Arab J Geosci
DOI 10.1007/s12517-014-1580-y

ORIGINAL PAPER

Estimation of soil salinity using three quantitative methods based on visible and near-infrared reflectance spectroscopy: a case study from Egypt

Said Nawar · Henning Buddenbaum · Joachim Hill

Received: 28 January 2014 / Accepted: 14 August 2014

Abstract Soil salinization is a progressive soil degradation process that reduces soil quality and decreases crop yields and agricultural production. This study investigated a method that provides improved estimations of soil salinity by using visible and near-infrared reflectance spectroscopy as a fast and inexpensive approach to the characterisation of soil salinity. Soil samples were collected from the El-Tina Plain on the northwestern Sinai Peninsula in Egypt and measured for electrical conductivity (EC_e) using a saturated soil-paste extract. Subsequently, the samples were scanned with an Analytical Spectral Devices spectrometer (350–2,500 nm). Three spectral formats were used in the calibration models derived from the spectra and ECe: (1) raw spectra (R), (2) first-derivative spectra smoothened using the Savitzky–Golay technique (FD-SG) and (3) continuum-removed reflectance (CR). The spectral indices (difference index (DI), normalised difference index (NDI) and ratio index (RI)) of all of the band–pair combinations of the three types of spectra were applied in linear regression analyses with the EC_e. A ratio index that was constructed from the first-derivative spectra at 1,483 and 1,918 nm with an SG filter produced the best predictions of the EC_e for all of the band–pair indices (R^2=0.65). Partial least-squares regression models using the CR of the 400–2,500 nm spectral region resulted in R^2=0.77. The multivariate adaptive regression splines calibration model with CR spectra resulted in an improved performance (R^2=0.81) for estimating the EC_e. The results obtained in this study have potential value in the field of soil spectroscopy because they can be applied directly to the mapping of soil salinity using remote sensing imagery in arid regions.

Keyword Soil salinity · PLSR · MARS · Salinity index · Reflectance spectroscopy

Introduction

Soil salinization is a universal problem, especially in extensively irrigated areas that are poorly drained. Current estimations of the proportion of salt-affected soils in irrigated lands for several countries are 20 % for Australia, 27 % for India, 28 % for Pakistan, 50 % for Iraq and 30 % for Egypt (Stockle 2013). The accumulation of soluble salts in the root zone greatly affects plant growth, resulting in lower crop yields and adversely affecting the soil fertility (Qadir et al. 2007; Matinfar et al. 2011; Li et al. 2013). Therefore, reliable information on the nature and spatial extent of soil salinity is a prerequisite for restoring fertility and preventing further degradation.

Soil salinity is typically assessed by measuring the soil electrical conductivity (EC_e) in saturated paste extracts or by using extracts with different soil-to-water ratios (Amezketa 2006; Sonmez et al. 2008). Because conventional laboratory methods are time-consuming and relatively costly and soils have high spatial variability, particularly in terms of soil salinity (Akramkhanov et al. 2011), numerous remote sensing data have been used to identify and monitor salt-affected soils (Metternicht and Zinck 2003; Farifteh et al. 2007; Mulder et al. 2011), and several studies have been conducted in Egypt (Goossens et al. 1994; Ghabour 1997; Masoud and Koike 2006; Ibrahim and El Falaky 2013). Soil salinity is related to different parameters that have been derived from multispectral images using several soil and vegetation-based indices (Khan

S. Nawar (✉)
Department of GIS, Cartography and Remote Sensing, Jagiellonian University, Krakow 30-387, Poland
e-mail: snawar@gis.geo.uj.edu.pl

H. Buddenbaum · J. Hill
Environmental Remote Sensing and Geoinformatics, Trier University, 54286 Trier, Germany

Published online: 22 August 2014

et al. 2005; Fernández-Buces et al. 2006; Eldeiry and Garcia 2008; Badreldin et al. 2013; Othman et al. 2013). The efficiency of these methods has been restricted by various factors, such as the spectral and spatial resolution of the images and vegetation coverage (Farifteh et al. 2006; Metternicht and Zinck 2008). Because of the contiguous nature of hyperspectral signatures, hyperspectral remote sensing may overcome several shortcomings of multispectral remote sensing and enable the discrimination of fine differences between materials (Chang 2003; Campbell 2011). Therefore, hyperspectral remote sensing has been commonly used to study soil salinity (Dehaan and Taylor 2002; Tamas and Lenart 2006; Farifteh et al. 2007; Weng et al. 2008; Bilgili et al. 2011; Mashimbye et al. 2012).

Hyperspectral visible and near-infrared reflectance spectroscopy (VNIRRS) displays promise as a result of its performance, accuracy and cost effectiveness in the determination of most soil properties in laboratories (Shepherd and Walsh 2002; Waiser et al. 2007; Bilgili et al. 2010). Once the calibration models between soil reflectance spectra and soil variables have been established, they can be used to predict unidentified parameters. With the development of imaging spectrometry, reflectance spectroscopy has recently been applied at larger scales to map various soil properties, such as texture, clay content (Brown et al. 2006; Stenberg et al. 2010; Divya et al. 2013), mineralogy (Balasubramanian et al. 2012; Tiwari et al. 2013), organic carbon (Patzold et al. 2008; Stevens et al. 2010) and salinity (Dehaan and Taylor 2002; Farifteh et al. 2007; Weng et al. 2008).

The performance of the models is usually high and explains more than 81 % of the variability (Farifteh et al. 2007; Weng et al. 2008). Several regression methods based on VNIRRS have been used to estimate soil salinity, and partial least-squares regression (PLSR) is the most common (Farifteh et al. 2007; Weng et al. 2008; Bilgili et al. 2011). The PLSR approach has inference capabilities that are useful for modelling a probable linear relationship between the measured reflectance spectra and salt content in soils (Farifteh et al. 2007). The multivariate adaptive regression splines (MARS) method is considered a non-parametric method that estimates complex nonlinear relationships among independent and dependent variables (Friedman 1991), and it has been effectively applied in different fields (Luoto and Hjort 2005; Bilgili et al. 2010; Felicísimo et al. 2012; Samui 2012) and generally exhibits high performance results compared with other linear and non-parametric regression models, such as principal component regressions, classification and regression trees and artificial neural networks. Bilgili et al. (2010, 2011) used MARS to model soil salinity and reported that it provided better estimations for the EC_e of air-dried soils compared with the more frequently used PLSR method. Importantly, the choice of spectral pre-processing method (Rinnan et al. 2009) was found to be essential for the performance of multivariate calibration (Buddenbaum and Steffens 2012).

Because the correct method can enhance the predictive capability of the models (Bilgili et al. 2010, 2011; Mashimbye et al. 2012), this research aims to advance the use of reflectance spectroscopy for assessments of soil salinity based on a case study in the selected area. The objectives of the present study were as follows: (1) develop a soil salinity index based on various types of spectra pre-processing, (2) model and estimate soil salinity using linear (PLSR) and nonlinear (MARS) modelling methods based on soil spectra and (3) compare the accuracy of selected indices and models for estimating soil salinity.

Materials and methods

Study area

The study area is the El-Tina Plain, which is located on the northwestern Sinai Peninsula in Egypt between longitudes 32°20′35″ and 32°33′10″ E and latitudes 30°57′25″ and 31°04′28″ N and has an area of approximately 175 km^2 (Fig. 1). The El-Tina Plain is characterised by arid conditions, with annual rainfall ranging from 33.4 to 70.3 mm. The mean air temperatures range from 7.5 to 23.3 °C in winter and between 16.3 and 35.6 °C in summer. The mean evaporation is high and ranges from 3.6 to 7.3 mm/day. The land surface is nearly flat and ranges in elevation from below sea level to 5 m above sea level. The soil texture varies from loamy sand to clay, and the soil salinity varies from non-saline to highly saline. Nawar et al. (2011) classified the soils of the El-Tina Plain into the two orders Entisols and Aridisols, which include the eight subgroups Typic Aquisalids, Typic Haplosalids, Aquic Torriorthents, Typic Torriorthents, Aquic Torripsamments, Typic Torripsamments, Gypsic Aquisalids and Gypsic Haplosalids.

Soil sampling and analysis

Ninety-four soil samples (0–20 cm) were collected from the study area based on a previous soil salinity mapping (Nawar et al. 2011). Satellite navigation measurements with a Garmin 12XL were used to obtain the geographic locations of the soil samples. The collected soil samples were air-dried, crushed and passed through a 2-mm sieve.

The resulting fine earth (>2 mm) was retained for analysis, and the particle size distribution was measured with the pipette method (Kilmer and Alexander 1949). The soil reaction (pH) was measured in a 1:2.5 soil–water suspension, and the EC was measured in a soil paste extract according to the method of Jackson (1973). The soluble salts in the soil paste extract were determined using the gravimetric method. The cation exchange capacity and exchangeable sodium percentage

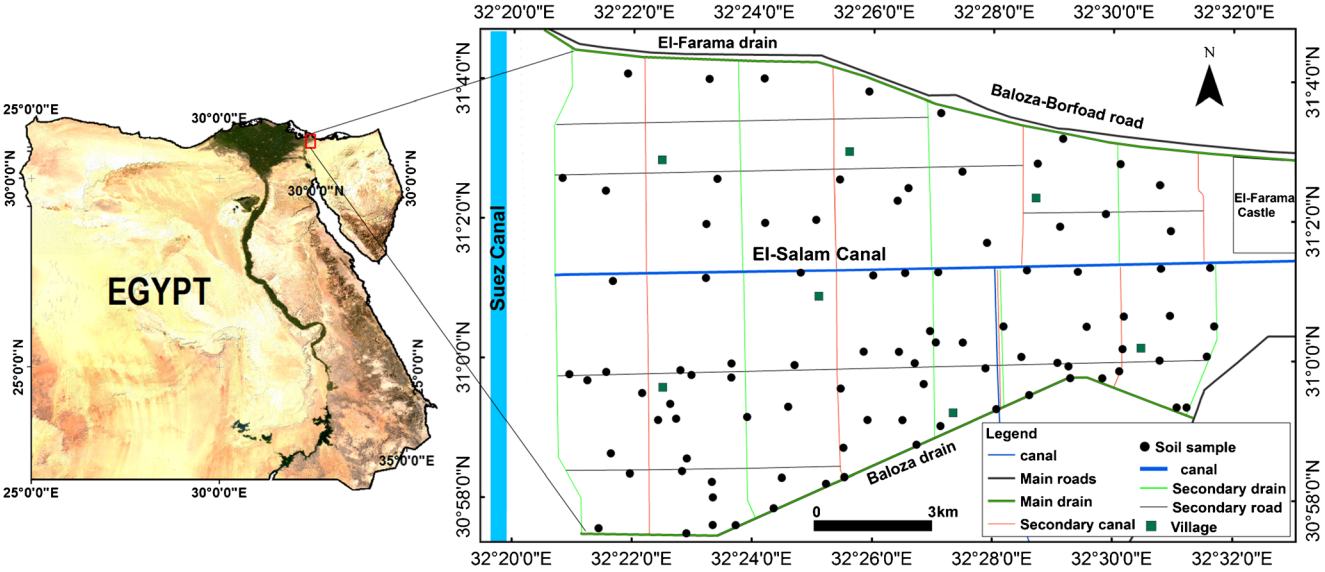

Fig. 1 Study area and soil sample locations

(ESP) were determined according to the method of Richards (1954). Organic matter content was determined using the modified Walkley and Black method (Page et al. 1982).

Spectral measurements

Soil reflectance spectra were collected using a portable spectroradiometer (FieldSpec-FR, Analytical Spectral Devices), which measures reflectance over a range from 350 to 2,500 nm with a resolution of approximately 10 nm and sampling interval of 3 nm in the short-wave infrared domain. The measurements were conducted in a dark laboratory environment. Plastic dishes were used to contain the soil samples, which were levelled off to a thickness of 2.0 cm (Mouazen et al. 2007). A tungsten quartz halogen lamp was set 45 cm away from the samples, and it illuminated the dishes for the initial measurements.

To measure each sample's spectral reflectance, the soil samples were measured with a viewing angle of 30° from the nadir at a distance of 15 cm. The central area of each sample was targeted when measuring the spectral reflectance, and three spectral measurements were performed. The final spectra measurement was attained by averaging each of the curves. The reflected radiance from a white reference panel with known reflectance was recorded before scanning each sample. To calculate the absolute reflectance of the samples, the radiance from each sample was divided by the radiance from the white reference panel and multiplied by the reflectance of the reference panel.

Pre-processing transformations

Three types of spectra (Fig. 2) were used to develop the models used to estimate the EC_e: (1) the raw reflectance spectra (R), (2) first-derivative spectra smoothened with the SG smoothing technique (Savitzky–Golay technique (FD-SG)) and (3) continuum removed reflectance (CR).

The FD-SG is a method that eliminates the baseline from spectra and enhances the absorption features. The FD-SG was calculated using a Savitzky–Golay smoothing technique (Savitzky and Golay 1964). In their study, Vasques et al. (2008) found that the Savitzky–Golay derivative procedure consistently yielded the best transformations in pre-processing. We used a second-order polynomial, which was fit to 101-point-width spectral windows. As noted by Ertlen et al. (2010), using derivatives of the spectra may allow for relevant information to be extracted from the near-infrared range.

To calculate the CR (Kokaly and Clark 1999), a convex hull (which is the continuum line) is fitted to the spectrum, and the spectrum is then divided at each wavelength by the hull. The wavelength regions that lie upon the convex hull (e.g. the first and last bands) receive a value of 1, whereas regions that lie within the absorption bands receive values between 0 and 1. This means that the CR will minimise any brightness differences and emphasise the spectra's absorption bands. The ENVI 5.0 software (Exelis Visual Information Solutions 2012) was used to perform the continuum removal.

Data analysis

Spectral indices

For an exploratory analysis of the relationship between the soil EC_e and corresponding reflectance spectra, 2-D correlograms of the coefficients of determination (R^2) were

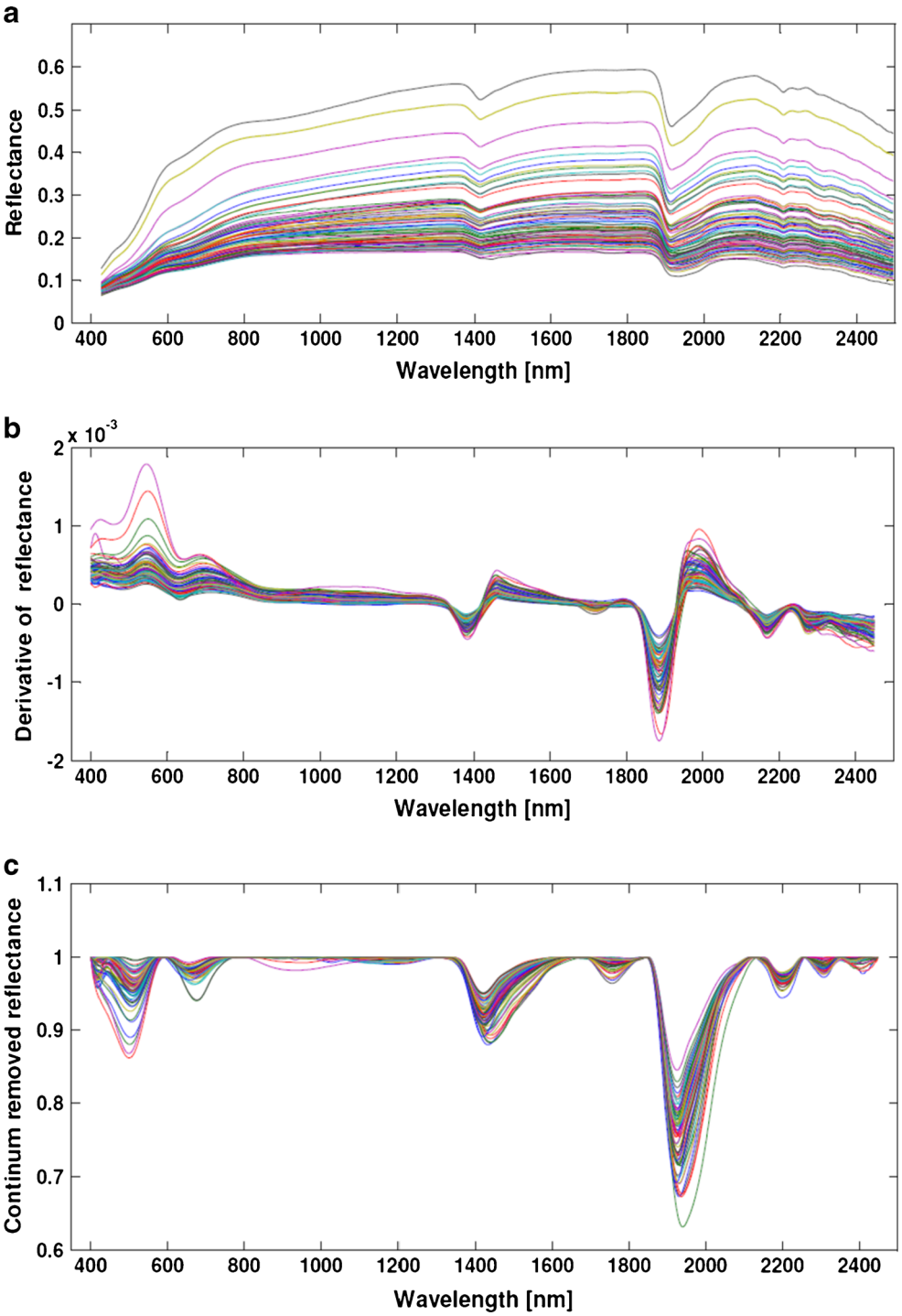

Fig. 2 Raw reflectance spectra (R) (**a**), first-derivative spectra smoothened with SG (FD-SG) (**b**) and continuum removed spectra (CR) (**c**) of 70 soil samples

calculated using a linear regression of the difference indices (DI) in Eq. (1), normalised difference indices (NDI) in Eq. (2) and ratio indices (RI) in Eq. (3) against the EC_e, including all of the possible band–pair combinations of narrow bands ranging between 400 and 2,500 nm. The index values were subsequently correlated with the EC_e from the soil samples to obtain the most effective spectral index for estimating the EC_e. The R_i and R_j values were used as the reflectances at waveband $i$nm and waveband $j$nm, respectively. The analyses were performed using MATLAB 8.0 software (MathWorks 2011).

$$DI(R_i, , R_j) = R_i - R_j \quad (1)$$

$$\mathrm{NDI}(\mathrm{R}_i\,,\,,\,\mathrm{R}_j) = \frac{(\mathrm{R}_i - \mathrm{R}_j)}{(\mathrm{R}_i + \mathrm{R}_j)} \quad (2)$$

$$\mathrm{RI}(\mathrm{R}_i\,,\,,\,\mathrm{R}_j) = \frac{\mathrm{R}_i}{\mathrm{R}_j} \quad (3)$$

Partial least-squares regression (PLSR)

The PLSR method is a popular regression method that is often applied in chemometrics, and it was introduced and statistically described in Geladi and Kowalski (1986) and Wold et al. (2001). PLSR is frequently used to conduct quantitative spectral analyses (Bilgili et al. 2011; Farifteh et al. 2007), and the algorithm uses a linear multivariate model to relate the predictor (X) and response (Y) variables and select successive orthogonal (latent) factors, thereby maximising the X and Y covariance, or the covariance between the spectra (X) and a measured soil property (Y). Compared with multiple linear regression, PLSR is an appropriate method for managing data with severe co-linearity in the independent variables, particularly in cases where the sample size is small. To determine the number of latent factors when using PLSR, leave-one-out cross-validation (LOOCV) was used (Efron and Tibshirani 1994) to prevent over- or under-fitting the data, which may produce models with poor performance. The root mean squared error (RMSE) of the predictions along with the coefficient of determination (R^2) were ascertained to identify the optimal cross-validated calibration model. Generally, the model with the highest cross-validated R^2 value and lowest RMSE value is selected. The PLSR process was performed using MATLAB 8.0 software.

Multivariate adaptive regression splines (MARS)

Developed by Friedman (1991), MARS is a non-parametric regression technique used for fitting the relationship between dependent and independent variables via the splines theory. Recently, MARS has been applied as a regression method in several disciplines, such as estimations of soil salinity (Bilgili et al. 2010, 2011), predictions of soil pH, organic carbon and clay content (Shepherd and Walsh 2002), simulations of pesticide transport in soils (Yang et al. 2003) and mappings of landslide susceptibility (Felicísimo et al. 2012). This technique has been shown to perform consistently better than traditional statistical methods. The MARS analysis uses basis functions to model the predictor and response variables (Hastie et al. 2001). MARS creates basis functions that can serve as new predictor variables in modelling. The basis functions are created by splitting the data into splines (or sub-regions), which have varied interval ending knots, at the points where the regression coefficients are altered. By utilising adaptive piecewise linear regressions, this process is also effective for the data in every sub-region. Each of the basis functions created by MARS may include linear combinations, nonlinear interaction factors and variable interaction factors of the second or third order. The number of basis functions and knots are defined using a forward stepwise process to choose certain spline basis functions. Next, backward stepwise algorithm elimination is applied until the best set is found to a smoothing procedure, which gives the final MARS approximation a particular level of continuity (Friedman 1991).

The final MARS model is composed of a group of basis functions that are defined based on the generalised cross-validation (GCV) criterion (Vidoli 2011). Predictions are improved, and over-fitting is avoided through the use of GCV, which involves the one-at-a-time removal of repetitious basis functions through a backward stepwise procedure. After applying the GCV, the basis functions that can be left out of the model and those that should be incorporated into the model become clear. The MARS analysis is performed using the ARESLab toolbox (Jekabsons 2011) with selected adaptations from the MATLAB 8.0 software.

Prediction accuracy

The R^2, RMSE and ratio of performance to deviation (RPD) values were used to assess the performance of the soil salinity prediction models. The RPD was classified into three classes by Chang et al. (2001): Category A (RPD>2) includes models that accurately predict a given property; category B (1.4< RPD<2) has limited predictive ability, and category C (RPD <1.4) has no predictive ability.

Results

Salinity parameters

The chemical analysis results for the 94 samples (Table 1) show that the soil salinity is high and exhibits a broad range from 3.3 to 166.8 dS/m. The predominant anion in the soil is Cl^- (74.4 % of the total anions), and the predominant cation is Na^+ (70.5 % of the total cations). The correlation coefficients between Cl^- and Na^+ and between Cl^- and Mg^{2+} are 0.98 and 0.91, respectively (Table 2). Moderate correlations were also found between EC_e and Cl^-, Na^+ and Mg^{2+} (correlation coefficients of

Table 1 Descriptive statistics of the soil parameters

	EC_e	$CaCO_3$	OM	Clay	Silt	Sand	pH	Na^+	Mg^{++}	Cl^-	SO_4^-	ESP
	dSm^{-1}			%					$meql^{-1}$			%
Min	3.30	0.00	0.00	0.00	0.50	16.00	7.10	317.00	172.00	360.00	160.00	17.60
Max	166.80	21.90	2.30	54.30	34.60	100.00	8.50	1,460.00	450.00	1,597.00	544.00	58.00
Mean	33.06	2.97	0.83	27.22	20.81	50.68	7.86	918.27	344.31	971.60	338.41	39.31
SD	31.31	3.07	0.52	16.77	10.21	26.63	0.29	336.72	62.07	339.65	78.09	11.61
CV (100)[a]	94.72	103.26	63.03	61.62	49.04	52.55	3.7	36.67	18.03	23.07	29.55	29.55

[a] CV=SD×100/mean

0.49, 0.50 and 0.45, respectively). These results indicate that the dominant soluble salts are NaCl and $MgCl_2$.

Spectral characteristics of salt-affected soils

The spectral reflectances of the selected 70 soil samples, R, FD-SG and CR are plotted in Fig. 2. Based on the saline soil classification (Nawar et al. 2011), the soil sample spectra were divided into five classes. In each class, an average spectrum was calculated (Fig. 3a), and the plots show that the reflectance curves display two deep absorption regions at 1,415 and 1,915 nm and several weak absorption regions near 494, 673, 1,748, 2,207 and 2,385 nm. In comparison, the absorption region depth varies with the level of soil salinity (Fig. 3b). These features suggest that the soil moisture content increases with increasing salinisation. Because the salts in this area are represented primarily by highly hygroscopic salts such as $MgCl_2$, which can absorb water vapour, increases in the soil moisture content can occur. These results are consistent with those of Weng et al. (2008) and Sidike et al. (2014).

Relationship between soil salinity and the spectral indices

A close correlation between the DI, NDI and RI and soil salinity occurred mostly in the visible and near-infrared ranges (Fig. 4). Although the performance of the three spectral indices as predictors of soil salinity appeared to vary with wavelength, constant basic patterns have emerged. Wavelength combinations in the 478–1978 nm region for R spectra (Fig. 4a) showed a strong correlation between the RI and soil salinity. For the DI and NDI, good wavelength combinations were observed in the 673–1963 nm region and 523–1963 nm region, which have R^2 values of 0.59 and 0.65, respectively (Table 3). The spectral indices that included the DI (R_{673}, R_{1963}), NDI (R_{523}, R_{1963}) and RI (R_{478}, R_{1978}) showed the best performance of the three combinations of R spectra. The RI (R_{478}, R_{1978}) displayed a medium correlation with soil salinity and had the highest R^2 (0.65).

Table 2 Pearson's correlation coefficients between measured soil variables

	pH	EC_e	$CaCO_3$	OM	Clay	Silt	Sand	Na^+	Mg^{++}	Cl^-	SO_4^-	ESP
pH	1.00											
EC_e	–0.44[a]	1.00										
$CaCO_3$	–0.35 [a]	0.00	1.00									
OM	–0.66[a]	0.38[a]	0.40[a]	1.00								
Clay	–0.58[a]	0.55[a]	0.40[a]	0.61[a]	1.00							
Silt	–0.54[a]	0.49[a]	0.20[b]	0.55[a]	0.88[a]	1.00						
Sand	0.59[a]	–0.53[a]	–0.32[a]	–0.59[a]	–0.97[a]	–0.95[a]	1.00					
Na^+	–0.40[a]	0.50[a]	0.25[b]	0.48[a]	.079[a]	0.84[a]	–0.82[a]	1.00				
Mg^{++}	–0.40[a]	0.45[a]	0.26[b]	0.52[a]	0.80[a]	0.87[a]	–0.85[a]	0.89[a]	1.00			
Cl^-	–0.39[a]	0.49[a]	0.23[b]	0.48[a]	0.78[a]	0.85[a]	–0.83[a]	0.91[a]	0.98[a]	1.00		
SO_4^-	–0.51[a]	0.41[a]	0.40[a]	0.55[a]	0.70[a]	0.71[a]	–0.72[a]	0.75[a]	0.75[a]	0.67[a]	1.00	
ESP	–0.51[a]	0.49[a]	0.29[a]	0.54[a]	0.84[a]	0.90[a]	–0.88[a]	0.85[a]	0.95[a]	0.92[a]	0.78[a]	1.00

[a] Significant at the 0.01 probability level

[b] Significant at the 0.05 probability level

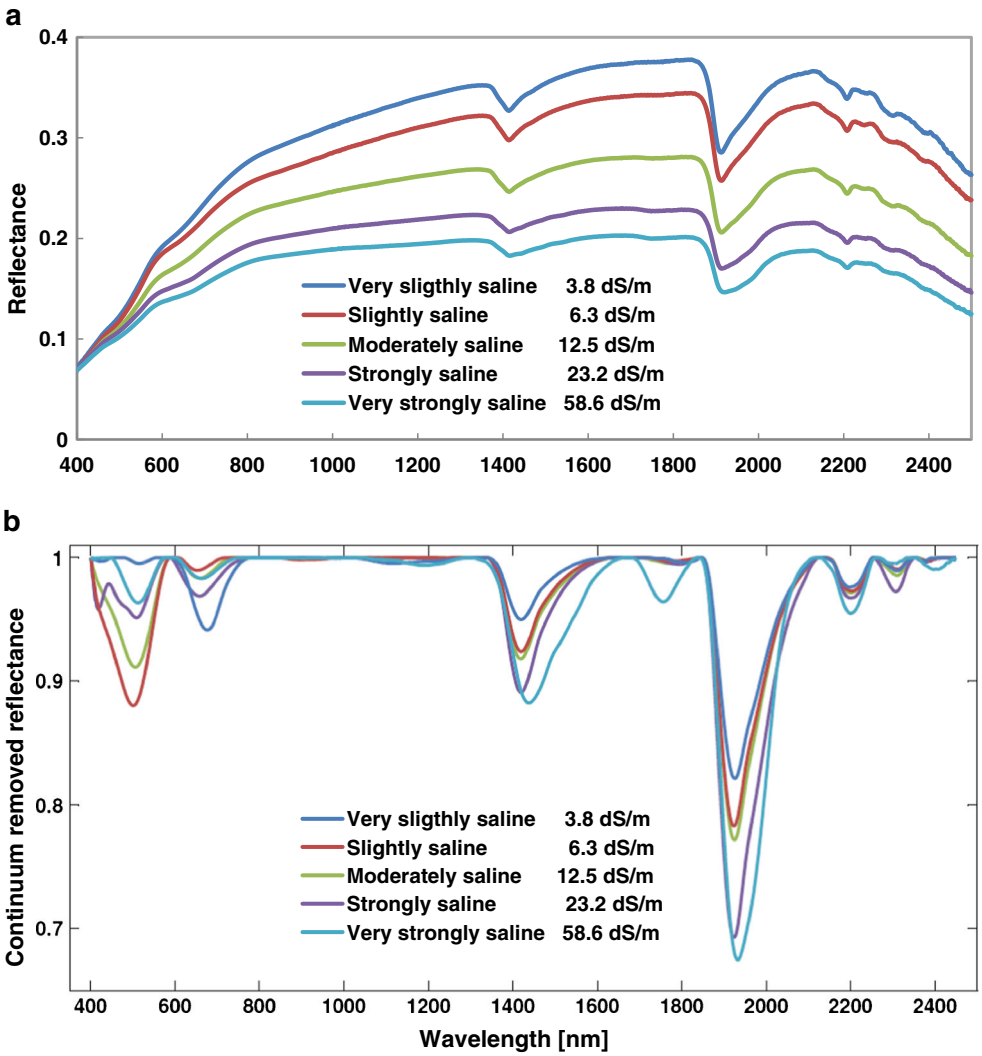

Fig. 3 Raw reflectance spectra (R) (**a**); continuum-removed spectra (CR) (**b**) of five salinity classes

An additional analysis was performed for the relationship between the EC_e and indices constructed from two FD-SG and CR spectra in all of the possible band combinations in the 400–2,500 nm range (Fig. 4b and c). The results indicated a good correlation between the EC_e and band–pair combinations of first derivative spectra located in the 1,483–1,918 nm region. The results based on the CR spectra did not perform, as well as those based on the FD-SG. Among the FD-SG wavelength combinations for the RI, NDI, and DI, those that performed the best were from 1,483 to 1,918, 1,498–1,918, and 1,378–1,933 nm, respectively. The RI (FD-$SG_{1,483}$, FD-SG_{1918}), NDI (FD-SG_{1498}, FD-SG_{1918}) and DI (FD-SG_{1378}, FS-DG_{1978}) were the best indices for FD-SG. RI (CR_{572}, CR_{2222}), DI (CR_{554}, CR_{1398}) and NDI (CR_{638}, CR_{786}) were the best indices for the CR spectra. The RI (FD-SR_{1483}, FD-SR_{1918}) showed strong correlations with the soil salinity, with R^2 values of 0.80 for the first derivative spectra with an SG filter (Table 3).

Performance of PLSR

Table 4 shows the calibration models derived from applying the PLSR model to various pre-processing methods. By analysing the R^2 and RMSE values that were derived from the calibration models, it was possible to estimate the EC_e after identifying the optimal number of latent factors. It is apparent from Table 4 that the PLSR models based on different spectral pre-processing methods yielded significant differences without the use of any auxiliary variables (Fig. 5). The soil clay contents that showed significant correlations with the EC_e were combined with the soil reflectance as an auxiliary predictor to improve the cross-validation predictions for the modelled EC_e (Fig. 6). Estimations of the EC_e were improved by up to 8.7 %, producing higher R^2 and RPD values (0.75 and 2.02, respectively) than with the use of soil spectra alone (R^2=0.69 and RPD=1.80). The

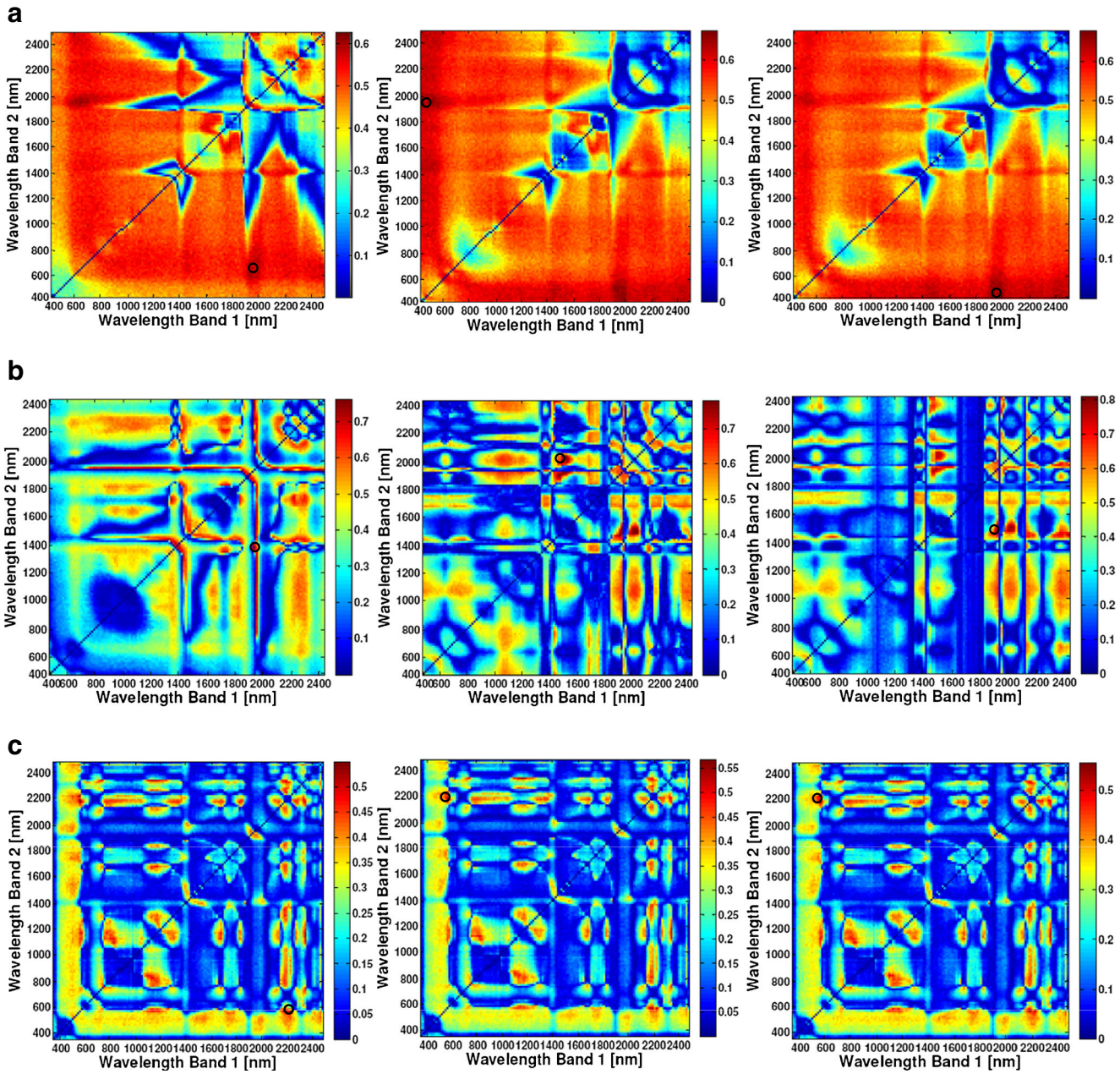

Fig. 4 The 2-D correlogram of R^2 between ECe and spectral indices DI (*left*), NDI (*middle*) and RI (*right*) based on R (**a**), FD-SG (**b**) and CR (**c**) spectra

improvement was small, however, and reflects the moderate degree of correlation between the clay content and EC_e (Table 2).

Performance of MARS

The MARS models based on the three types of spectra (R, FD-SG and CR) displayed significant differences (Fig. 7). The MARS model provided good correlations between the soil spectra and the soil EC_e. Table 5 summarises the MARS cross-validation statistics for the EC_e indicators. The best estimation was achieved using the CR spectra, and the results showed that 16 basis functions gave the best performance for predicting the soil EC_e.

Testing of the soil salinity estimation models

The results of the spectral indices showed that indices based on the FD-SG spectra performed better than those composed of R and CR spectra. Consequently, a difference index based on the band pair of first-derivative spectra at 1,483 and 1,918 nm with the FD-SG method yielded the best results of

Table 3 Quantitative relationship of EC_e (y) with different spectral indices (x)

Spectral index	Regression equation	R^2
DI (R_{673}, R_{1963})	$y=-315.49x+34.209$	0.59
NDI (R_{523}, R_{1963})	$y=113.49x+51.754$	0.65
RI (R_{478}, R_{1978})	$y=93.044x-24.099$	0.65
DI (FD-SG_{1378}, FD-SG_{1933})	$y=99.842x+41.806$	0.74
NDI (FD-SG_{1498}, FD-SG_{1918})	$y=155.6x-41.407$	0.77
RI (FD-SG_{1483}, FD-SG_{1918})	$y=135.05x+72.657$	0.80
DI (CR_{572}, $CR_{2,222}$)	$y=1596.2x-5.9836$	0.53
NDI (CR_{572}, CR_{2222})	$y=-3133.9x-5.7996$	0.53
RI (CR_{572}, CR_{2222})	$y=1540.8x-1546.4$	0.53

all spectral indices, with R^2 of 0.80, RMSE of 6.20 and RPD of 2.06. The RMSE and RPD values proved a better fit for the models that were based on the application of PLSR and MARS techniques. Using the CR spectra yielded the best calibration models with respect to estimates of the EC_e, which generated R^2 values of 0.77, 0.81 and 0.81; RMSE values of 7.20, 6.50 and 6.55; and RPD values of 2.07, 2.29 and 2.30 for PLSR without auxiliary variables, PLSR with clay content as the auxiliary predictor and MARS, respectively.

Discussion

Numerous studies have examined the relationship between sensitive spectral wavebands and soil salinity content (Farifteh et al. 2007; Weng et al. 2008; Bilgili et al. 2011; Mashimbye et al. 2012). For example, Farifteh et al. (2007) reported that the best-performing bands for different scales ranging from the field, experiment and image datasets were found in the NIR and SWIR spectral regions. Using a linear regression model with an unprocessed spectral band found at 2,257 nm demonstrates the possibilities for estimating the EC_a. The reflectance spectra within the NIR and SWIR regions were viewed as the best spectral region for estimating the EC_a (Farifteh et al. 2007; Weng et al. 2008; Mashimbye et al. 2012). It was found that certain spectral indices correlated well with the EC_e, including the first-derivative spectral regions at 1,483–1,918, 1,498–1,918 and 1,378–1,933 nm and spectral indices that were composed of wavebands in the 523–1,963 nm spectral range. This suggests that for EC_e evaluations in arid and semi-arid environments, the raw spectral reflectance in the 478–1,978 nm waveband and first-derivative spectra in the 1,483–1,918 nm waveband should provide adequate sensitivity.

The results of this study show that compared with the difference and normalised indices, the ratio index might be the optimal spectral index type for estimating the soil EC_e. Furthermore, the ratio indices composed of FD-SG spectra may further improve the estimation of soil salinity because the RI (FD-$SG_{1,483}$, FD-SG_{1918}) provided the best estimation results, which might be attributable to the effective removal of specific interfering factors, such as soil particle size, in the spectra pre-treatment. Moreover, in the RI spectral index (FD-SG_{1483}, FD-SG_{1918}) for the soil EC_e, the bands at 1,483 and 1,918 nm are found in the NIR reflectance region where the reflectance is primarily affected by the water content. The strong co-variation between the EC_e and NIR reflectance of the soil samples can easily be visualised in the form of numerous peaks of the PLSR coefficients (Figs. 5 and 6). The NIR band appears to be a good indicator of soil salinity, and this is especially true for the first-derivative and continuum-removed spectral forms. This result is largely consistent with the findings of earlier studies (Farifteh et al. 2007; Weng et al. 2008; Bilgili et al. 2011; Mashimbye et al. 2012). However, the applicability of the RI (FD-SG_{1483}, FD-SG_{1918}) in estimating the EC_e must be assessed in future investigations.

The PLSR and MARS methods were also used to estimate the EC_e in the present study. However, the accuracy of the PLSR and MARS models is generally affected by variations in the soil texture and moisture content (Farifteh et al. 2007), and a successful pre-processing method for the spectral data may improve the performance of such models (Vasques et al. 2008; Rinnan et al. 2009). Compared with published results that used PLSR (R^2=0.80, 0.74 and 0.74, which was reported by Farifteh et al. 2007; Bilgili et al. 2011 and Mashimbye et al. 2012, respectively) and MARS (R^2=0.39 and 0.77, which was reported by Bilgili et al. 2010 and 2011, respectively), the EC_e calibration models used in the present study may be more reliable and precise in modelling the EC_e. The estimation quality of the EC_e was dependent on the different pre-treatment and calibration methods. For instance, the MARS model without spectral pre-treatment produced better results than did the PLSR (Table 5). Similarly, averaging the spectra consistently optimised the calibrations built using PLSR.

Table 4 Cross-validation results of PLSR models of EC_e with different spectra

	Pre-processing	Number of factors	R^2	RMSE
PLSR	R	6	0.69	8.30
	FD-SG	6	0.75	7.40
	CR	6	0.77	7.20
PLSR+AP	R	6	0.75	7.40
	FD-SG	6	0.77	7.20
	CR	6	0.81	6.50

R Raw reflectance spectra, *FD-SG* first-derivative spectra smoothing with a Savitzky–Golay filter, *CR* continuum removed spectra, *AP* auxiliary predictor (clay content)

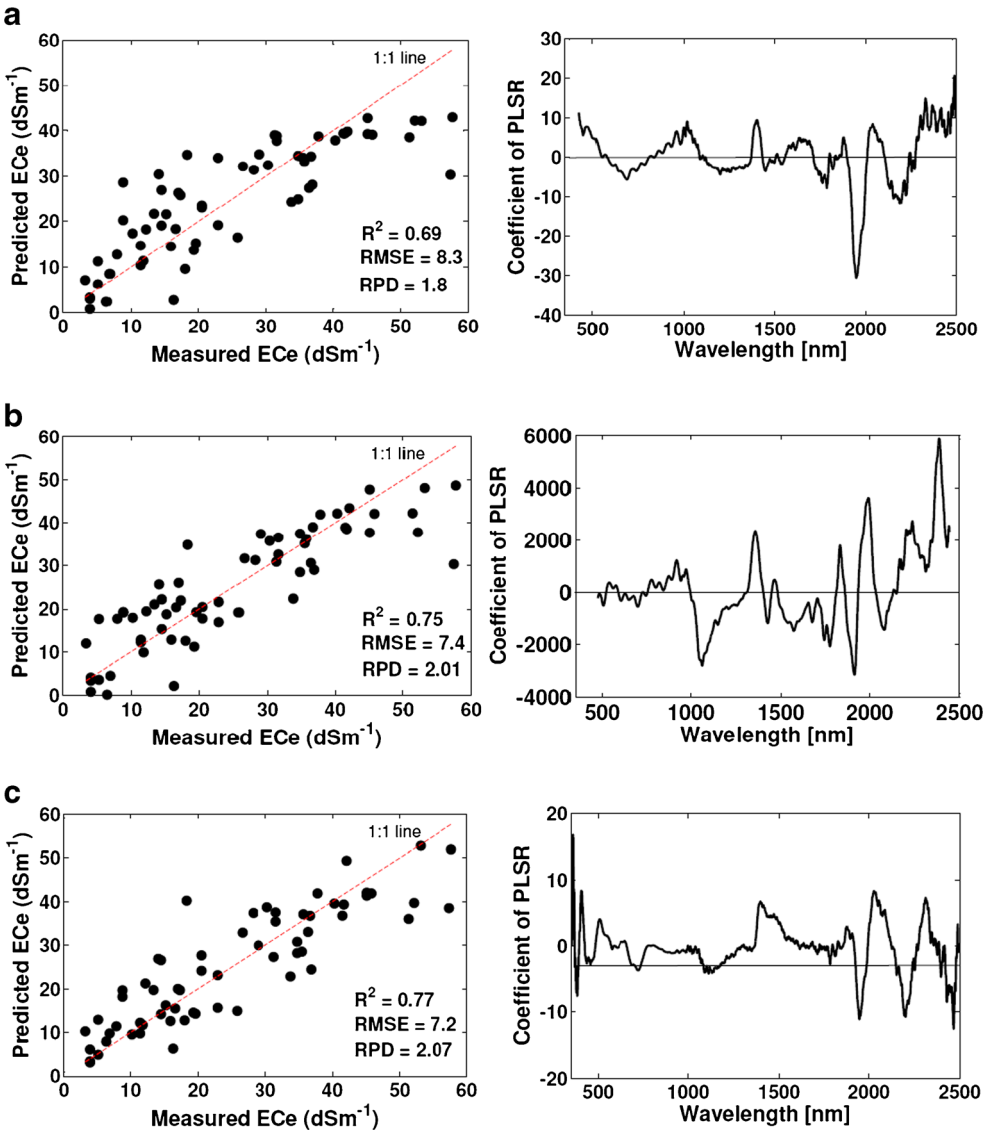

Fig. 5 PLSR models for R (**a**), FD-SG (**b**) and CR (**c**) spectra (*left*) with regression coefficients for PLSR (*right*)

Averaging smooths the spectra, removes noise and may eliminate factors that lead to nonlinearities (Bilgili et al. 2010). The MARS method typically yields better results in a nonlinear relationship, whereas the PLSR model fits linear relationships only (Brown et al. 2006).

The MARS method provided better estimation results depending on the different pre-treatment method that was used. The best model performance (RPD=2.30) was obtained using MARS for continuum-removed spectra (CR). Continuum removal generally emphasises the spectral absorption features (Clark and Roush 1984) and may improve the estimations. Weng et al. (2008) recorded improvements in the estimation of the soil salinity content using reflectance spectroscopy after implementing CR pre-processing on the soil spectra.

This study found that CR generally improved these estimations compared with estimations from the raw spectra. Overall, using any of the pre-treatment methods produced better results compared with the results from using the raw spectra (Figs. 5 and 6). A significant correlation was found between the soil clay content and EC_e (r=0.55), which is shown in Table 2. The clay content combined with the soil reflectance spectra improved the estimations of the EC_e (Fig. 6) by up to 8.7 % and produced higher R^2 and RMSE values (0.75 and 7.45, respectively) compared with estimations from only the raw spectra (R^2=0.69, RMSE= 8.29). This improvement was small, however, and reflected the moderate degree of correlation between the clay content parameter and soil EC_e (Table 2). Bilgili et al. (2011) and Brown et al. (2006) also used auxiliary predictors in addition to the soil reflectance for estimations of different soil properties. For example, including the soil salinity produced better results than when the soil spectra

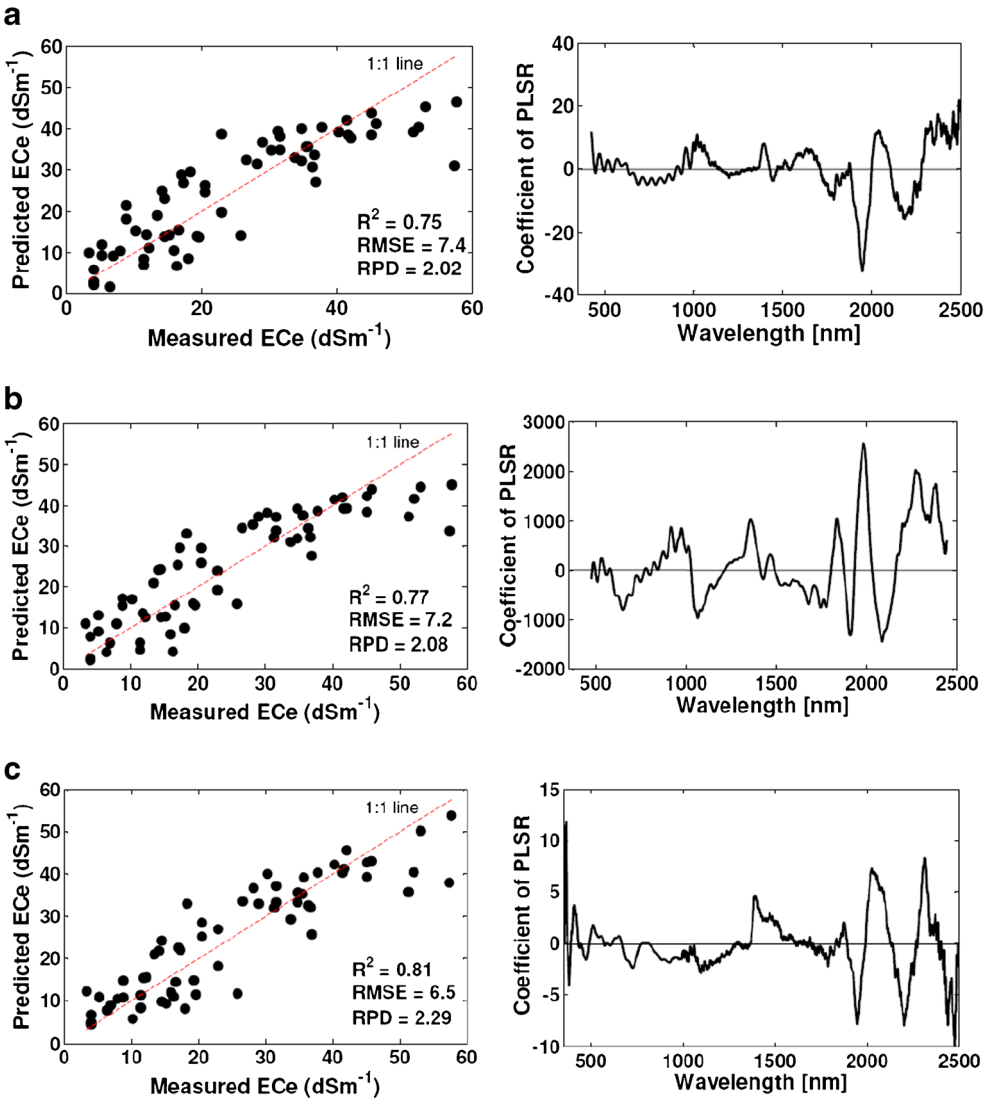

Fig. 6 PLSR models for R (**a**), FD-SG (**b**) and CR (**c**) spectra including the auxiliary predictor (clay content) (*left*) with regression coefficients for PLSR (*right*)

was used alone. Bilgili et al. (2011) combined topographical variables that showed significant correlations with the EC_e, such as elevation, and included soil reflectance to increase the model performance. The performance accuracy was improved by up to 12 % and yielded higher R^2 and RPD values (0.74 and 1.89, respectively) than were found using only the spectra or topographical indicators (R^2= 0.21, RMSE=5.4 dSm-^{1}).

The predictions from the MARS and PLSR models were more accurate than those of the RI (FD-SG_{1483}, FD-SG_{1918}) in estimating the EC_e in this study. Still, the EC_e estimation model based on the RI (FD-SG_{1483}, FD-SG_{1918}) was simple and effective because it only required information from two derivative spectral bands. Thus, the PLSR and MARS models provide higher prediction accuracy if the full spectrum is available, but the RI (FD-SG_{1483}, FD-SG_{1918}) model offers a potentially fast and reliable spectral index that accurately estimates the soil salinity. In addition, the spectral pre-processing techniques such as FD-SG and CR successfully removed the spectral noise and were applicable for modelling in this study. Considerably more work is required to develop these models to evaluate soils from different arid and semi-arid regions.

Conclusions

The results suggest that spectral indices, PLSR and MARS provide opportunities for estimating soil salinity. The analysis of the relationship between the EC_e and corresponding reflectance spectra in the soil samples provided the basis for a spectral index to estimate the soil salinity and construct EC_e estimation models using the PLSR and MARS techniques. The RI (FD-SG_{1483}, FD-

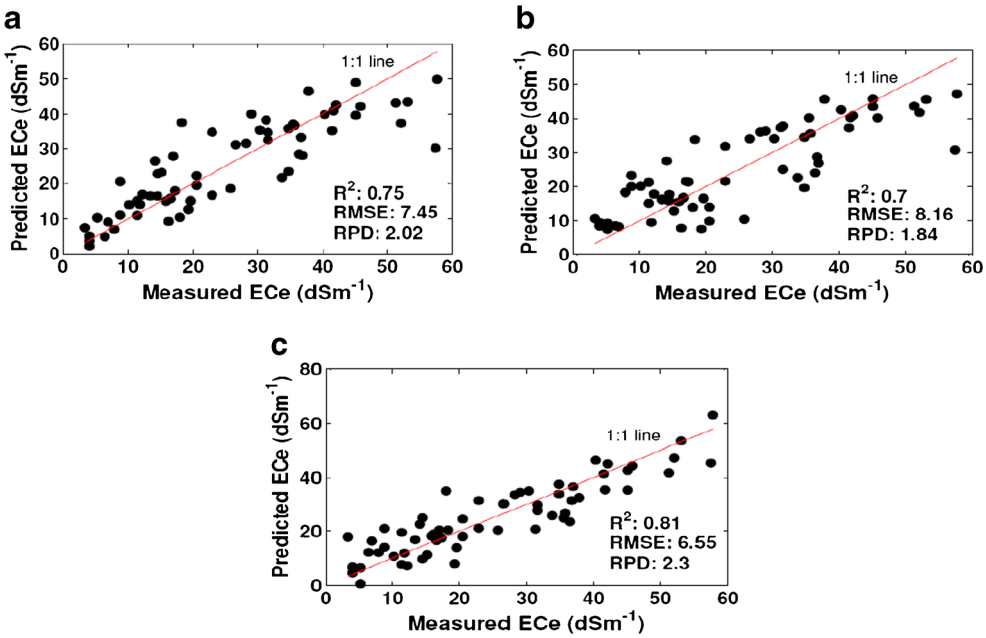

Fig. 7 MARS models for R (**a**), FD-SG (**b**) and CR (**c**) spectra

SG_{1918}), which produced R^2, RMSE and RPD values of 0.80, 6.20 and 2.06, respectively, is recommended for reliable estimations of the soil EC_e. The MARS calibration model estimated the EC_e better than the frequently used PLSR model, yielding optimal cross-validation R^2, RMSE and RPD values of 0.81, 6.55 and 2.30, respectively. In summary, the visible and near-infrared spectra offer the potential to efficiently estimate soil salinity, but factors such as the soil texture and mineral composition may complicate these estimations.

These estimation models should be subjected to further examination and optimisation before broad application in soil salinity modelling and mapping. The soil salinity predictive models used in the current study may become more accurate through the selection of optimal waveband regions for the model calibration rather than the full spectral range (400–2,500 nm). Our models were constructed based on soil spectra measured in the laboratory. Future research should focus on possible integration between spectra obtained in the field and laboratory and derived from satellite imagery. This integration could be more useful for accurate soil salinity mapping that employs PLSR and MARS models. Moreover, sufficient atmospheric correction and elimination of spectral noise and vegetation influences might enable these models to be used with spectral data collected from airborne or satellite platforms for more efficient modelling and mapping of soil salinity.

Table 5 Cross-validation results of EC_e estimation models of different spectral pre-processing methods

Estimation model	Methods	R^2	RMSE	RPD
DI (R_{673}, R_{1963})	R	0.59	8.06	1.57
NDI (R_{523}, R_{1963})	R	0.65	7.46	1.71
RI (R_{478}, R_{1978})	R	0.65	7.36	1.73
DI (FD-SG_{1378}, FD-SG_{1933})	FD+SG	0.74	6.96	1.87
NDI (FD-SG_{1498}, FD-SG_{1918})	FD+SG	0.77	6.28	2.03
RI (FD-SG_{1483}, FD-SG_{1918})	FD+SG	0.80	6.20	2.06
DI (CR_{572}, CR_{2222})	CR	0.54	8.60	1.48
NDI (CR_{572}, CR_{2222})	CR	0.54	8.60	1.48
RI (CR_{572}, CR_{2222})	CR	0.54	8.59	1.49
PLSR (350–2,500)	R	0.69	8.30	1.80
PLSR (350–2,500)	FD+SG	0.75	7.40	2.01
PLSR (350–2,500)	CR	0.77	7.20	2.07
PLSR+AP (350–2,500)	R	0.75	7.40	2.02
PLSR+AP (350–2,500)	FD+SG	0.77	7.20	2.08
PLSR+AP (350–2,500)	CR	0.81	6.50	2.29
MARS (350–2,500)	R	0.75	7.45	2.02
MARS (350–2,500)	FD+SG	0.70	8.16	1.84
MARS (350–2,500)	CR	0.81	6.55	2.30

R raw reflectance spectra, *FD+SG* first-derivative spectra with Savitzky–Golay smoothing, *CR* continuum removed spectra, *AP* clay content as the auxiliary predictor

Acknowledgements This study was in part funded by the Egyptian Ministry of Higher Education and by the Institute of Geography and Spatial Management of the Jagiellonian University, Poland. The corresponding author would like to thank Dr. El-Sayed Omran and soil and water department staff at Suez Canal University, Ismailia, Egypt, for assistance in laboratory analyses. The authors also thank Prof. Jack

Kozak, Jagiellonian University, for giving valuable suggestions and revisions for this paper.

References

Akramkhanov A, Martius C, Park SJ, Hendrickx JMH (2011) Environmental factors of spatial distribution of soil salinity on flat irrigated terrain. Geoderma 163:55–62

Amezketa E (2006) An integrated methodology for assessing soil salinization, a pre-condition for land desertification. J Arid Environ 67: 594–606

Badreldin N, Frankl A, Goossens R (2013) Assessing the spatiotemporal dynamics of vegetation cover as an indicator of desertification in Egypt using multi-temporal MODIS satellite images. Arab J Geosci. doi:10.1007/s12517–013–1142–8

Balasubramanian UABR, Saravanavel J, Gunasekaran S (2012) Ore mineral discrimination using hyperspectral remote sensing—a field-based spectral analysis. Arab J Geosci 6(12):4709–4716

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 L Res Manag 25:19–37

Bilgili AV, 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. J Arid Environ 74:229–238

Brown DJ, Shepherd KD, Walsh MG, Dewayne Mays M, Reinsch TG (2006) Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132:273–290

Buddenbaum H, Steffens M (2012) The effects of spectral pretreatments on chemometric analyses of soil profiles using laboratory imaging spectroscopy. Appl Environ soil sci 2012:1–12

Campbell JB (2011) Introduction to remote sensing, 5th edn. Guilford Press, New York

Chang CI (2003) Hyperspectral imaging: techniques for spectral detection and classification. Kluwer Academic/Plenum Publishers, New York

Chang CW, Laird DA, Mausbach JMJ, Hurburgh CR (2001) Near-infrared reflectance spectroscopy-principal components regression analysis of soil properties. Soil Sci Soc Am J 65:480–490

Clark RN, Roush T (1984) Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. J Geophys Res 89: 6329–6340

Dehaan RL, Taylor GR (2002) Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced soil salinization. Remote Sens Environ 80:406–417

Divya Y, Sanjeevi S, Ilamparuthi KA (2013) Study on the hyperspectral signatures of sandy soils with varying texture and water content. Arab J Geosci. doi:10.1007/s12517–013–1015–1

Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. Monographs on statistics and applied probability, vol. 57. Boca Raton, FL, CRC Press, p. 436

Eldeiry AA, Garcia LA (2008) Detecting soil salinity in alfalfa fields using spatial modeling and remote sensing. Soil Sci Soc Am J 72: 201–211

Ertlen D, Schwartz D, Trautmann M, Webster R, Brunet D (2010) Discriminating between organic matter in soil from grass and forest by near-infrared spectroscopy. Eur J Soil Sci 61:207–216

Exelis Visual Information Solutions (2012) ENVI 5.0. Exelis Inc, McLean, Virginia

Farifteh J, Farshad A, George RJ (2006) Assessing salt affected soils using remote sensing, solute modelling, and geophysics. Geoderma 130:191–206

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 Sens Environ 110:59–78

Felicísimo ÁM, Cuartero A, Remondo J, Quirós E (2012) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslide 10:175–189

Fernández-Buces N, Siebe C, Cram S, Palacio JL (2006) Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico. J Arid Environ 65:644–667

Friedman JH (1991) Multivariate adaptive regressions splines. Ann Stat 19:1–67

Geladi P, Kowalski BR (1986) Partial least-squares regression: a tutorial. Anal Chim Acta 185:1–17

Ghabour ThK (1997) Soil salinity classification of the north Nile delta region based on remotely sensed data. Int. Symposium on Salt Affected Soils, 22–26 Sep., Cairo, Egypt

Goossens R, De Dapper M, El Badawi MA, Ghabour T (1994) A simulation model to monitor the soil salinity in irrigated arable land in Arid areas based upon remote sensing and GIS. Earsel Adv Remote Sens 2:3–11

Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning: datamining, inference, and prediction. Springer, New York

Ibrahim HM, El Falaky AA (2013) Soil salinity mapping in the Sinai Peninsula of Egypt using geographic information system and remote sensing techniques. In: Shahid SA, Abdelfattah MA, Taha FK (eds) Developments in soil salinity assessment and reclamation, innovative thinking and use of marginal soil and water resources in irrigated agriculture, pp. 113–125. Springer Science + Business Media, B.V., Berlin, p 808

Jackson ML (1973) Soil chemical analysis. Prentice-Hall of India private limited, New Delhi, 498 pp

Jekabsons G (2011) ARESLab: Adaptive regression splines toolbox for Matlab/Octave. available at http://www.cs.rtu.lv/jekabsons/

Khan NM, Rastoskuev VV, Sato Y, Shiozawa S (2005) Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. Agric Water Manag 77:96–109

Kilmer VJ, Alexander LT (1949) Methods of making mechanical analysis of soils. Soil Sci 68:15–24

Kokaly RF, Clark RN (1999) Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. Remote Sens Environ 67:267–287

Li HY, Shi Z, Webster R, Triantafilis J (2013) Mapping the three-dimensional variation of soil salinity in a rice-paddy soil. Geoderma 195–196:31–41

Luoto M, Hjort J (2005) Evaluation of current statistical approaches for predictive geomorphological mapping. Geomorphology 67:299–315

Mashimbye ZE, Cho MA, Nell JP, De Clercq WP, Van Niekerk A, Turner DP (2012) Model-based integrated methods for quantitative estimation of soil salinity from hyperspectral remote sensing data: a case study of selected South African soils. Pedosphere 22:640–649

Masoud AA, Koike K (2006) Arid land salinization detected by remotely-sensed landcover changes: a case study in the Siwa region, NW Egypt. J Arid Environ 66:151–167

MathWorks (2011) MATLAB 8.0. MathWorks, Inc., Natick, MA

Matinfar HR, Alavi Panah SK, Zand F, Khodaei K (2011) Detection of soil salinity changes and mapping land cover types based upon remotely sensed data. Arab J Geosci 6:913–919

Metternicht GI, Zinck JA (2003) Remote sensing of soil salinity: potentials and constraints. Remote Sens Environ 85:1–20

Metternicht GI, Zinck JA (2008) Remote sensing of soil salinization: Impact on land management. CRC Press, Boca Raton, FL

Mouazen AM, Maleki MR, De Baerdemaeker J, Ramon H (2007) On-line measurement of some selected soil properties using a VIS-NIR sensor. Soil Till Res 93:13–7

Mulder VL, de Bruin S, Schaepman ME, Mayr TR (2011) The use of remote sensing in soil and terrain mapping—a review. Geoderma 162(1–2):1–19

Nawar S, Reda M, Farag F, El-Nahry A (2011) Mapping soil salinity in El-Tina plain in Egypt using geostatistical approach. Geoinformatics Forum, Salzburg, pp 81–90

Othman AA, Al-Saady YI, Al-Khafaji AK, Gloaguen R (2013) Environmental change detection in the central part of Iraq using remote sensing data and GIS. Arab J Geosci. doi:10.1007/s12517–013–0870–0

Page AL, Miller RH, Keeney DR (1982) Chemical and microbiological properties, part 2. Society of agronomy, American

Patzold S, Mertens FM, Bornemann L, Koleczek B, Franke J, Feilhauer H, Welp G (2008) Soil heterogeneity at the field scale: a challenge for precision crop protection. Precis Agric 9:367–390

Qadir M, Oster JD, Schuber S, Noble AD, Sahrawat KL (2007) Phytoremediation of sodic and saline-sodic soils. Adv Agron 96: 197–247

Richards LA (1954) Diagnosis and improvement of saline and alkaline Soils. U.S. Department of Agriculture, Hand Book, No, 60

Rinnan A, Berg FVD, Engelsen SB (2009) Review of the most common pre-processing techniques for near-infrared spectra. Trends Anal Chem 28:1201–1222

Samui P (2012) Multivariate adaptive regression spline (MARS) for prediction of elastic modulus of jointed rock mass. Geotech Geol Eng 31:249–253

Savitzky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least-squares procedures. Anal Chem 36: 1627–1639

Sidike A, Zhao S, Wen Y (2014) Estimating soil salinity in Pingluo County of China using QuickBird data and soil reflectance spectra. Int J Appl Earth Obs Geoinf 26:156–175

Shepherd KD, Walsh MG (2002) Development of reflectance spectral libraries for characterization of soil properties. Soil Sci Soc Am J 66: 988–998

Sonmez S, Buyuktas D, Okturen F, Citak S (2008) Assessment of different soil to water ratios (1:1, 1:2.5, 1:5) in soil salinity studies. Geoderma 144:361–369

Stenberg B, Viscarra Rossel RA, Mouazen AM, Wetterlind J (2010) Visible and near infrared spectroscopy in soil science. Adv Agron 107:163–215

Stevens A, Udelhoven T, Denis A (2010) Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. Geoderma 158:32–45

Stockle CO (2013) Environmental impact of irrigation: a review.<http://www.swwrc.wsu.edu/newsletter/fall2001/IrrImpact2.pdf> Accessed 10 October.2013

Tamas J, Lenart C (2006) Analysis of a small agricultural watershed using remote sensing techniques. Int remote sens 27:3727–3738

Tiwari PS, Garg RD, Sen AK, Dadhwal VK (2013) Spectral delineation of albite zone using ASTER data in Khetri Copper Belt. Arab J Geosci. doi:10.1007/s12517–013–1087-y

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

Vidoli F (2011) Evaluating the water sector in Italy through a two stage method using the conditional robust nonparametric frontier and multivariate adaptive regression splines. Eur J Oper Res 212:583–595

Waiser TH, Morgan CLS, Brown DJ, Hallmark TC (2007) In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. Soil Sci Soc Am J 71:389–396

Weng YL, Gong P, Zhu ZL (2008) Reflectance spectroscopy for the assessment of soil salt content in soils of the Yellow River Delta of China. Int Remote Sens 29:5511–5531

Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemometr Intell Lab 58:109–130

Yang CC, Prasher SO, Lacroix R, Kim SH (2003) A multivariate adaptive regression splines model for simulation of pesticide transport in soils. Biosyst Eng 86:9–15