Modeling and Mapping of Soil Salinity with Reflectance Spectroscopy and Landsat Data Using Two Quantitative Methods (PLSR and MARS)

Said Nawar 1,2,*, Henning Buddenbaum 3, Joachim Hill 3, and Jacek Kozak 1

1 Institute of Geography and Spatial Management, Jagiellonian University, Krakow 30-387, Poland; E-Mail: jkozak@gis.geo.uj.edu.pl
2 Faculty of Agriculture, Suez Canal University, Ismailia 41522, Egypt
3 Environmental Remote Sensing and Geoinformatics, Trier University, 54286 Trier, Germany; E-Mails: buddenbaum@uni-trier.de (H.B.); hillj@uni-trier.de (J.H.)

* Author to whom correspondence should be addressed; E-Mail: snawar@gis.geo.uj.edu.pl; Tel.: +48-797-941-230; Fax: +48-126-645-385.

Abstract: The monitoring of soil salinity levels is necessary for the prevention and mitigation of land degradation in arid environments. To assess the potential of remote sensing in estimating and mapping soil salinity in the El-Tina Plain, Sinai, Egypt, two predictive models were constructed based on the measured soil electrical conductivity (ECe) and laboratory soil reflectance spectra resampled to Landsat sensor’s resolution. The models used were partial least squares regression (PLSR) and multivariate adaptive regression splines (MARS). The results indicated that a good prediction of the soil salinity can be made based on the MARS model (R^2 = 0.73, RMSE = 6.53, and ratio of performance to deviation (RPD) = 1.96), which performed better than the PLSR model (R^2 = 0.70, RMSE = 6.95, and RPD = 1.82). The models were subsequently applied on a pixel-by-pixel basis to the reflectance values derived from two Landsat images (2006 and 2012) to generate quantitative maps of the soil salinity. The resulting maps were validated successfully for 37 and 26 sampling points for 2006 and 2012, respectively, with R^2 = 0.72 and 0.74 for 2006 and 2012, respectively, for the MARS model, and R^2 = 0.71 and 0.73 for 2006 and 2012, respectively, for the PLSR model. The results indicated that MARS is a more suitable technique than PLSR for the estimation and mapping of soil salinity, especially in areas with high levels of
1. Introduction

Salinization is a worldwide problem that affects the physical and chemical properties of soil, leading to the loss of crop productivity [1,2]. 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 [3]. As salt-affected soils are common in arid and semiarid climates with precipitation rates lower than evapotranspiration rates [4], salinization is a prevalent environmental hazard, posing serious socio-economic and environmental implications. To manage salt-affected soils better, soil salinity must be first monitored and mapped [5]. Typically, soil salinity is assessed by measuring the soil electrical conductivity (ECe) in saturated paste extracts or using different soil to water ratios [6]; however, the complexity and cost of traditional methods for regional monitoring of soil salinity has motivated the exploration of various remote sensing techniques, which have exhibited promising results in mapping the variations in surface soil salinity [2].

For the last two decades, remotely sensed imagery has demonstrated its ability to monitor salinity changes effectively from surface characteristics in real-time and at various scales [7–10]. Several studies have shown the visible, near infrared, or short-wave infrared spectral bands from the optical sensors to be promising for the detection of surface soil salinity [11–15]. In addition, hyperspectral data have been successfully used in several studies on soil salinity, enabling quantitative assessment of salt-affected soils [16–22]. However, practical limitations associated with hyperspectral imagery, including the availability of orbital data and the limited spatial coverage of the existing satellite sensors, still limit its potential for regional monitoring of salt-affected soils [23]. The recent launch of Landsat 8 has extended opportunities to map soil salinity, with Landsat data now having been freely available for several years. Therefore, this study found it appropriate, economical, and sustainable to use Landsat images for monitoring salt-affected soils.

Several regression methods based on image reflectance have been used to estimate the soil salinity, with partial least-squares regression (PLSR) being the most common one [18,24]. The PLSR approach has inference capabilities, which are useful to model a probable linear relationship between the reflectance spectra and the salt content in soils [18], and a capacity to include several response variables simultaneously while effectively addressing strongly collinear and noisy predictor variables [25]. Multivariate adaptive regression splines (MARS), typically known as a nonparametric method that estimates complex nonlinear relationships among independent and dependent variables [26], has been effectively applied in different fields [27–30] and generally exhibits high-performance results compared with other linear and non-parametric regression models, such as principal component regression (PCR) and artificial neural networks (ANN) [31]. Bilgili et al. [28,32] used MARS to model soil salinity and reported that MARS provided better estimations for the soil ECe than the generally used PLSR method, yielding the best cross-validation $R^2$ and RPD values for air-dried soils. These results were confirmed.
by Nawar et al. [33], who modelled soil salinity using MARS and obtained similarly high $R^2$ and RPD values (0.81 and 2.3, respectively).

A robust method for estimating the soil salinity on a regional scale must provide sufficiently accurate measurements and be cost-effective without excessive complexities, given the high temporal and spatial variability of soil salinization [23,34]. Although several of the studies estimated soil salinity using measured reflectance spectra and multispectral remote sensing data, effective and reliable methods for the accurate estimation of low salinity using linear and non-linear predictive models have not been implemented to date [35–37]. To address this gap, the aim of this study was to develop a robust methodology to estimate and map the soil salinity in arid areas using freely available Landsat 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) data. This aim was addressed through the following activities: (i) building linear (PLSR) and non-linear (MARS) predictive models based on laboratory soil reflectance spectra resampled to Landsat TM and ETM+ spectral resolutions and the measured EC$_e$; (ii) applying the resulting models to TM and ETM+ data to generate soil salinity maps; (iii) estimating the accuracy of these models in soil salinity mapping; and (iv) evaluating their usefulness in the monitoring of salt-affected soils. The paper builds on a previous study [33], in which soil salinity was successfully related to the soil spectra measured in laboratory conditions via the PLSR and MARS models.

2. Materials and Methods

2.1. 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 (Figure 1). This area is characterized by arid conditions and the annual rainfall varies between 33.4 mm and 70.3 mm. The mean air temperatures range from 7.5 °C to 23.3 °C in winter and from 16.3 °C to 35.6 °C in summer, hence mean evaporation is high, from 3.6 mm/day to 7.3 mm/day. The soil salinity varies from slightly saline to highly saline and the dominant salt minerals are NaCl and MgCl$_2$ [33]. The soil texture varies from loamy sand to clay. Nawar et al. [38] classified the soils of the El-Tina Plain into two orders: Entisols and Aridisols, which include eight subgroups: Typic Aquisalids, Typic Haplosalids, Aquic Torriorthents, Typic Torriorthents, Aquic Torripsamments, Typic Torripsamments, Gypsic Aquisalids, and Gypsic Haplosalids.

Figure 1 shows locations of soil sampling sites and physiographic features of the study area. This region is characterized by the presence of sabkhas (salt flats), which are characterized by a crusty surface consisting of evaporate deposits (including salt, gypsum, and calcium carbonate). Wet sabkhas are located mostly in the northwestern part of the study area and are marked with a dark brown color. Dry sabkhas are located mostly in the northeast and are shown in light gray. Fish ponds and other water bodies occupy up to half of the total area and appear in dark blue and black in the satellite image. Vegetated areas (marked green) are located in the south and southeast in small and scattered areas cultivated mainly with seasonal crops such as maize (Zea mays), sesame (Sesamum indicum), and peanuts (Arachis hypogaea). Sandy soils in the southwestern part of the study area are shown with a pale yellow color.
The El-Tina Plain surface is nearly flat, with a gentle incline from south to north. Elevations in the area barely reach 5 m above sea level, and areas below sea level are also present [39]. The water table is persistently high in the northern part of the plain, ranging from 50 to 75 cm below the ground surface due to the subsurface seepage of Mediterranean seawater [38]. Because of small elevation gradients in the region, high salinity of the ground water table (the mean salinity is 79 dSm$^{-1}$), and unreasonable irrigation, salinization has become a major problem in the El-Tina Plain, resulting in reduced crop yields [40]. In recent years, with an increase of the pumping irrigation using water with a slight to moderate salinity (0.9–1.4 dSm$^{-1}$), increased rates of leakage and groundwater recharge are leading to an acceleration of the salinization rate in the southern part of the study area. Due to the increasing salinization hazard, an accurate estimation of the soil salinity level is urgently required for the prevention of further degradation of soils in the study area. As our observations and analysis of 81 soil profiles (3–4 layers per soil profile) across the study area proved that soil salinity is homogeneous in the soil profile through all its layers, estimating the salinity level in the soil surface layer should be sufficient to characterize the entire profile. This justifies the use of remote sensing to monitor changes in soil salinity of the study area.

2.2. Soil Samples

Seventy topsoil samples (0–20) were collected in the study area between 28 August and 4 September 2012 (Figure 1). The soil salinity ($EC_e$) was measured in a soil paste extract according to the method of
Jackson [41]. Next, spectral measurements of soil samples were performed at the laboratory of Trier University in Germany. In addition, field spectroscopy was used at some locations to measure surface reflectivity. Due to the limited number of measurements during the field campaign, field spectroscopy was used only for calibration and validation of the resulting image spectra. For model calibration, only the laboratory-measured reflectance spectra were used.

The ECe ranged from 3.3 to 166 dSm$^{-1}$; the mean salinity value was 33.03 dSm$^{-1}$ (Table 1). The relatively high mean soil salinity indicated that the soils in the study area were highly salt-affected. Very slightly and slightly saline soil samples accounted for only 15.7% of all samples. Moderately saline, strongly saline, and very strongly saline soil samples accounted for 30% and 54.3%, respectively, of all samples. The soil pH values ranged between 7.1 and 8.5; the mean pH value was 7.86, suggesting that the soils in the investigated area are slightly alkaline. The clay content of the soil samples ranged from undetectable to 54%, and soil samples containing >40% of clay accounted for 30% of all the soil samples. The soil organic matter (OM) content was low, with mean and maximum values of 0.83% and 2.3%, respectively. Samples with OM content > 1.5% accounted for 10% of all samples, and the overall OM content was low (lower than 1.5%).

Table 1. Descriptive statistics of the soil parameters.

<table>
<thead>
<tr>
<th>pH</th>
<th>EC$_e$ dSm$^{-1}$</th>
<th>CaCO$_3$ %</th>
<th>OM %</th>
<th>Clay %</th>
<th>Silt %</th>
<th>Sand %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>7.10</td>
<td>3.30</td>
<td>0.01</td>
<td>0.00</td>
<td>0.50</td>
<td>16.00</td>
</tr>
<tr>
<td>Max</td>
<td>8.50</td>
<td>166.80</td>
<td>21.90</td>
<td>2.30</td>
<td>54.30</td>
<td>34.60</td>
</tr>
<tr>
<td>Mean</td>
<td>7.86</td>
<td>33.03</td>
<td>2.97</td>
<td>0.83</td>
<td>27.22</td>
<td>20.81</td>
</tr>
<tr>
<td>St.dev</td>
<td>0.29</td>
<td>31.33</td>
<td>3.07</td>
<td>0.52</td>
<td>16.77</td>
<td>10.21</td>
</tr>
<tr>
<td>CV(100) *</td>
<td>3.7</td>
<td>94.85</td>
<td>103.26</td>
<td>63.03</td>
<td>61.62</td>
<td>49.04</td>
</tr>
</tbody>
</table>

* CV = St.dev $\times$ 100/mean.

2.3. Soil Spectrometry and Spectral Characteristics of the Soils

Soil reflectance spectra were collected using a portable spectroradiometer (FieldSpec-FR, ASD), which measures the reflectance over the range of 350–2500 nm, with a resolution of 3 nm at 700 nm and 10 nm at 1400 and 2100 nm, and a 1.4 to 2 nm sampling interval in the short-wave infrared domain. The measurements were conducted in a dark laboratory environment. Plastic dishes contained the soil samples, which were leveled off to a 2.0-cm thickness [42]. Each sample was measured three times, and the final spectral measurement was attained by averaging the curves [33]. To calculate the absolute reflectance of the samples, the radiance from each sample was divided by the radiance from the white reference panel and then multiplied by the reflectance of the reference panel.

The reflectance curves display a shoulder near 800 nm, show two deep absorption features at 1415 nm and 1915 nm, and have several weak absorption features near 494 nm, 673 nm, 1748 nm, 2207 nm and 2385 nm. The soil albedo and the absorption feature depth vary with the level of soil salinity (Figure 2a). The deeper water absorption features at 1415–1444 nm and 1915–1932 nm likely represent the combined results of O–H stretches, H–O–H bending combination of molecular free water and overtones, and, as the salt content increases, the features becoming more extended and asymmetrical (Figure 2b).
Figure 2. Raw reflectance spectra (R) (a); continuum-removed spectra (CR) (b) of five soil salinity classes.

2.4. Soil Salinity Modeling

In the previous study [33], two methods were used and tested for modeling the soil salinity and for measuring the soil spectra in laboratory conditions: PLSR and MARS. PLSR is a popular regression method that is frequently applied in chemometrics [25,42–47]. The algorithm uses a linear multivariate model to relate two data matrices: the predictor variable X and the response variable Y. This algorithm selects successive orthogonal (latent) factors, thereby maximizing the X and Y covariance—the covariance between the spectra (X) and a measured soil property (Y). To prevent over- or under-fitting of the data, the optimum number of latent factors was identified by performing leave-one-out cross-validation (LOOCV); the minimum root mean squared error (RMSE) with the highest coefficient of determination (R²) was used as the decision criterion.

Developed by Friedman [26], MARS is a non-parametric regression technique used for fitting the relationship between dependent and independent variables via application of the splines theory. Recently, MARS was applied as a regression method in several disciplines [28,29,32,48,49] and was consistently noted to perform better than traditional statistical methods. The MARS analysis uses basis functions to model the predictor and response variables [50]. MARS creates basis functions that can serve as new predictor variables in modeling; it does so by splitting the data into splines (or sub-regions), which have varied interval ending knots at the points where the regression coefficients differ.
By utilizing adaptive piecewise linear regressions, this process also suits 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 that are of 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 for a smoothing procedure, which gives the final MARS approximation a particular level of continuity. To determine which basis functions should be included in the model, MARS uses a modified form of generalized cross-validation (GCV) [26].

The strength of the soil salinity prediction models can be evaluated following the suggestions from Chang et al. [51]. They grouped Ratio of Performance to Deviation (RPD) values into three classes: category A (RPD > 2) are models that accurately predict a given property; category B (1.4 < RPD < 2) are models that have limited prediction ability; and category C (RPD < 1.4) are models with no prediction ability. The PLSR and MARS analyses were performed using MATLAB 8.0 software [52].

2.5. Estimation of Soil Salinity Based on Landsat Data

The results received in [33] 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. The suggested hypothesis is that if the predictive model based on the measured laboratory reflectance spectra proved accurate, the Landsat data could be used as an input in the soil salinity estimating model [35]. For this purpose, the PLSR and MARS models were further developed for soil salinity modeling and mapping using soil spectra resampled to Landsat spectral resolution. The measured reflectance spectra were resampled to Landsat 7 ETM+ and Landsat 5 TM spectral sensor resolutions using ENVI’s spectral resampling routine [53]. To perform the resampling process, the pre-defined filter function method for ETM+ and TM was selected. The models were applied to the image scale of selected pre-processed Landsat images, pixel by pixel, to produce the soil salinity maps.

2.5.1. Landsat Data Pre-Processing

A level 1 Terrain Corrected (L1T) Landsat 7 Enhanced Thematic Mapper Plus (ETM+) satellite image (acquired on 24 August 2012) and an L1T Landsat 5 Thematic Mapper (TM) satellite image (acquired on 22 May 2006), which are freely available from the Landsat archive at the United States Geological Survey (USGS) (http://earthexplorer.usgs.gov/), were used in this study. The image acquisition dates were selected to be as close to the soil sampling campaigns as possible. The dates were representative of the dry season in Egypt (summer season) with scarce vegetation and dry soils. Because the soil is bare and dry, information about the surface’s constituents can be collected by the satellite sensors [54]. Six spectral bands with a spatial resolution of 30 m and 8-bit radiometric resolution were used, three in the visible (0.45–0.52 µm, 0.52–0.60 µm, and 0.63–0.69 µm), one in the near-infrared (0.76–0.90 µm), and two in the mid-infrared (1.55–1.75 µm and 2.08–2.35 µm) parts of the spectrum. The gaps in the ETM+ SLC-off image of 24 August 2012 were filled using the localized linear histogram match (LLHM) method [55]. Landsat 7 ETM+ SLC-off, 8 August 2012 was used to fill the gaps due to the good quality of the image, the short lag time (20 days) between the two images, and the fact that the gaps were not overlapping. ENVI 5.0 software [53] was used to perform the gap filling.
Accurate image registration is an important aspect in studies examining the temporal change in the environment, and for temporal comparison, images should be co-registered with sub-pixel accuracy [56]. As the accuracy of geometric correction for the L1T product of Landsat depends on the quality of the control points and the resolution of the DEM used, the two images were additionally co-registered using 20 GCPs selected for each scene that were well-distributed over the images, with resultant RMS error (RMSE) of less than one pixel (0.21). A nearest neighbor resampling procedure was used in the rectification so that the digital numbers (DNs) representing the brightness values of the pixels for each band were not altered.

The atmospheric correction process is very important to remove the effects of atmospheric distortion on the soil chromophores; hence their response was registered by a sensor [57]. The use of a proper atmospheric correction model is therefore a mandatory step to ensure successful imaging spectroscopy during the approximation of laboratory-derived spectra with image spectra. To reduce atmospheric effects, the Landsat images were atmospherically corrected to obtain surface reflectance using the ATmospheric CORection and haze Reducion (ATCOR 2) model, which is based on the MODerate resolution atmospheric TRANsmission (MODTRAN 5) code [58]. The following sensor and physical environmental parameters were used to run the atmospheric correction model: Earth-Sun Distance = 1.00934 Astronomical Unit (AU), Average Surface Altitude = 0.002 km, Atmospheric model = Dry (26 °C), and Aerosol type = Rural.

To compensate for the spectral interference that may arise from the surroundings, the analysis was limited to bare soils. Features other than bare soils (water, vegetation) were hidden using supervised classification with the Spectral Angle Mapper (SAM) method available in ENVI and trained by classes carefully selected from the two scenes. The classification result was verified through comparison with information collected during the field work. The blacked-out area comprised 7.52% of the study area in 2006, and 11.66% in 2012. The increase was related mostly to the expansion of water bodies due to the rise of the water table between 2006 and 2012. Finally, spatial subsets were created and various stacks of color composites were produced to aid in the interpretation of the results.

2.5.2. Application of the PLSR and MARS Models to Landsat Data

PLSR and MARS models were constructed based on the soil ECe and laboratory soil reflectance spectra that were resampled to Landsat TM and ETM+ spectral resolutions. The resulting models were assessed using $R^2$, RMSE, and RPD parameters. After the Landsat images were pre-processed and optimized, the resulting models were used to estimate the image-scale soil salinity (ECe) and to map the spatial variability of soil salinity. The resulting soil salinity maps for 2006 and 2012 were validated using 37 soil samples collected between 28 May and 8 June 2006, and 26 soil samples collected in September 2012 (Figure 1). According to geographical coordinates of soil samples measured in the field, the corresponding ECe was extracted from Landsat-based soil salinity maps. The predicted values were then compared with the true (laboratory) values using $R^2$ and RMSE parameters.
3. Results

3.1. Evaluation of ETM+ Data

The trends in the spectral reflectance derived from ETM+ bands and the resampled measured reflectance spectra were consistent with variations in the soil salinity values (Figure 3). The spectral reflectance values from the ETM+ data’s six bands for all of the soil samples were lower than the corresponding field and laboratory-measured spectral reflectance. This result may be attributed to the differences between the Landsat sensor and the sensor that was used for measuring soil spectra. In addition, the accuracy of atmospheric correction is another factor impacting the spectral reflectance of the ETM+ data [35]. From the data in Figure 3, it is apparent that there was a high correlation between the measured spectral reflectance and the spectral reflectance derived from ETM+ data.

**Figure 3.** Comparison of the field-measured, laboratory, and Landsat 7 ETM+ spectral reflectance of soil samples with different salinity values (ECe).  

In Figure 4, the results of the analysis of the regression coefficients for a PLSR model with six latent variables are shown. The NIR (i.e., Band 4) and SWIR-1 (i.e., Band 5) bands had the highest contribution to the estimation of soil salinity, followed by the red (i.e., Band 3), SWIR-2 (i.e., Band 7), green band
Remote Sens. 2014, 6

(i.e., Band 2) and the blue band (i.e., Band 1), which made the lowest contribution to the estimation of soil salinity. There was a high correlation between the reflectance values for resampled spectra and the Landsat spectra ($R^2 > 0.93$). These results indicated that the measured spectral reflectance and the reflectance values derived from Landsat have great potential for predicting and mapping soil salinity.

Figure 4. PLS regression coefficients of the six bands of Landsat data.

3.2. Soil Salinity Estimation Using the PLSR and MARS Models

The application of the PLSR and MARS models based on resampled measured spectra resulted in the calibration models in Figure 5. For the PLSR model, the optimal number of latent factors used to estimate ECe was defined by analyzing the $R^2$, RMSE, and RPD values. It is apparent from the results that the PLSR model exhibited good performance, with the values of $R^2$ of 0.70, RMSE of 6.95, and RPD of 1.82 (Figure 5a). The MARS model resulted in good performance, with $R^2$, RMSE, and RPD values of 0.73, 6.53, and 1.96, respectively (Figure 5b). This result indicates that the MARS model is suitable for the estimating and mapping of soil salinity.

Using the Landsat images, the PLSR and MARS calibration models were used to produce soil salinity maps for 2006 and 2012 (Figure 6). The maximum estimated ECe values were the same and equal to 56 dSm$^{-1}$, the minimum values were $-1.6$ dSm$^{-1}$ and $-1.8$ dSm$^{-1}$, the mean values were 15.4 dSm$^{-1}$ and 15.8 dSm$^{-1}$, and the standard deviations were 12.3 dSm$^{-1}$ and 13.6 dSm$^{-1}$, for the PLSR and MARS resulting maps for 2006, respectively (Figure 6a,b). For 2012 the maximum estimated ECe values were 51 dSm$^{-1}$ and 50 dSm$^{-1}$, the minimum values were the same and equal to $-1.2$ dSm$^{-1}$, the mean values were 13.6 dSm$^{-1}$ and 14.6 dSm$^{-1}$, and the standard deviation were 11.6 dSm$^{-1}$ and 12.4 dSm$^{-1}$, for the PLSR and MARS maps, respectively (Figure 6c,d). The negative values are not reasonable for both models and all maps, comprised less than 3% of all pixels. These negative values may be explained by a number of different factors. One factor is the uncertainty of the atmospheric correction. Another factor is that the PLSR model was established based on laboratory-measured spectral data that are different from the pixel reflectance in the Landsat images, which are influenced by spectral mixture [35].
The received soil salinity values were grouped into five classes (<4, 4–8, 8–16, 16–32, and >32 dS/m) (Figure 7). The soil salinity distribution in Figure 7 is consistent with our prior knowledge of the area.
Remote Sens. 2014, 6

and field investigation. The spatial distribution of soil salinity had the same pattern for all four maps (PLSR and MARS models for 2006 and 2012). The areas with high salinity were located mostly in the northern and western parts of the study area, and the areas with low salinity occupied the southern and eastern parts of the study area; the soil salinity tended to increase from south to north in the direction of the El-Salam Canal and the Mediterranean Sea, which validated the trend of the effect caused by the seepage from the canals and the Mediterranean sea. The soil salinity in 2012 was higher than in 2006, mostly in the southern part of the study area, primarily due to the expansion of classes 4–8 dSm$^{-1}$, 8–16 dSm$^{-1}$, and 16–32 dSm$^{-1}$, and an accompanying decrease of the class with the lowest salinity (<4 dSm$^{-1}$). At the same time, there was a decrease of the class with the highest salinity (>32 dSm$^{-1}$) in the northern part of the study area in 2012, due to the cultivation and leaching in these areas.

**Figure 7.** (a) PLSR results for 2006; (b) MARS results for 2006; (c) PLSR results for 2012; and (d) MARS results for 2012.
The MARS model showed higher salinity values than the PLSR model in both 2006 and 2012. The proportion of classes 4–8, 8–16, 16–32 and >32 dSm\(^{-1}\) was 53.54% for PLSR and 62.54% for MARS in 2006 (Figure 7a,b), consistent with the results for 2012 (63.35% for PLSR and 76.95% for MARS; Figure 7c,d). Consequently, the very slightly saline class (<4 dSm\(^{-1}\)) covered a larger area in the PLSR and MARS results and decreased in area over the period 2006–2012.

The quantitative soil salinity maps produced by the PLSR and MARS models were successfully validated with 37 and 26 soil samples for the 2006 and 2012 maps, respectively. The predicted and measured values are plotted in Figure 8. The results indicated a satisfactory relationship between the measured data and predicted EC\(_e\) values (R\(^2\) = 0.71 and 0.72 and RMSE = 5.37 and 5.30), for PLSR and MARS maps in 2006, respectively (Figure 8a,b). The validation results for MARS (R\(^2\) = 0.74 and RMSE = 5.23) and PLSR (R\(^2\) = 0.73 and RMSE = 5.18) maps for 2012 were of similar quality (Figure 8c,d). The results indicate that the MARS model using Landsat reflectance data is stable, and slightly better than the PLSR model for mapping and monitoring the soil salinity. It is apparent from Figure 8 that when the EC\(_e\) values are high, some of the predicted values are deviated further from the measured EC\(_e\) values, implying a nonlinear relationship between the EC\(_e\) and the spectral data. The MARS model showed an ability to model the nonlinear relationship better than the PLSR model (Figure 8b,d).

**Figure 8.** Scatterplots of measured versus extracted EC\(_e\) values for (a) PLSR results for 2006; (b) MARS results for 2006; (c) PLSR results for 2012; and (d) MARS results for 2012.
4. Discussion

4.1. Soil Salinity Estimation Using the PLSR and MARS Models

Two predictive models, PLSR and MARS, based on measured reflectance spectra and Landsat data, were applied to estimate and map soil salinity in the El-Tina Plain, Egypt. The performances, as well as their usefulness in monitoring the soil salinity of these models, were compared. PLSR was successfully used for predicting and mapping soil salinity elsewhere in the world, for example, in New South Wales, Australia [19], the Yellow River delta region in China [24], and Pingluo County, China [35]. MARS was successfully used for estimating soil salinity and other soil properties in laboratory work [28,32,45]. In this study, the results revealed that there was potential for the MARS and PLSR predictive models to improve soil salinity mapping using remote sensing data. The obtained results indicate that using MARS as a nonlinear modeling approach has improved predictive quality compared to the PLSR method. These results support the findings of Bilgili et al. [32] and Nawar et al. [33], which demonstrated that the MARS predictive models provided more robust predictions of soil salinity than the PLSR models. This difference is probably because MARS is a non-linear and flexible modeling method, capable of fitting complex and non-linear relationships and specifying the interaction effects as well as the linear combinations of variables [26,28,59]. Several studies have demonstrated that the prediction of high soil salinity levels implies some non-linear relationship between the measured soil salinity and the soil reflectance spectra [18,35,60]. MARS enables the reduction of the effects of the multistep process and any other unknown nonlinearity, and therefore is more effective than the PLSR model [59]. Although the PLSR method assumes a linear relationship between the ECe and soil reflectance spectra, a small deviation from linearity is acceptable and can be readily suppressed by including additional modeling factors [61].

The efficiency of the selected MARS and PLSR models to predict the soil salinity is demonstrated by all the indicators used in model testing (R^2 = 0.73, RMSE = 6.53 dSm\(^{-1}\), and RPD = 1.96, for MARS; and R^2 = 0.70, RMSE = 6.95 dSm\(^{-1}\), and RPD = 1.82 for PLSR). This efficiency is in part due to the high correlation between the measured soil salinity and the soil spectra. The PLSR model in this study exhibited superiority in the prediction power (R^2 = 0.70) of soil salinity over those reported by Shamsi et al. [62] (R^2 = 0.39) and recently by Allbed et al. [63] (R^2 = 0.65), which were developed using a simple regression model and multispectral satellite images. Nonetheless, differences between the simple regression and the predictive models can have a high impact on the accuracy of estimating and mapping the soil salinity. Our findings further support the idea that the prediction of soil salinity based on the MARS and PLSR models and measured spectra yield better results than those based on simple regression and multispectral image spectra alone [64].

In this research, the method of resampling measured spectral reflectance was an important step in the processing. There are differences between the spectral reflectance of Landsat data and the simulated spectral bands. The method that was used for resampling the measured reflectance spectra only considered the spectral resolution of the Landsat sensor. Moreover, other factors such as radiometric resolution, spatial resolution, sensor geometry, and signal-to-noise ratio were not considered and should be involved in the simulation process in future research. Sidike et al. [35] used the resampled spectral reflectance with the PLSR model for estimating the soil salinity using QuickBird data; only the spectral
resolution of the QuickBird sensor was considered. Other advanced simulation methods, such as Liu et al. [65], which include effective parameters such as sensor geometry and signal-to-noise ratio, should be tested in further research. The field spectroscopy measurements were used only to assess the resulting image spectra after the atmospheric correction pre-processing due to the limitation of the number of field spectrometry measurements. Only the laboratory-measured reflectance spectra were used for statistical modeling in this study.

The good results of the predictive models can be attributed to: (1) the good correlation between the soil salinity and the clay content [28,33], and (2) a strong decrease in the soil albedo with increasing soil salinity [18,66]. The influence of several factors, such as vegetation, soil moisture, and organic matter (OM), which could decrease the predictive capabilities of the models, was reduced. The vegetated areas were blacked out in the image classification process. Because this study was based on dry soils, the influence of moisture content on the soil spectra was also likely negligible. As long as the organic matter (OM) content is lower than 2%, the soil spectra are hardly affected by OM [20,67]. Therefore, in this study, the OM could not have affected the spectral reflectance because 97.15% of the soil samples in this study had an OM content of less than 2%, while the average OM content was 0.82%. According to Farag [68] and Reda [69], the dominant clay mineral type in the study area is montmorillonite, so the influence of clay mineral type on spectra may be the same for all soil samples.

4.2. Soil Salinity: Mapping and Assessment

The mapping of soil salinity with satellite imagery is based on the premise that the spectral reflectance of soil salinity is affected by the physical and chemical properties of the soil: quantity and mineralogy of salts, moisture content, soil texture, organic matter, and surface roughness [2]. Several researchers have studied the spectral response of salt crusts and other deposits recorded by Landsat sensors with simple regression models [23,64,70], and their findings justify why predictive regression models based on Landsat surface reflectance data have potential for use in the mapping of soil salinity. The good performance and accuracy of the predictive models in this study may be attributed to the effectiveness of atmospheric correction in removing radiometric distortions and the retrieval of true reflectance values [35,71]. Next, this study demonstrated that salt-affected soils exhibit a clear decrease in the spectral reflectance (Figure 2), consistent with the results of Metternicht and Zinck [2], Schmid et al. [72], and Sidike et al. [35]. In addition, our findings indicate that there is a high correlation between the measured reflectance spectra and the reflectance values derived from multispectral image data (Figure 4). This result indicates that the measured spectral reflectance and the Landsat data have great potential for predicting and mapping the soil salinity. This finding is in agreement with Sidike et al. [35], who found that a PLSR model based on measured reflectance spectra with multispectral image data is a promising tool for estimating and mapping the soil salinity. However, more investigation into factors influencing modeling and mapping accuracy—for example, types of salt minerals, types of clay minerals, soil moisture, and texture—should be considered in future research.

The PLSR models tended to underestimate ECe at high salinity levels, but matched well at a low salinity level (<16 dSm\(^{-1}\)), suggesting a non-linear relationship between the soil salinity and the soil reflectance [18]. The high R\(^2\) values obtained from the maps indicated that, similar to the MARS models, the PLSR models also learned well and had good generalization ability (Figure 5). The validation results
of ECe maps for 2006 and 2012 (0.71 < R² < 0.73 and 5.18 < RMSE < 5.37) were found to be good. Both the MARS and PLSR results (Figure 8) indicate that Landsat imagery provided suitable data for predicting and mapping the soil salinity over large areas if an accurate and robust predictive model is available.

The descriptive statistics derived from the soil salinity maps indicate that the predicted ECe values in the selected areas are lower than the observed ECe values. The reason behind the tendency of the MARS models to overestimate the measured ECe is not clear, but it can probably be attributed to spectral characteristics of the imagery or model generalization [18]. In spite of this tendency, the MARS models are satisfactory and resulted in very good R² values for all maps, indicating that the models mostly learned well and had good generalization ability (Figure 8b,d). The validation results of ECe maps for 2006 and 2012 (0.72 < R² < 0.74 and 5.23 < RMSE < 5.30) were very good.

Factors causing soil salinization include soil texture, land management practices (e.g., leaching, cultivation, and irrigation), a rising water table, and climate [73–75]. Therefore, the excess accumulation of salts at the soil surface is probably varied due to the distribution of different factors across the study area. For example, Akramkhanov et al. [74] found that the spatial distribution of soil salinity in the Khorezm Province, Uzbekistan, is likely due to soil texture, land management practices, and water table. We found that the spatial distribution of soil salinity was variable over the investigated areas. Patches with very strongly saline soil (>32 dSm⁻¹) were found mainly in non-vegetated and clayey wet areas. Wet areas, shown in dark brown and brown (Figure 1), form in the El-Tina Plain due to a rising water table through subsurface seepage of seawater and leakage from irrigation and drainage canals. In addition, dissolved salts in the irrigation water can gradually accumulate in the soil and shallow water table. Thus, the rising water table within two meters causes potentially very high salinization rates [73,76]. However, the salinization of dry land occurs when saline groundwater approaches the soil surface, especially during drier seasons; the high evaporation rate causes the accumulation of salts at the soil surface [70]. Thus, the findings of this study suggest that the spatial variation in soil salinity over the investigated area may be explained mostly by the dominance of clayey soils, the rising saline groundwater, and the accumulation of salts at the soil surface combined with a high evaporation rate. The confirmation of this explanation requires further work in the study area. Similar results in an arid region, Al Hassa Oasis, Saudi Arabia, were found by Albed et al. [63].

5. Conclusions

The study reveals that estimation of soil salinity with predictive models based on the resampled laboratory spectra is promising. Good results were achieved for predicting soil salinity based on the PLSR model (R² = 0.70, RMSE = 6.95, and RPD = 1.82) and even better for the MARS model (R² = 0.73, RMSE = 6.53, and RPD = 1.96). The efficiency of both models is in part due to the high correlation between the measured soil salinity and the soil spectra. The models were subsequently applied on a pixel-by-pixel basis to the reflectance values derived from two Landsat images (2006 and 2012) to map soil salinity in the study area of El-Tina Plain in Egypt. The accuracy assessment of the resulting maps with independent soil samples provided satisfactory results for both images and both models (MARS: R² = 0.72 and 0.74 for 2006 and 2012, respectively; PLSR: R² = 0.71 and 0.73 for 2006
and 2012, respectively). Overall, MARS was found to be a more suitable technique for the estimation and mapping of soil salinity than PLSR, especially in areas with high levels of salinity.

Previous studies (e.g., [28,32,33]) showed that MARS and PLSR allows for modeling of soil salinity under laboratory conditions with sufficient accuracy. This study extended these findings and demonstrated how these predictive models, coupled with freely available satellite imagery, could successfully map soil salinity over the large agricultural area of the El-Tina Plain in Egypt. Mapping the soil salinity with remotely sensed data, the successful approach presented here, can help decision makers by providing up-to-date digital soil maps to better manage salt-affected soils in the studied area. The main strength of this method, besides its accuracy in soil salinity mapping, is the lower cost as compared to conventional field-based approaches. Consequently, the predictive models described in this study can be used to model and map other soil properties, such as cation exchange capacity, total nitrogen and phosphorus, soil pH, and organic matter, clay, and CaCO₃ contents. Furthermore, using soil spectra resampling methods, the approach is robust enough to use various sensors for mapping and monitoring of agricultural arid areas, for example, the sensors onboard the newly launched Landsat 8. In future work we assume that the approach to mapping soil salinity presented in this paper should be tested with very high spatial resolution multispectral images, such as those provided by the WorldView-3 satellites. With their increasing availability, the spatial resolution and accuracy of salinity maps may most likely be improved, provided the approach elaborated in this study is fine-tuned to the specific properties of the very high resolution data.

Acknowledgments

This study was funded in part by Jagiellonian University, Poland, and by the Suez Canal University, Ismailia, Egypt. The corresponding author is grateful for the support from the Egyptian Ministry of Higher Education. The corresponding author would like to thank Mohamed Ismail and Abdellatif Deyab, Remote Sensing and GIS unit, Soil Research Center, Cairo, Egypt, for helping in spectroscopy field measurements.

Author Contributions

Said Nawar conceived and designed the experiments with advice of Henning Buddenbaum, Joachim Hill and Jacek Kozak. Said Nawar performed the acquisition, analysis and interpretation of the data. Said Nawar and Henning Buddenbaum performed the spectroscopy measurement and statistical modeling. All authors discussed the basic structure of the manuscript, and Said Nawar finished the first draft. Jacek Kozak, Henning Buddenbaum and Joachim Hill reviewed and edited the draft. All authors read and approved the submitted manuscript, agreed to be listed and accepted the version for publication.

Conflicts of Interest

All authors declare no conflict of interest.
References


© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).