : 1277. DOI: 10.1007/s12665-016-6042-2

Farifteh J, Van der Meer F, Atzberger C, Carranza EJM. 2007. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment* **110** : 59–78. DOI: 10.1016/j.rse.2007.02.005

Friedman JH. 1991. Multivariate adaptive regression splines. *The annals of statistics* **19** : 1–67

Gupta UC. 1967. A simplified method for determining hot-watersoluble boron in podzol soils. *Soil Science* **103** : 424–428

Hakansson I. 1990. A method for characterizing the state of compactness of the plough layer. *Soil and Tillage Research* **16** : 105–120. DOI: 10.1016/0167-1987(90)90024-8

Hanway JJ, Heidel H. 1952. Soil analysis method as used in Iowa State College, Soil Testing Laboratory. *Iowa Agriculture* **54** : 1–31

He Y, Huang M, Garcia A, Hernandez A, Song H. 2007. Prediction of soil macronutrients content using near-infrared spectroscopy. *Computers and Electronics in Agriculture* **58** : 144–153. DOI: 10.1016/j.compag.2007.03.011

Hively WD, McCarty GW, Reeves JB, Lang MW, Oesterling RA, Delwiche SR. 2011. Use of Airborne Hyperspectral Imagery to Map Soil Properties in Tilled Agricultural Fields. *Applied and Environmental Soil Science* **2011** : 1–13. DOI: 10.1155/2011/358193

HU X-Y. 2013. Application of Visible/Near-Infrared Spectra in Modeling of Soil Total Phosphorus. *Pedosphere* **23** : 417–421. DOI: 10.1016/S1002-0160(13)60034-X

Islam K, Singh B, McBratney A. 2003. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research* **41** : 1101–1114. DOI: 10.1071/SR02137

Jackson ML. 1973. Soil Chemical Analysis,(2nd Indian Print) Prentice-Hall of India Pvt. *Ltd. New Delhi* 38–336

Janik LJ, Merry RH, Skjemstad JO. 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture* **38** : 681. DOI: 10.1071/EA97144

Kodaira M, Shibusawa S. 2013. Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping.*Geoderma* **199** : 64–79. DOI: 10.1016/j.geoderma.2012.09.007

Kuang B, Mouazen AM. 2011. Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *European Journal of Soil Science* **62** : 629–636. DOI: 10.1111/j.1365-2389.2011.01358.x

Kusumo BH, Hedley CB, Hedley MJ, Hueni A, Tuohy MP, Arnold GC. 2008. The use of diffuse reflectance spectroscopy for in situ carbon and nitrogen analysis of pastoral soils. *Soil Research* **46** : 623. DOI: 10.1071/SR08118

Lindsay WL, Norvell WA. 1978. Development of a DTPA soil test for zinc, iron, manganese, and copper 1. *Soil science society of America journal* **42** : 421–428

Mahajan GR, Manjunath BL, Latare AM, D'Souza R, Vishwakarma S, Singh NP. 2016. Microbial and Enzyme Activities and Carbon Stock in Unique Coastal Acid Saline Soils of Goa. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences* **86** : 961–971. DOI: 10.1007/s40011-015-0552-7

Malmir M, Tahmasbian I, Xu Z, Farrar MB, Bai SH. 2019. Prediction of soil macro- and micro-elements in sieved and ground air-dried soils using laboratory-based hyperspectral imaging technique. *Geoderma***340** : 70–80. DOI: 10.1016/j.geoderma.2018.12.049

Martin PD, Malley DF, Manning G, Fuller L. 2002. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Canadian Journal of Soil Science* **82** : 413–422. DOI: 10.4141/S01-054

Minasny B, McBratney AB, Tranter G, Murphy BW. 2008. Using soil knowledge for the evaluation of mid-infrared diffuse reflectance spectroscopy for predicting soil physical and mechanical properties.*European Journal of Soil Science* **59** : 960–971. DOI: 10.1111/j.1365-2389.2008.01058.x

Moreira CS, Brunet D, Verneyre L, Sa SMO, Galdos M V., Cerri CC, Bernoux M. 2009. Near infrared spectroscopy for soil bulk density assessment.*European Journal of Soil Science* **60** : 785–791. DOI: 10.1111/j.1365-2389.2009.01170.x

Morellos A, Pantazi X-E, Moshou D, Alexandridis T, Whetton R, Tziotzios G, Wiebensohn J, Bill R, Mouazen AM. 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering***152** : 104–116. DOI: 10.1016/j.biosystemseng.2016.04.018

11

Mouazen AM, Kuang B, De Baerdemaeker J, Ramon H. 2010. 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* **158** : 23–31. DOI: 10.1016/j.geoderma.2010.03.001

Nawar S, Buddenbaum H, Hill J. 2015. Estimation of soil salinity using three quantitative methods based on visible and near-infrared reflectance spectroscopy: a case study from Egypt. *Arabian Journal of Geosciences* **8** : 5127–5140. DOI: 10.1007/s12517-014-1580-y

Nawar S, Buddenbaum H, Hill J, Kozak J. 2014. Modeling and mapping of soil salinity with reflectance spectroscopy and landsat data using two quantitative methods (PLSR and MARS). *Remote Sensing* **6** : 10813–10834. DOI: 10.3390/rs61110813

Nawar S, Buddenbaum H, Hill J, Kozak J, Mouazen AM. 2016. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil and Tillage Research* **155** : 510–522. DOI: 10.1016/j.still.2015.07.021

Nawar S, Reda M, Farag F, El-Nahry A. 2011. Mapping soil salinity in El-Tina plain in Egypt using geostatistical approach.

R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria

Ramaroson VH, Becquer T, Sa SO, Razafimahatratra H, Delariviere JL, Blavet D, Vendrame PRS, Rabeharisoa L, Rakotondrazafy AFM. 2018. Mineralogical analysis of ferralitic soils in Madagascar using NIR spectroscopy. *CATENA* **168** : 102–109. DOI: 10.1016/j.catena.2017.07.016

Reeves JB, McCarty GW. 2001. Quantitative Analysis of Agricultural Soils Using near Infrared Reflectance Spectroscopy and a Fibre-Optic Probe. *Journal of Near Infrared Spectroscopy* **9** : 25–34. DOI: 10.1255/jnirs.291

Reeves JB, McCarty GW, Meisinger JJ. 1999. Near Infrared Reflectance Spectroscopy for the Analysis of Agricultural Soils. *Journal of Near Infrared Spectroscopy* **7** : 179–193. DOI: 10.1255/jnirs.248

SAEKI K, TANABE K, MATSUMOTO T, UESAKA H, AMANO T, FUNATSU K. 2003. Prediction of Polyethylene Density by Near-Infrared Spectroscopy Combined with Neural Network Analysis. *Journal of Computer Chemistry, Japan* **2** : 33–40. DOI: 10.2477/jccj.2.33

SAS Institute. 2012. SAS User's guide. SAS Institute, IC, Cary.

Schirrmann M, Gebbers R, Kramer E. 2013. Performance of Automated Near-Infrared Reflectance Spectrometry for Continuous in Situ Mapping of Soil Fertility at Field Scale. *Vadose Zone Journal* **12** . DOI: 10.2136/vzj2012.0199

Shepherd KD, Walsh MG. 2002. Development of Reflectance Spectral Libraries for Characterization of Soil Properties. *Soil Science Society of America Journal* **66** : 988. DOI: 10.2136/sssaj2002.0988

Sidike A, Zhao S, Wen Y. 2014. Estimating soil salinity in Pingluo County of China using QuickBird data and soil reflectance spectra. *International Journal of Applied Earth Observation and Geoinformation* **26** : 156–175. DOI: 10.1016/j.jag.2013.06.002

Siebielec G, McCarty GW, Stuczynski TI, Reeves JB. 2004. Near- and Mid-Infrared Diffuse Reflectance Spectroscopy for Measuring Soil Metal Content. *Journal of Environment Quality* **33** : 2056. DOI: 10.2134/jeq2004.2056

Singh K. 2016. Microbial and Enzyme Activities of Saline and Sodic Soils. *Land Degradation & Development* **27** : 706–718. DOI: 10.1002/ldr.2385

Stenberg B, Viscarra Rossel RA, Mouazen AM, Wetterlind J. 2010. Visible and Near Infrared Spectroscopy in Soil Science. *Advances in Agronomy* , 163–215. DOI: 10.1016/S0065-2113(10)07005-7

Sun B, Zhou S, Zhao Q. 2003. Evaluation of spatial and temporal changes of soil quality based on geostatistical analysis in the hill region of subtropical China. *Geoderma* **115** : 85–99. DOI: 10.1016/S0016-7061(03)00078-8

Sun W, Zhang X, Zou B, Wu T. 2017. Exploring the Potential of Spectral Classification in Estimation of Soil Contaminant Elements. *Remote Sensing* **9** : 632. DOI: 10.3390/rs9060632

Suuster E, Ritz C, Roostalu H, Reintam E, Kolli R, Astover A. 2011. Soil bulk density pedotransfer functions of the humus horizon in arable soils. *Geoderma* **163** : 74–82. DOI: 10.1016/j.geoderma.2011.04.005

Udelhoven T, Emmerling C, Jarmer T. 2003. No Title. *Plant and Soil* **251** : 319–329. DOI: 10.1023/A:1023008322682

van Groenigen JW, Mutters CS, Horwath WR, van Kessel C. 2003. NIR and DRIFT-MIR spectrometry of soils for predicting soil and crop parameters in a flooded field. *Plant and Soil* **250** : 155–165. DOI: 10.1023/A:1022893520315

Vasques GM, Grunwald S, Sickman JO. 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. *Geoderma* **146** : 14–25. DOI: 10.1016/j.geoderma.2008.04.007

Vasques GM, Grunwald S, Sickman JO. 2009. Modeling of Soil Organic Carbon Fractions Using Visible–Near-Infrared Spectroscopy. *Soil Science Society of America Journal* **73** : 176. DOI: 10.2136/sssaj2008.0015

Vendrame PRS, Marchao RL, Brunet D, Becquer T. 2012. The potential of NIR spectroscopy to predict soil texture and mineralogy in Cerrado Latosols. *European Journal of Soil Science* **63** : 743–753. DOI: 10.1111/j.1365-2389.2012.01483.x

Viscarra Rossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **131** : 59–75. DOI: 10.1016/j.geoderma.2005.03.007

Vohland M, Besold J, Hill J, Frund H-C. 2011. Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy. *Geoderma***166** : 198–205. DOI: 10.1016/j.geoderma.2011.08.001

Volkan Bilgili A, van Es HM, Akbas F, Durak A, Hively WD. 2010. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *Journal of Arid Environments* **74** : 229–238. DOI: 10.1016/j.jaridenv.2009.08.011

Walkley A, Black IA. 1934. AN EXAMINATION OF THE DEGTJAREFF METHOD FOR DETERMINING SOIL ORGANIC MATTER, AND A PROPOSED MODIFICATION OF THE CHROMIC ACID TITRATION METHOD. *Soil Science* **37** : 29–38. DOI: 10.1097/00010694-193401000-00003

Wang J, Ding J, Abulimiti A, Cai L. 2018. Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS–NIR) spectroscopy, Ebinur Lake Wetland, Northwest China.*PeerJ* **6** : e4703. DOI: 10.7717/peerj.4703

Weng Y-L, Gong P, Zhu Z-L. 2010. A Spectral Index for Estimating Soil Salinity in the Yellow River Delta Region of China Using EO-1 Hyperion Data. *Pedosphere* **20** : 378–388. DOI: 10.1016/S1002-0160(10)60027-6

Weng Y, Gong P, Zhu Z. 2008. Soil salt content estimation in the Yellow River delta with satellite hyperspectral data. *Canadian Journal of Remote Sensing* **34** : 259–270

Wenjun J, Zhou S, Jingyi H, Shuo L. 2014. In situ measurement of some soil properties in paddy soil using visible and near-infrared spectroscopy. *PLoS ONE* **9** . DOI: 10.1371/journal.pone.0105708

Wijewardane NK, Ge Y, Wills S, Libohova Z. 2018. Predicting Physical and Chemical Properties of US Soils with a Mid-Infrared Reflectance Spectral Library. *Soil Science Society of America Journal* **82** : 722. DOI: 10.2136/sssaj2017.10.0361

Williams C, Steinbergs A. 1959. Soil sulphur fractions as chemical indices of available sulphur in some Australian soils. *Australian Journal of Agricultural Research* **10** : 340. DOI: 10.1071/AR9590340

Williams PC, Norris K. 1997. Near-Infrared Technology in the Agricultural and Food Industries. In: Williams, P.C., Norris K (ed)*Near-Infrared Technology: In the Agricultural and Food Industries* . Amer Assn of Cereal Chemists, 312

Wu Y, Chen J, Wu X, Tian Q, Ji J, Qin Z. 2005. Possibilities of reflectance spectroscopy for the assessment of contaminant elements in suburban soils. *Applied Geochemistry* **20** : 1051–1059. DOI: 10.1016/j.apgeochem.2005.01.009

Xu S, Zhao Y, Wang M, Shi X. 2018. Comparison of multivariate methods for estimating selected soil properties from intact soil cores of paddy fields by Vis–NIR spectroscopy. *Geoderma* **310** : 29–43. DOI: 10.1016/j.geoderma.2017.09.013

Yuan B-C, Li Z-Z, Liu H, Gao M, Zhang Y-Y. 2007. Microbial biomass and activity in salt affected soils under arid conditions. *Applied Soil Ecology* **35** : 319–328. DOI: 10.1016/j.apsoil.2006.07.004

Table 1 Descriptive statistics of the properties of the salt affected acid soils

Property	Mean	Standard error	Minimum
Full data set (n=372)	Full data set (n=372)	Full data set (n=372)	Full data set (n=372)
pH	5.42	0.05	3.47
EC (dS m^{-1})	6.91	0.3	0.1
SOC (%)	1.11	0.04	0.02
BD (Mg m^{-3})	1.44	0.01	0.89
N (kg ha^{-1})	201.76	4.78	37.63
P (kg ha^{-1})	20.05	1.02	0.59
K (kg ha^{-1})	975.7	46.24	2.84
Na (cmol(p$^+$) kg^{-1})	27.93	1.02	0.04
Ca (cmol(p$^+$) kg^{-1})	3.87	0.15	0.09
Mg (cmol(p$^+$) kg^{-1})	4.37	0.19	0.14
S (ppm)	23.23	0.99	0.63
Fe (ppm)	57.86	3.87	0.65
Mn (ppm)	7.24	0.41	0.06
Cu (ppm)	5.58	0.23	0.03
Zn (ppm)	1.97	0.07	0.01
B (ppm)	9.79	0.55	0.15
Calibration data set (n=260)	Calibration data set (n=260)	Calibration data set (n=260)	Calibration data set (n=260)
pH	5.45	0.07	3.51
EC (dS m^{-1})	7.08	0.36	0.1
SOC (%)	1.11	0.04	0.02
BD (Mg m^{-3})	1.44	0.02	0.89
N (kg ha^{-1})	200.89	5.8	37.63
P (kg ha^{-1})	20.02	1.27	0.59
K (kg ha^{-1})	990.29	55.84	2.84
Na (cmol(p$^+$) kg^{-1})	28.51	1.26	0.04
Ca (cmol(p$^+$) kg^{-1})	3.9	0.18	0.09
Mg (cmol(p$^+$) kg^{-1})	4.3	0.22	0.14
S (ppm)	24.36	1.24	0.8

Property	Mean	Standard error	Minimum
Fe (ppm)	61.54	5.05	0.7
Mn (ppm)	7.36	0.48	0.06
Cu (ppm)	5.75	0.28	0.06
Zn (ppm)	2.02	0.09	0.07
B (ppm)	9.74	0.63	0.15
Validation data set (n=112)	Validation data set (n=112)	Validation data set (n=112)	Validation data set (n=112)
pH	5.31	0.09	3.47
EC (dS m^{-1})	6.54	0.54	0.11
SOC (%)	1.13	0.07	0.07
BD (Mg m^{-3})	1.42	0.03	0.89
N (kg ha^{-1})	203.81	8.44	75.26
P (kg ha^{-1})	20.1	1.7	1.21
K (kg ha^{-1})	941.13	82.61	7.32
Na (cmol(p$^+$) kg^{-1})	26.61	1.7	0.68
Ca (cmol(p$^+$) kg^{-1})	3.8	0.27	0.16
Mg (cmol(p$^+$) kg^{-1})	4.54	0.39	0.39
S (ppm)	20.45	1.5	0.63
Fe (ppm)	48.6	4.73	0.65
Mn (ppm)	6.96	0.78	0.07
Cu (ppm)	5.17	0.38	0.03
Zn (ppm)	1.83	0.12	0.01
B (ppm)	9.91	1.07	0.63

pH, soil pH; EC, Electrical conductivity; SOC, Soil organic carbon; BD, Bulk density; N, Soil available nitrogen; P, Soil available phosphorus; K, Ammonium acetate extractable potassium; Na, Ammonium acetate extractable sodium; Ca, Ammonium acetate extractable calcium; Mg, Ammonium acetate extractable magnesium; S, soil available sulphur; Fe, soil available iron; Mn, Soil available manganese; Cu, Soil available copper; Zn, Soil available zinc; B, Soil available boron.

Table 2 Results on testing of equal means, variance, distribution and coefficient of variation for the calibration and validation dataset developed (value in the cell indicate the p value for the test and parameter)

	pH	EC	SOC	BD	N	P	K	Na	Ca	Mg	S	Fe	Mn	Cu
t-test	0.08	0.33	0.68	0.35	0.44	0.83	0.37	0.68	0.88	0.69	0.28	0.55	0.58	0.29
F-test	0.06	0.92	0.74	0.90	0.21	0.51	0.51	0.02	0.28	0.01	0.05*	0.72	0.20	0.07
Kolmogorov-Smirnov test	0.19	0.11	0.63	0.33	0.31	0.24	0.31	0.26	0.81	0.59	0.69	0.49	0.11	0.45
Flinger-Kileen test	0.08	0.22	0.13	0.30	0.01	0.17	0.26	0.28	0.37	0.27	0.21	0.11	0.20	0.29

pH, soil pH; EC, Electrical conductivity; SOC, Soil organic carbon; BD, Bulk density; N, Soil available nitrogen; P, Soil available phosphorus; K, Ammonium acetate extractable potassium; Na, Ammonium acetate extractable sodium; Ca, Ammonium acetate extractable calcium; Mg, Ammonium acetate extractable magnesium; S, soil available sulphur; Fe, soil available iron; Mn, Soil available manganese; Cu, Soil available copper; Zn, Soil available zinc; B, Soil available boron.

* and ** indicates significance at 0.05 and 0.01 level. Rest of the values are non-significant.

Table 3 Correlation matrix of the soil properties of the salt-affected soils (n=372)

	pH	EC	SOC	BD	N	P	K	Na	Ca	Mg	S	Fe
pH	1.00											
EC	0.37**	1.00										
SOC	-0.40**	-0.11*	1.00									
BD	-0.10^{ns}	-0.34**	-0.32**	1.00								
N	-0.03^{ns}	0.42**	0.19**	-0.47**	1.00							
P	0.19**	0.20**	-0.02^{ns}	-0.09^{ns}	0.17**	1.00						
K	0.33**	0.37**	0.001^{ns}	-0.34**	0.14*	0.13*	1.00					
Na	-0.15**	-0.17**	0.23**	0.06^{ns}	-0.26**	-0.09^{ns}	0.30**	1.00				
Ca	0.03^{ns}	0.005^{ns}	-0.05^{ns}	0.13*	-0.08^{ns}	0.03^{ns}	0.02^{ns}	-0.01^{ns}	1.00			
Mg	0.40**	0.29**	-0.18**	-0.22**	0.15**	0.09^{ns}	0.37**	-0.21**	0.45**	1.00		
S	-0.29**	-0.02^{ns}	0.34**	0.01^{ns}	-0.27**	-0.07^{ns}	0.05^{ns}	0.27**	0.05^{ns}	-0.19**	1.00	
Fe	-0.46**	-0.30**	0.41**	0.10^{ns}	-0.26**	-0.19**	-0.06^{ns}	0.30**	-0.04^{ns}	-0.27**	0.60**	1.00
Mn	0.13*	0.27**	0.04^{ns}	-0.28**	0.33**	0.001^{ns}	0.18**	-0.18**	-0.07^{ns}	0.17**	-0.17**	-0.18
Cu	0.23**	0.40**	0.06^{ns}	-0.37**	0.46**	0.13*	0.32**	-0.21**	-0.01^{ns}	0.26**	-0.20**	-0.2
Zn	0.10^{ns}	0.45**	0.04^{ns}	-0.29**	0.39**	0.13*	0.17**	-0.13*	-0.12*	0.12*	-0.09^{ns}	-0.10
B	0.44**	0.56**	-0.03^{ns}	-0.44**	0.39**	0.21**	0.35**	-0.24**	-0.12*	0.34**	-0.20**	-0.30

pH, soil pH; EC, Electrical conductivity; SOC, Soil organic carbon; BD, Bulk density; N, Soil available nitrogen; P, Soil available phosphorus; K, Ammonium acetate extractable potassium; Na, Ammonium acetate extractable sodium; Ca, Ammonium acetate extractable calcium; Mg, Ammonium acetate extractable magnesium; S, soil available sulphur; Fe, soil available iron; Mn, Soil available manganese; Cu, Soil available copper; Zn, Soil available zinc; B, Soil available boron.

* and ** indicates significance at 0.05 and 0.01 level and ns indicates not significant.

Table 4 Prediction accuracies of different multivariate models for calibration and validation to predict properties of the salt affected acid soils

Property	Model	Calibration (n=260)	Calibration (n=260)	Calibration (n=260)	Calibration (
		Optimization parameter(s)	R^2_c	MBE$_c$	RMSE$_c$
pH	MARS	-	0.90	-0.003	0.34
log10	PCR	PC = 16	0.75	-0.011	0.50
	PLSR	PC = 11	0.75	0.003	0.50
	RF	mtry = 94	0.63	-0.070	0.63
	SVR	γ =0.125 , c = 2	0.89	-0.008	0.35
EC	MARS	-	0.95	-0.05	1.45
BoxCox	PCR	PC = 17	0.78	-0.14	2.79
	PLSR	PC = 13	0.78	-0.05	2.74
	RF	mtry = 210	0.78	-0.26	2.99
	SVR	γ =0.0019, c = 128	0.92	-0.03	1.92
SOC	MARS	-	0.88	1.43	1.50
BoxCox	PCR	PC = 36	0.65	1.38	1.46
	PLSR	PC = 36	0.73	1.39	1.46
	RF	mtry = 140	0.34	1.38	1.51
	SVR	γ =0.0078 , c =32	0.83	1.39	1.44
BD	MARS	-	0.89	0.00	0.10
log10	PCR	PC = 23	0.55	0.00	0.20
	PLSR	PC = 18	0.60	0.00	0.19
	RF	mtry = 140	0.36	0.00	0.24
	SVR	γ =0.125 , c = 2	0.99	-0.01	0.05

Left margin rotated text:

Property	Model	Calibration (n=260)	Calibration (n=260)	Calibration (n=260)	Calibration (
N	MARS	-	0.87	0.10	35.16
log10	PCR	PC = 27	0.67	-2.55	53.73
	PLSR	PC = 16	0.68	-2.59	52.86
	RF	mtry = 210	0.33	-1.33	75.86
	SVR	γ =0.0019 , c = 256	0.75	-5.95	50.52
P	MARS	-	0.80	-1.12	10.75
log10	PCR	PC = 39	0.49	-3.14	16.55
	PLSR	PC = 27	0.54	-2.52	15.95
	RF	mtry = 94	0.28	-3.52	19.50
	SVR	γ =0.0019, c = 4	0.46	-3.09	16.14
K	MARS	-	0.88	-43.62	346.90
BoxCox	PCR	PC = 24	0.45	-104.36	692.04
	PLSR	PC = 16	0.48	-106.58	685.32
	RF	mtry = 140	0.39	-166.50	744.53
	SVR	γ =0.125, c = 1	0.95	-27.59	246.34
Na	MARS	-	0.87	-0.98	10.65
BoxCox	PCR	PC = 14	0.34	-3.29	20.34
	PLSR	PC = 8	0.29	-1.96	18.94
	RF	mtry = 140	0.38	-2.20	20.54
	SVR	γ =0.0019, c = 64	0.74	-0.66	15.67
Ca	MARS	-	0.81	-0.09	1.60
BoxCox	PCR	PC = 37	0.56	-0.28	2.15
	PLSR	PC = 23	0.51	-0.17	2.28
	RF	mtry = 210	0.34	-0.14	2.35
	SVR	γ =0.0078, c = 16	0.77	-0.06	1.49
Mg	MARS	-	0.89	-0.09	1.37
BoxCox	PCR	PC = 33	0.67	0.03	2.02
	PLSR	PC = 22	0.67	-0.09	2.09
	RF	mtry = 63	0.42	-0.20	2.65
	SVR	γ =0.0039, c = 64	0.85	-0.08	1.51
S	MARS	-	0.93	-0.08	6.58
BoxCox	PCR	PC = 48	0.81	-1.21	10.38
	PLSR	PC = 25	0.80	-0.62	10.04
	RF	mtry = 94	0.50	-1.89	15.26
	SVR	γ =0.0039, c = 128	0.96	-0.41	5.79
Fe	MARS	-	0.94	1.88	22.57
BoxCox	PCR	PC = 28	0.77	-2.48	46.91
	PLSR	PC = 36	0.80	-1.02	48.00
	RF	mtry = 94	0.92	-0.65	29.60
	SVR	γ =0.015, c = 32	0.92	0.18	26.39
Mn	MARS	-	0.93	-0.25	3.19
BoxCox	PCR	PC = 42	0.66	-1.10	6.44
	PLSR	PC = 24	0.74	-1.25	6.56
	RF	mtry = 63	0.17	-0.99	8.69
	SVR	γ =0.125 , c = 2	0.99	-0.72	2.16
Cu	MARS	-	0.92	-0.08	1.39
BoxCox	PCR	PC = 34	0.65	-0.09	2.69
	PLSR	PC = 18	0.59	-0.13	2.92
	RF	mtry = 140	0.46	-0.18	3.35
	SVR	γ =0.0019, c =256	0.77	-0.11	2.21

Property	Model	Calibration (n=260)	Calibration (n=260)	Calibration (n=260)	Calibration (
Zn	MARS	-	0.79	0.01	0.72
BoxCox	PCR	PC = 24	0.60	-0.05	0.94
	PLSR	PC = 21	0.66	-0.03	0.88
	RF	mtry = 210	0.43	0.02	1.08
	SVR	γ =0.0019, c = 256	0.78	-0.01	0.69
B	MARS	-	0.88	0.06	4.03
BoxCox	PCR	PC = 38	0.79	-0.63	5.16
	PLSR	PC = 19	0.77	-0.62	5.37
	RF	mtry = 210	0.59	-1.00	6.68
	SVR	γ = 0.001953, c = 128	0.81	-0.11	4.95

R^2, coefficient of regression; MBE, Mean bias error; RMSE, Root mean square error; RPD, ratio of performance to deviation.

MARS, Multivariate adaptive regression splines; PCR, Principal component regression; PLSR, Partial least square regression; RF, Random forest; SVR, Support vector regression.

pH, soil pH; EC, Electrical conductivity; SOC, Soil organic carbon; BD, Bulk density; N, Soil available nitrogen; P, Soil available phosphorus; K, Ammonium extractable potassium; Na, Ammonium extractable sodium; Ca, Ammonium extractable calcium; Mg, Ammonium extractable magnesium; S, soil available sulphur; Fe, soil available iron; Mn, Soil available manganese; Cu, Soil available copper; Zn, Soil available zinc; B, Soil available boron.

PC, Principal component; mtry, ; γ, sigma ; c, cost.

Figures

Figure 1 . Sampling locations in the present study (Geographic Coordinate System: GCS_WGS_1984).

Figure 2 . Average spectral signature of calibration and validation dataset of the salt-affected soils

Figure 3 . Soil salinity class-wise (a) raw reflectance spectra and (b) continuum-removed spectra

Figure 4 . Relationship between observed and predicted (a) soil pH, (b) electrical conductivity (EC), (c) Soil organic carbon (SOC), (d) Bulk density(BD), (e) Soil available nitrogen (N), (f) Soil available phosphorus (P), (g) Ammonium extractable potassium (K), (h) Ammonium extractable sodium (Na) (i) Ammonium extractable calcium (Ca), (j) Ammonium extractable magnesium (Mg), (k) Soil available sulphur (S), (l) Soil available iron (Fe), (m) Soil available manganese (Mn), (n) Soil available copper (Cu), (o) Soil available zinc (Zn), (p) Soil available boron (B); R^2, coefficient of regression; MBE, Mean bias error; RMSE, Root mean square error; RPD, ratio of performance to deviation; SVR, Support vector regression; PLSR, Partial least square regression; PCR, Principal component regression; RF, Random forest; Multivariate adaptive regression splines MARS

(a)

(b)