Salinity Properties Retrieval from Sentinel-2 Satellite Data and Machine Learning Algorithms

Nada Mzid [1,*,†], **Olfa Boussadia** [2,†], **Rossella Albrizio** [3,*], **Anna Maria Stellacci** [4], **Mohamed Braham** [2] **and Mladen Todorovic** [5]

1 Department of Agriculture Forestry and Nature (DAFNE), University of Tuscia, 01100 Viterbo, Italy
2 Olive Institute, Avenue Ibn Khaldoun Tafala, Sousse 4000, Tunisia
3 Institute for Mediterranean Agricultural and Forestry Systems, National Research Council of Italy, P. le Enrico Fermi 1, 80055 Portici, Italy
4 Department of Soil Plant and Food Sciences, University of Bari Aldo Moro, Via G. Amendola 165/a, 70126 Bari, Italy
5 CIHEAM—Mediterranean Agronomic Institute of Bari, 70010 Valenzano, Italy
* Correspondence: nada.mzid@unitus.it (N.M.); rossella.albrizio@cnr.it (R.A.)
† These authors contributed equally to this work.

Abstract: The accurate monitoring of soil salinization plays a key role in the ecological security and sustainable agricultural development of semiarid regions. The objective of this study was to achieve the best estimation of electrical conductivity variables from salt-affected soils in a south Mediterranean region using Sentinel-2 multispectral imagery. In order to realize this goal, a test was carried out using electrical conductivity (EC) data collected in central Tunisia. Soil electrical conductivity and leaf electrical conductivity were measured in an olive orchard over two growing seasons and under three irrigation treatments. Firstly, selected spectral salinity, chlorophyll, water, and vegetation indices were tested over the experimental area to estimate both soil and leaf EC using Sentinel-2 imagery on the Google Earth Engine platform. Subsequently, estimation models of soil and leaf EC were calibrated by employing machine learning (ML) techniques using 12 spectral bands of Sentinel-2 images. The prediction accuracy of the EC estimation was assessed by using k-fold cross-validation and computing statistical metrics. The results of the study revealed that machine learning algorithms, together with multispectral data, could advance the mapping and monitoring of soil and leaf electrical conductivity.

Keywords: Mediterranean region; olive orchard; soil and leaf electrical conductivity; Google Earth Engine; spectral vegetation indices

1. Introduction

Soil salinization is a critical environmental problem in arid and semiarid regions globally since it seriously affects the ecological sustainability of limited land resources. As a form of land degradation, soil salinization greatly impacts on ecosystem services [1]. Thus, soil salinization is restricting agriculture's global development and affecting the social economy's growth [2,3]. At the same time, soil salinization is one of the most important factors causing direct adverse effects on soil characteristics, as it gravely affects soil resources, decreases both soil fertility and soil microbial activity; all this results in a sharp decline of soil productivity and nutrient availability [4]. Meanwhile, soil salinization can accelerate the desertification process and inhibit plants' absorption of water and nutrients, thereby affecting plants' physiological status [5].

Semiarid and arid regions of the Mediterranean are facing extraordinary pressure on already-degraded land and water resources, while the need for irrigation is increasing sharply [6]. This is due to the scarcity of rainfall, massive evaporation, high water-soluble

salt content, and increasing demand for food, water, and living materials from the rapidly growing population [7,8]. A severe threat to the sustainable development of regional agriculture and the economy is imposed. The timely and accurate acquisition of soil salinization information has paramount practical significance for irrigation and drainage management, for setting water and environmental policies [4,9], and for a selection of the most appropriate and adapted crops to salinity challenges.

Traditional methods for monitoring soil salinization rarely obtain large-scale distribution information [10]. It is always convenient to consider electrical conductivity as a standard measurement of salinity [11,12], since the electrical conductivity (EC) of a soil, vegetation, or water sample is influenced by the concentration and composition of dissolved salts [13,14]. Remote sensing techniques present great advantages in providing spectral soil property information at large spatial scales, and repeatedly with short temporal intervals [15]. In recent years, soil spectral characteristics have been used to estimate soil organic matter [16], total nitrogen, heavy metals [17], and soil moisture content [18]. Currently, there are many satellites with sensors with moderate to high spatial and temporal resolutions, which provide new opportunities for monitoring the spatial distribution of soil salinization using digital soil mapping techniques [15,16,19]. Indeed, the soil spectrum is a comprehensive reflection of various soil physical and chemical properties [17,18,20,21].

The theoretical basis for remote sensing monitoring is building upon the spectral characteristics of saline soils. The soil reflectance increases with the increase in soil salinization in visible, near-infrared, and shortwave infrared bands [22,23]. However, in practice, it is complicated to obtain the pure spectral information of saline soils using remote sensing due to interference from other factors, such as soil moisture, vegetation cover, and data acquisition time. Overall, the monitoring accuracy in the dry season is usually higher than in the wet season [24].

Various machine learning (ML) models are assessed to find the model that maximizes the prediction accuracy for a specific phenomenon. In this regard, Xu et al. [25] proposed a new method for simultaneously identifying the hyper-parameters and input features of the support vector machine regression algorithm based on an adaptive genetic algorithm for the quantitative evaluation of soil salinization. The authors of [3] combined Sentinel-2 Multispectral Imager (MSI) data and MSI-derived covariates with measured soil salinity data and applied three machine learning algorithms (support vector machines, artificial neural network, and random forest) to estimate and map the soil salinity in a sample study area in China and provided a scientific basis for the simulation of soil salinization scenarios in arid areas. Recently, Xiao et al. [11] evaluated the performance of three machine learning models, i.e., random forest, support vector machine, and extreme gradient boosting, in predicting soil salinity variables (such as total dissolved ionic matter, potential salinity, sodium adsorption ratio, exchangeable sodium percentage, residual sodium carbonate, and magnesium adsorption ratio in soils) and thus optimized the variable input combinations. Furthermore, the main objective of [26] was to map soil salinity intrusion using Sentinel-1 Synthetic Aperture Radar C-band data combined with five machine learning models. The authors concluded that advanced machine learning models could be used for mapping soil salinity.

The groundwater resources of Tunisia are of low quality, and only 50% of these resources have water salinity below 1.5 g/L. Tunisia is a country that faces the risk of soil salinization through the combination of its arid climate and the poor quality of its waters and soils [27]. Salt-affected soils cover about 1.5 million hectares (10% of the land) [28], which are located mainly in the central part of the country [29]. This risk will be aggravated in the upcoming years, along with increased water scarcity. In order to avoid the risk of salinization, it is important, on the one hand, to control the soil salinity and to keep it below plant salinity tolerance thresholds and, on the other hand, to comprehensively analyze the area where water is applied.

Even though previous studies [30,31] explored the ability of some models to estimate soil salinity, they have yet to be applied for soil and leaf EC prediction in the Tunisian case. In this study, we investigated the potential of recently developed retrieval algorithms designed to quantify EC traits from the spectroscopic imagery of the multispectral instrument Sentinel-2 (S-2 MSI) satellite in a southern Mediterranean country (central Tunisia). We selected an experimental agricultural area in the Kondar region and tested different processing approaches using parametric regression methods and machine learning algorithms. In this regard, this research performed three primary investigations, which comprise (i) an evaluation of the performance of a large number of machine learning algorithms; (ii) the use of the whole spectral band set of the multispectral imager satellite instead of a selection of specific wavelengths; and (iii) the retrieval of electrical conductivity at both the soil and leaf levels.

2. Materials and Methods

2.1. Study Area and Experimental Setup

The experiments were conducted at the olive orchard of a Kondar farm, located in the Sousse region, Tunisia (35°56′ N, 10°14′ E). The climate is of the Mediterranean type, having an average annual rainfall of 271 mm. Average temperatures reach 20–21 °C. The soil is sandy loam, with a field capacity and wilting point of 15.6% and 8.6%, respectively, according to the USDA methodology and granulometric analysis conducted in the laboratory.

The area of the experimental farm covers 30 ha. The olive orchard under study included 80 trees (*Olea europaea* cv. Koroneiki) with a spacing of 6 m × 7 m. Among them, 48 trees were monitored for experimental measurements. Olive trees were cultivated in two growing seasons (March–November 2017 and March–November 2018) under three irrigation treatments: full irrigation (FI), deficit irrigation (DI), and rainfed treatment (RF). Each treatment was replicated four times. A completely randomized block design was adopted; the three irrigation treatments were randomly distributed in each block, which led to a total of twelve plots. Water was supplied to plots via a drip irrigation system. Irrigation water was withdrawn from a well within the farm with a water electrical conductivity equal to 7.3 mS/cm. The soil EC measured on the saturated paste extract was 5 mS/cm.

The crop water balance and irrigation scheduling were managed using an Excel-based model [32]. The model estimates crop evapotranspiration, irrigation water requirements, and relative yield through the standard procedure proposed in FAO Irrigation and Drainage Paper 56 [33]. Water requirements were considered for the following fruit growth stages: cluster formation (stage C), full bloom (stage F1), fruit set (stage H), fruit growth stage 1 (stage I), and fruit growth stage 2 (stage I1) (Tables 1 and 2).

Table 1. Observed phenological stages during the growing season March–November 2017.

Dates	8 April 2017	6 May 2017	21 May 2017	2 June 2017	23 July 2017
Julian days	98	126	141	153	174
Stage	Cluster formation	Full bloom	Fruit set	Fruit growth (stage 1)	Fruit growth (stage 2)
Symbol	C	F1	H	I	I1

Table 2. Observed phenological stages during the growing season March–November 2018.

Dates	11 April 2018	25 April 2018	24 May 2018		25 June 2018	31 July 2018
Julian days	101	115	136	144	176	212
Stage	Cluster formation	Full bloom	Fruit set		Fruit growth (stage 1)	Fruit growth (stage 2)
Symbol	C	F1	H		I	I1

Based on the soil water balance calculation, the effective net irrigation supply for the fully irrigated trees (FI) was 801 m³/ha and 620 m³/ha for 2017 and 2018, respectively. Half of these volumes were supplied in the case of deficit irrigation (DI).

2.2. Ground-Based Field Measurements

2.2.1. Soil Electrical Conductivity

Soil electrical conductivity (EC_{Soil}) measurements were performed periodically (once per week) throughout the irrigation season by collecting a composite sample for each experimental plot. The protocol followed involved collecting soil samples located within an Elementary Sample Unit (ESU) scheme set up to fit the Sentinel-2 spatial resolution, i.e., according to 10 by 10 m quadrats. Each ESU contained 5 sampling points, defined by establishing a 14 m transect (diagonal) in each plot (center of the plot and 4 points positioned 3.5 m away from the center). At each point, three soil samples were collected with an Ejkelkamp auger in the 0–40 cm depth layer. Soil samples were collected from a total of 12 ESUs for each data set, which implies a total of 120 ESUs in the two growing seasons. Soil sampling was carried out to include the active layer affected by agricultural practices and was performed on dates close to the pass of the satellite.

After the field work, soil electrical conductivity was measured in the laboratory through the saturated paste extract method. The soil samples were collected and dried in the air. The soil was sieved in the laboratory through a 2 mm mesh, and 200 g of soil for each sample was weighed for the preparation of the saturated paste. Subsequently, the saturated paste was introduced in a pressure pump for filtration, and soil electrical conductivity was measured through a portable conductivity meter [34].

2.2.2. Leaf Electrical Conductivity

Leaf samples were periodically (once per week) collected throughout the irrigation season in four replicates for each treatment, which led to a total of twelve plants selected from the investigated plots. Data were taken across the experimental orchard from the cluster formation stage until the end of the harvest in both growing seasons. As a consequence, a total of 120 leaf samples were collected during both growing seasons. Measurements were made after the irrigation and on dates close to the pass of the satellite.

The collection was carried out by sampling leaves from each plant at 1.5–1.8 m from the soil surface on the four sides, N-S-E-W. Leaves were fully intact, clean, dry, green, and free of signs of disease or damage.

For each composite leaf sample, 2 g leaf discs 0.6 cm in diameter were prepared with a paper hole puncher and put in glass tubes in the laboratory. Then, 20 mL of distilled water was added to each tube, and the whole sample was left at ambient temperature for 24 h. Afterward, leaf electrical conductivity (EC_{Leaf}) (mS/cm) was measured using a portable conductivity meter [35].

2.3. Satellite Remote Sensing Analysis

In order to obtain the corresponding spectral acquisitions for the in-situ data, Sentinel-2 Level-2A orthorectified atmospherically corrected surface reflectance (L2A) images were requested using the GEE catalog. Each Level 2A product is composed of 100 × 100 km² tiles in cartographic geometry (UTM/WGS84 projection). The Sentinel-2

products used in this study corresponded to the MSI covering 12 spectral bands (443–2190 nm), with a swath width of 290 km and spatial resolutions of 10 m (four visible and near-infrared bands), 20 m (six red-edge and shortwave infrared bands), and 60 m (coastal aerosol, water vapor, and cirrus bands).

The processing and analysis steps of the remote sensing datasets were performed using a self-prepared script written on the Google Earth Engine platform.

The web-based IDE for the Earth Engine JavaScript Application Programming Interface (API) package *ee* provides functions that allow the extraction of any available information layers over a specific area of interest (AOI; Kondar site) and the very efficient processing of the resulting datasets. In this regard, the shapefile of the trial, including plots with different water treatments, was uploaded using the Imports section of the GEE code.

In all subsequent steps, all processing was bounded by the AOI.

Sentinel-2 images for the study area should be selected on specific dates, considering the crop phenological stages (and in consequence, the ground sampling dates) on the Google Earth Engine (GEE) code editor platform for the period between 1 March 2017 and 31 November 2018 (Table 3).

Table 3. Selected dates (based on crop phenological stages) of Sentinel-2 images for the study area located in the south Mediterranean–Kondar, Tunisia.

Sentinel-2 Selected Dates	
Agricultural Year 2016–2017	**Agricultural Year 2017–2018**
01 April 2017	11 April 2018
11 April 2017	06 May 2018
31 May 2017	26 May 2018
10 June 2017	30 June 2018
25 July 2017	30 July 2018

During this step, the satellite data (Sentinel-2) were imported using the *ee.ImageCollection* function, and a filter was applied to the image collection to include only relevant data that support the purpose of the research. The filter considers the dates (*filterDate*), cloud cover (*Filter.lt*), spatial extent (*filterBounds*), and selective bands (*select*). As a result, (i) ten out of all available Sentinel-2 MSI images were considered, (ii) a maximum of 20% cloud presence was set, and (iii) from the 12 available Sentinel-2 MSI bands, eight were employed for the calculation of the spectral indices, covering the visible, near-infrared (NIR), vegetation red-edge (VREdge 1 and VREdge 4), and shortwave infrared (SWIR1 and SWIR2) bands with central wavelengths of 490 nm, 560 nm, 665 nm, 842 nm, 705 nm, 865 nm, 1610 nm, and 2190 nm.

Since the selected bands have different spatial resolutions (10 m and 20 m), the corresponding 20 m bands were resampled to the 10 m band spatial resolution via the *resample* function.

Furthermore, the study area was divided into two parts, (I) olive grove area and (II) non-olive-grove area, to separate the exposed olive trees and bare soil area from other land covers in the experimental site (e.g., uncultivated land). The outline of the olive area was defined using the land cover map of Tunisia coordinated by the Observatoire du Sahara et du Sahel. Maps of the different functional land cover traits were obtained by importing the corresponding shapes into the GEE environment. The other classes of maps were excluded.

2.4. Assessment of the Appropriate Spectral Indices for Electrical Conductivity Property Estimation

A preliminary assessment of the most widely investigated spectral indices was carried out in order to select those that have already been recognized as good descriptors of salinity status.

For this purpose, a set of different salinity-, chlorophyll-, vegetation-, and water-related spectral index families were selected. Definitions, formulas, and references are summarized in Table 4. For each considered band, the specific central wavelengths are those indicated in the original formulation proposed. NDSI, S1, SI, NDVI, NMDI, SAVI, EVI, TCARI/OSAVI, MCARI, CVI, GDVI, SIWSI, and MSI were selected for the assessment of the EC property estimation using the available dataset.

Table 4. List of selected spectral indices related to salinity estimation.

Index		Formula		Reference
		Salinity spectral indices		
NDSI	Normalized Differential Salinity Index	$(R_{665} - R_{842})/(R_{665} + R_{842})$	$(R_{Red} - R_{NIR})/(R_{Red} + R_{NIR})$	[36]
S1	Salinity Index 1	R_{490}/R_{665}	R_{Blue}/R_{Red}	[37]
SI	Salinity Index	$\sqrt{R_{490} * R_{665}}$	$\sqrt{R_{Blue} * R_{Red}}$	[38]
		Vegetation spectral indices		
NDVI	Normalized Difference Vegetation Index	$\dfrac{(R_{842} - R_{665})}{(R_{842} + R_{665})}$	$\dfrac{(R_{NIR} - R_{Red})}{(R_{NIR} + R_{Red})}$	[39]
SAVI	Soil-Adjusted Vegetation Index	$\dfrac{1.5(R_{842} - R_{665})}{(R_{842} + R_{665} + 0.5)}$	$\dfrac{1.5(R_{NIR} - R_{Red})}{(R_{NIR} + R_{Red} + 0.5)}$	[40]
EVI	Enhanced Vegetation Index	$2.5\dfrac{(R_{842} - R_{665})}{(R_{842} + (6R_{665}) - (7.5R_{490}) + 1)}$	$2.5\dfrac{(R_{NIR} - R_{Red})}{(R_{NIR} + (6R_{Red}) - (7.5R_{Blue}) + 1)}$	[41,42]
GDVI	Generalized Difference Vegetation Index	$GDVI^n = (R_{842}{}^n - R_{665}{}^n)/(R_{842}{}^n + R_{665}{}^n)$	$GDVI^n = (SR^n - 1)/(SR^n + 1)$ $= (R_{NIR}{}^n - R_{Red}{}^n)/(R_{NIR}{}^n + R_{Red}{}^n)$ n: Power, an integer of the values of 1, 2, 3, 4…n. GDVI ranges from −1 to 1. SR: simple ratio = R_{NIR}/R_{Red}	[43]
		Water spectral indices		
NMDI	Normalized Multiband Drought Index	$\dfrac{(R_{860} - (R_{1640} - R_{2130}))}{(R_{860} + (R_{1640} - R_{2130}))}$	$\dfrac{(R_{VREdge4} - (R_{SWIR1} - R_{SWIR2}))}{(R_{VREdge4} + (R_{SWIR1} - R_{SWIR2}))}$	[44]
SIWSI	Shortwave Infrared Water Stress Index	$\dfrac{(R_{1640} - R_{850})}{(R_{1640} + R_{850})}$	$\dfrac{(R_{SWIR1} - R_{VREdge4})}{(R_{SWIR1} + R_{VREdge4})}$	[45]

MSI	Moisture Stress Index	$\dfrac{R_{1600}}{R_{820}}$	$\dfrac{R_{SWIR1}}{NIR}$	[46,47]
		Chlorophyll spectral indices		
TCARI/OSAVI	Transformed Chlorophyll Absorption Reflectance Index/ Optimized Soil Adjusted Vegetation Index	$\dfrac{3(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})\dfrac{R_{700}}{R_{670}}}{(1 + 0.16)\dfrac{R_{800} - R_{670}}{R_{800} + R_{670} + 0.16}}$	$\dfrac{3(R_{VREdge1} - R_{Red}) - 0.2(R_{VREdge1} - R_{Green})\dfrac{R_{VREdge1}}{R_{Red}}}{(1 + 0.16)\dfrac{R_{NIR} - R_{Red}}{R_{NIR} + R_{Red} + 0.16}}$	[48]
CVI	Chlorophyll vegetation index	$\dfrac{(R_{842} * R_{665})}{R_{560}^2}$	$\dfrac{(R_{NIR} * R_{Red})}{R_{Green}^2}$	[49]
MCARI	Modified Chlorophyll Absorption in Reflectance Index	$[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})](R_{700}/R_{670})$	$[(R_{VREdge1} - R_{Red}) - 0.2(R_{VREdge1} - R_{Green})](R_{VREdge1}/R_{Red})$	[38]

R: Reflectance for the specified spectral wavelength.

Following the steps mentioned in Section 2.3, from the Sentinel-2 scenes, the spectral indices were extracted at the location of the ESUs, i.e., within a 10 × 10 m Sentinel pixel, by first defining the expression of each index on GEE and then mapping it for each Sentinel-2A image. The required indices were downloaded to a personal Google Drive account.

Subsequently, simple linear regression models were computed with EC_{Soil} and EC_{Leaf} as response variables and the selected spectral indices as predictors. The optimal index was then deduced according to the maximum determination coefficient (R^2), the lowest root-mean-square deviation (RMSE), and the probability level (P). Data were statistically analyzed using R (Auckland, New Zealand) (4.0.2, 2018) and the packages *stats* and *metrics*.

Chlorophyll indices (TCARI/OSAVI and MCARI) based on 550 nm, 670 nm, 700 nm, and 800 nm were proposed here due to their sensitivity to salinity [50]. The chlorophyll indices were calculated by replacing the 550 nm, 670 nm, 700 nm, and 800 nm bands with the green, red, vegetation red edge 1, and NIR bands available in the Sentinel-2A band set, as was also confirmed by other authors for both TCARI/OSAVI [51–55] and MCARI [52,56–58] computation.

Water index retrieval from Sentinel-2A imagery for the study sites was carried out via the Sentinel-2A band set proposed in the literature specifically for NMDI [59–64], SIWSI [65–68], and MSI [69,70]. On the basis of calculating the existing hyperspectral indices, red edge band 4 of the Sentinel-2 imagery was introduced to replace the 860 nm band of the NMDI so as to calculate the water indices. The Shortwave Infrared Water Stress Index and Normalized Multiband Drought Index used the SWIR1 band instead of the 1640 nm band. Bands 820 nm, 850 nm, 1600 nm, and 2130 nm used bands 4, 8a, SWIR1, and SWIR2 of the Sentinel-2A band set, respectively, to replace the near-infrared, vegetation red-edge and shortwave infrared bands of the NMDI, SIWSI, and MSI.

The calculation method of the relevant spectral indices is shown in Table 4.

2.5. Electrical Conductivity Variables' Estimation Using Machine Learning Algorithms

Electrical conductivity property estimation models were built starting from measured EC_{Soil} and EC_{Leaf} based on soil and plant ground sampling in the field and the spectral data acquired by the Sentinel-2 satellite MSI sensor (processed using the GEE code script).

The spectral profiles of the corresponding Sentinel-2 satellite imagery were downloaded from the Google Earth Engine to a Google Drive account. Indeed, following the steps mentioned in Section 2.3, from the Sentinel-2 scenes, the reflectance was extracted at the location of the ESUs, i.e., within a 10 × 10 m Sentinel pixel, using the *batch.Download.ImageCollection.toDrive* function in GEE.

All estimation models were calibrated by testing machine learning algorithms available in the Automated Radiative Transfer Models Operator (ARTMO) [71].

ARTMO is a modular GUI toolbox developed in Matlab that embodies a suite of leaf and canopy radiative transfer models (RTMs), including PROSAIL and several retrieval toolboxes, i.e., a spectral indices toolbox [72], a machine learning regression algorithm toolbox [73], and a LUT-based inversion toolbox [74]. The full software framework can be freely downloaded from artmotoolbox.com. The machine learning regression algorithm (MLRA) assessment toolbox provides a suite of nonparametric techniques to enable semiautomatic mapping of the electrical conductivity variables.

The input module is the first mandatory step to be configured. Two types of input data (TXT in a matrix format) were required within the same file: the parameters to be estimated, i.e., EC_{Soil} and EC_{Leaf}, and the related spectra. Once the input data had been configured, the MLRA scenarios were identified. Indeed, the capabilities of the following statistical non-parametric regression algorithms were computed for the evaluation: least-squares linear regression, partial least-squares regression, regularized least-squares regression, principal component regression, Elastic Net regression, K-nearest neighbor regression, weighted k-nearest neighbor regression, regression tree, regression tree (LS boosting), boosting trees, bagging trees, Gradient Boosting/Boosted Trees, random forest (TreeBagger), Canonical Correlation Forests, Extreme Learning Machine, relevance vector machine, kernel ridge regression, kernel signal-to-noise ratio, Gaussian Process Regression, Sparse Spectrum Gaussian Process Regression, Warped Gaussian Process Regression, and VH Gaussian Process Regression.

The next step was the configuration of the k-fold cross-validation (k = 10). Once the input data had been provided and the MLRA settings (single-input) had been configured, these scenarios were run. Once all scenarios had been analyzed, an overview table with the best validation results indicated the degree of association between estimated and observed values of the same variable. To compare the estimation accuracy, the metrics included the Mean Absolute Error (MAE), the root-mean-square error (RMSE), the coefficient of correlation (R), the coefficient of determination (R^2), and the Nash–Sutcliffe efficiency (NSE).

3. Results

3.1. Descriptive Statistics of the Soil and Leaf Samples Electrical Conductivity (EC) Values

The EC_{Soil} for the whole dataset varied between 2.729 and 6.830 mS/cm. The mean and median values were 4.939 mS/cm and 4.800 mS/cm, respectively, with a standard deviation (SD) of 1.077 and a coefficient of variation (CV) of 21.805%. From a summary of the descriptive statistics of the dataset employed (Table 5), it appeared that a quite moderate variability in EC_{Soil} was observed in the field.

Table 5. Descriptive statistics of the soil and leaf samples' electrical conductivity (EC, mS/cm) dataset.

Variable	Min	Max	Mean	Median	SD	CV (%)	Skewness
EC_{Soil}	2.729	6.830	4.939	4.800	1.077	21.805	0.012
EC_{Leaf}	0.536	1.690	0.898	0.859	0.210	23.395	1.159

Regarding the EC_{Leaf}, the dataset varied between 0.536 and 1.690 mS/cm. The mean and median values were 0.898 mS/cm and 0.859 mS/cm, respectively, with an SD of 0.210 and a CV of 23.395%.

Due to the strong evaporation and low precipitation, it was difficult for salt to leach from the surface soil to soil layers below the 20 cm depth, and thus, the salts often accumulated on the soil surface.

3.2. Assessment of the Spectral Indices for Electrical Conductivity Property Estimation

The regression analysis was performed to examine the sensitivity of the Sentinel-2 MSI-derived spectral indices to the soil and leaf EC. The results of the statistical analysis are presented in Table 6.

Table 6. Coefficient of determination (R^2), root-mean-square error (RMSE), and *p*-level statistical significance for soil (EC_{Soil}) and leaf (EC_{Leaf}) electrical conductivity and spectral indices derived from Sentinel-2 multispectral images.

Spectral Indices		EC_{Soil}			EC_{Leaf}		
		R^2	RMSE	*p*	R^2	RMSE	*p*
Salinity indices	NDSI	0.157	5.30	***	0.154	1.17	***
	S1	0.702	4.07	***	0.052	0.28	*
	SI	0.573	4.84	***	0.263	0.72	***
Chlorophyll indices	TCARI/OSAVI	0.314	4.17	***	0.147	0.21	***
	MCARI	0.357	4.11	***	0.187	0.28	***
	CVI	0.415	4.34	***	0.030	0.29	n.s.
Vegetation indices	NDVI	0.001	4.76	n.s.	0.308	0.70	***
	SAVI	0.003	4.65	n.s.	0.341	0.59	***
	EVI	0.004	4.20	n.s.	0.309	0.44	***
	GDVI	0.310	5.19	***	0.460	0.70	***
Water indices	NMDI	0.002	4.45	n.s.	0.355	0.41	***
	SIWSI	0.076	4.66	*	0.014	0.48	n.s.
	MSI	0.180	4.27	***	0.450	0.32	***

* Significant at $p \le 0.05$. *** Significant at $p \le 0.001$. n.s. Not significant.

The results of the regression analysis between the salinity indices and EC_{Soil} were significant at $p \le 0.001$. The strongest relationship was obtained between S1 and EC_{Soil} (R^2 = 0.70). The regression models between the same electrical conductivity variable and the chlorophyll indices were also significant at $p \le 0.001$. However, for the vegetation and water spectral indices generated from Sentinel-2 MSI spectral data and EC_{Soil}, significant relationships were observed only for GDVI, SIWSI, and MSI (Table 5). It is worth noting that chlorophyll indices (TCARI/OSAVI, MCARI, and CVI) showed better results than the broadband vegetation indices in estimating EC_{Soil}, with R^2 values ranging between 0.31 and 0.41.

The relationships between the same spectral indices and EC_{Leaf} were significant at $p \le 0.001$ and $p \le 0.05$. From a list of thirteen spectral indices, only two, namely, CVI and SIWSI, showed non-significant relationships. However, the four categories of indices presented low (for salinity and chlorophyll indices) to moderate (for vegetation and water

indices) relationships. The best model was obtained between GDVI and EC$_{Leaf}$, with a coefficient of determination of 0.46.

3.3. Electrical Conductivity Variables Estimation Using Machine Learning Algorithms

The original Sentinel-2 MSI images and their spectral bands were adopted as remote sensing data sources to estimate both EC$_{Soil}$ and EC$_{Leaf}$. With the use of 12 spectral bands as the independent variables and with the soil and leaf EC data as the dependent variables, machine learning regression algorithms (MLRAs) were constructed (Tables 7 and 8). In Table 7 and Table 8, we present only the statistical metric outputs of the machine learning algorithms with acceptable ranges. The models that presented very low performance are not shown. The predicted EC$_{Soil}$ and EC$_{Leaf}$ based on different tested models were validated against the measured EC$_{Soil}$ and EC$_{Leaf}$ to evaluate the modeling effect and accuracy. Five statistical metric parameters, namely, MAE, RMSE, R, R^2, and NSE, were considered for the evaluation.

Table 7. Estimation accuracy of the best-performing prediction models of EC$_{Soil}$ derived from Sentinel-2 spectra. MAE: Mean Absolute Error; RMSE: root-mean-square error; R: coefficient of correlation; R^2: coefficient of determination; NSE: Nash–Sutcliffe efficiency.

MLRA	MAE	RMSE	R	R^2	NSE
Sparse Spectrum Gaussian Process Regression	0.601	0.720	0.716	0.513	0.510
Warped Gaussian Process Regression	0.620	0.762	0.716	0.4524	0.452
Canonical Correlation Forests	0.634	0.764	0.673	0.4834	0.449
Random forest (Tree Bagger)	0.614	0.766	0.695	0.4496	0.447
Bagging trees	0.622	0.769	0.671	0.4523	0.442
Gaussian Process Regression	0.651	0.794	0.673	0.4096	0.406
Boosting trees	0.616	0.796	0.640	0.4673	0.402
VH Gaussian Process Regression	0.664	0.796	0.684	0.4059	0.401
Regression tree	0.616	0.803	0.634	0.4578	0.392
Gradient Boosting/Boosted Trees	0.673	0.810	0.677	0.3981	0.381
Regression tree (LS boosting)	0.637	0.855	0.631	0.4228	0.310
Kernel signal-to-noise ratio	0.707	0.866	0.650	0.2995	0.293
Relevance vector machine	0.700	0.878	0.547	0.2766	0.273
Weighted k-nearest neighbor regression	0.712	0.879	0.526	0.2741	0.271
Kernel ridge Regression	0.754	0.894	0.524	0.3027	0.245
Elastic Net regression	0.770	0.944	0.550	0.1606	0.158
K-nearest neighbor regression	0.777	0.974	0.401	0.1381	0.105
Regularized least-squares regression	0.825	0.985	0.372	0.0966	0.084
Extreme Learning Machine	0.792	1.148	0.311	0.1768	-
Least-squares linear regression	-	-	-	-	-
Partial least-squares regression	-	-	-	-	-
Principal component regression	-	-	-	-	-
Adaptive Regression Splines	-	-	-	-	-
Support vector regression	-	-	-	-	-
Twin Gaussian process	-	-	-	-	-

Table 8. Estimation accuracy of the best-performing prediction models of EC_{Leaf} derived from Sentinel-2 spectra. MAE: Mean Absolute Error; RMSE: root-mean-square error; R: coefficient of correlation; R^2: coefficient of determination; NSE: Nash–Sutcliffe efficiency.

MLRA	MAE	RMSE	R	R^2	NSE
Gaussian Process Regression	0.097	0.132	0.790	0.625	0.624
Kernel ridge Regression	0.098	0.134	0.797	0.636	0.613
Canonical Correlation Forests	0.101	0.134	0.813	0.661	0.597
Relevance vector machine	0.106	0.140	0.761	0.579	0.577
VH Gaussian Process Regression	0.098	0.140	0.761	0.579	0.576
Kernel signal-to-noise ratio	0.104	0.146	0.740	0.547	0.541
Sparse Spectrum Gaussian Process Regression	0.107	0.151	0.725	0.526	0.508
Weighted k-nearest neighbor regression	0.116	0.164	0.662	0.439	0.415
Extreme Learning Machine	0.128	0.172	0.627	0.393	0.362
Adaptive Regression Splines	0.127	0.172	0.678	0.460	0.360
Bagging trees	0.121	0.176	0.649	0.421	0.331
Boosting trees	0.132	0.177	0.601	0.361	0.322
Random forest (Tree Bagger)	0.119	0.177	0.574	0.330	0.319
Gradient Boosting/Boosted Trees	0.127	0.183	0.572	0.327	0.278
K-nearest neighbor regression	0.130	0.184	0.568	0.323	0.264
Elastic Net regression	0.139	0.185	0.530	0.281	0.258
Regularized least-squares regression	0.141	0.186	0.522	0.272	0.253
Regression tree	0.165	0.267	0.090	0.008	-
Warped Gaussian Process Regression	0.151	0.275	0.035	0.001	-
Regression tree (LS boosting)	0.182	0.281	0.108	0.012	-
Principal component regression	-	-	-	-	-
Least-squares linear regression	-	-	-	-	-
Partial least-squares regression	-	-	-	-	-
Support Vector Regression	-	-	-	-	-
Twin Gaussian process	-	-	-	-	-

According to the prediction results, some machine learning algorithms exhibited good accuracy. Tables 7 and 8 indicate that Sparse Spectrum Gaussian Process Regression and Gaussian Process Regression were the most robust models among the twenty-five models tested for the estimation of EC_{Soil} and EC_{Leaf}, respectively. The statistical results of the model parameters showed that, regarding the Sparse Spectrum Gaussian Process Regression model for EC_{Soil} estimation, the R and R^2 values were equal to 0.72 and 0.51, respectively, while in the Gaussian Process Regression model for EC_{Leaf} estimation, the MAE, R, and R^2 were equal to 0.097, 0.79, and 0.62, respectively.

Among the machine learning models for EC_{Soil} estimation (Table 7), Canonical Correlation Forests (CCF), random forest (RF), and bagging trees (BT) exhibited high prediction accuracy, with the CCF model having a low RMSE (0.764), high R (0.673), and low MAE (0.634). The RF model was very close to BT (RF: RMSE = 0.766, R = 0.695, MAE = 0.614; BT: RMSE = 0.769, R = 0.671, MAE = 0.622). Warped Gaussian Process Regression also showed good fitting results, with a coefficient of correlation between predicted and observed values of 0.716.

Regarding the estimation accuracy of the best-performing prediction models of EC_{Leaf} derived from Sentinel-2 spectra (Table 8), Canonical Correlation Forests and Sparse Spectrum Gaussian Process Regression results slightly improved, whereas for bagging trees and random forest, the results were slightly worse. In more detail, in the case of CCF, the values of MAE, RMSE, R, and R^2 were equal to 0.101, 0.134, 0.813, and 0.661, respectively, whereas RF presented MAE, RMSE, and R values of 0.119, 0.177, and 0.574, respectively.

Regarding the best-performing prediction models of EC_{Soil} and EC_{Leaf} derived from Sentinel-2 spectral bands (Sparse Spectrum Gaussian Process Regression and Gaussian Process Regression, respectively), in Figure 1 the relationships between measured and estimated values are reported.
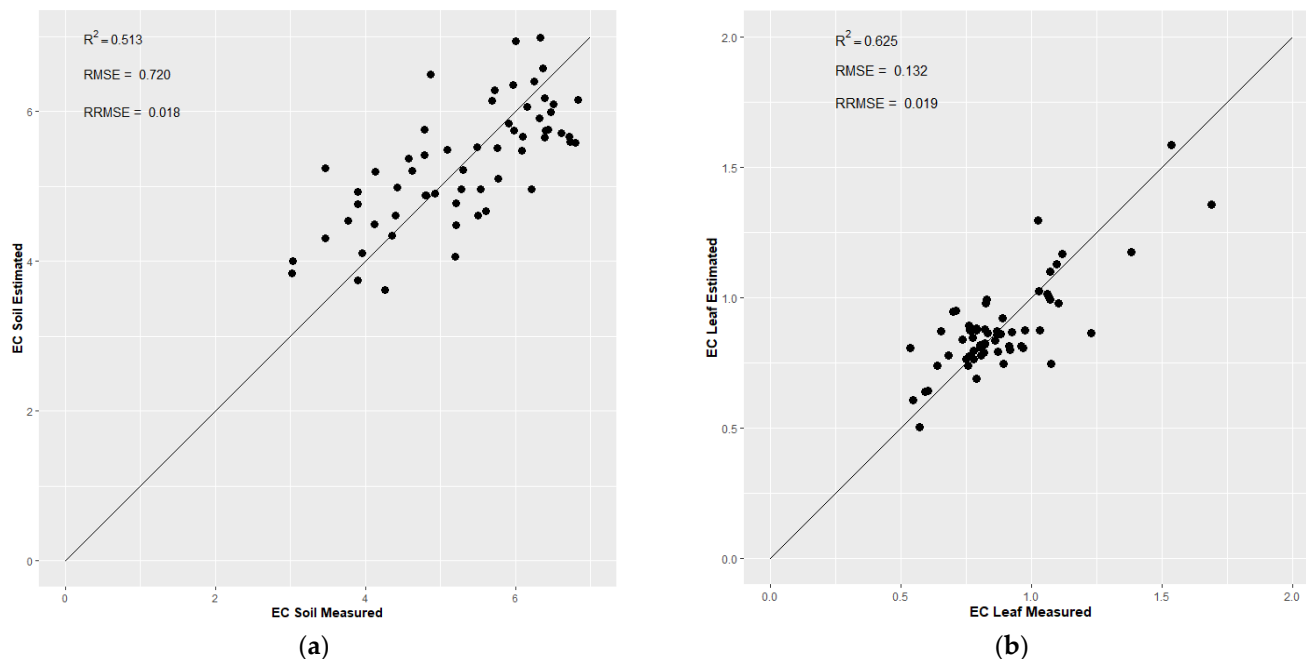

(a)　　　　　　　　　　　　　　　　　　　**(b)**

Figure 1. Best (**a**) soil and (**b**) leaf prediction models.

4. Discussion

Soil and leaf EC estimation models were established in this study based on the use of twenty-five machine learning algorithms. It was found that the Gaussian Process Regression models (Sparse Spectrum and Gaussian Processes) met the accuracy requirements and can be used for the quantitative estimation of the soil and leaf EC.

The response of vegetation spectral indices to soil EC is generally affected by many factors, including vegetation cover, salt tolerance, soil moisture, and soil type [75]. The results varied markedly, especially for leaf EC estimation, with better performance obtained from "vegetation indices" [76]. Indeed, in many previous studies, using salinity indices (salinity indices 1, 2, and 3) and vegetation indices (normalized difference infrared index, green normalized difference vegetation index, and a simple ratio) to estimate EC_{Soil} [77,78], the regions with vegetation coverage were directly identified as not salinized regions or slightly salinized regions [16,24,79]. As confirmed by [3], spectral indices are indispensable estimators for soil salinization monitoring. The relationship between spectral indices and electrical conductivity variables is established through fitting functions. However, the use of spectral indices is restricted to formulations that use only a few bands, with at most three or four bands, which implies a decrease in the complete-spectrum dataset [80]. Nowadays, this limits the strength of the estimation methods, as tens or hundreds of spectral bands are available, respectively, in current superspectral [81] or spaceborne hyperspectral sensors [82,83].

Soils with different salinity have different spectral characteristics, which is the basis of the remote sensing monitoring of soil salinization. The key to successfully estimating the soil and leaf salt content using spectral variables is to choose an effective model. Non-parametric regression algorithms (also referred to as machine learning regression algorithms—MLRAs) are data-driven approaches based on the definition of regression functions between the spectral information and the variables of interest. Machine learning algorithms are able to learn autonomously and can solve the problem of complex

nonlinear function approximation in estimation of soil and leaf electrical conductivity. The main use of these algorithms in this research is related to the possibility of training them with complete spectral information from the Sentinel-2 multispectral satellite. These adaptive algorithms can cope with the strong non-linearity inherent in remote sensing data [84,85]. As confirmed by [86,87], ML algorithms such as random forest (RF) [88] were demonstrated to be more robust to noisy features, small training sizes, and high dimensionality and collinearity. Indeed, in this case study, because the soil electrical conductivity dataset contains in situ sampling and the corresponding spectral characteristics, non-parametric regression models were applied. The methodology allows electrical conductivity prediction models adjusted to the input variables information and existing observations to be achieved.

As presented in the results (Section 3.3), the Gaussian Process Regression (GPR) models presented the best results and attained a higher performance and accuracy when estimating the electrical conductivity variables (EC_{Soil}: RMSE = 0.132 and R = 0.790; EC_{Leaf}: RMSE = 0.720 and R = 0.716). Indeed, GPR models are efficient machine learning regression algorithms for retrieving biophysical parameters [73]. GPR is based on a Bayesian approach for building generic regression models between the input (remote sensing data) and output (biophysical parameter) variables [81]. Gaussian Process Regression models are nonparametric as well as non-linear, while at the same time providing a ranking of relevant bands (features) from input spectral data obtained from remote sensing tools [89]. Gaussian Process Regression models are flexible enough to fit many types of data, including geospatial and time-series data. In the inference stage, every time a new observation is made, the model hypothesis (prior probability distribution) is updated considering the new observations [90]. For this reason, GPR was evaluated as top-performing in accurately reconstructing time-series datasets [91,92]. Our results agree with the findings of a previous study [26], revealing that GPR is a powerful tool that could be used for soil salinity mapping.

Moreover, both boosting and random forest algorithms presented close statistic results in the range of 0.7 and 0.6 for RMSE and R, respectively, when estimating EC_{Soil} and at the range of 0.2 and 0.6 for RMSE and R, respectively, when predicting EC_{Leaf}. From this perspective, some authors believed that the Gradient Boosting (GB) model added regularization terms, which were beneficial in reducing model complexity, avoiding the overfitting phenomenon, and increasing the generalization ability of the model [93]. In addition, GB supports the parallel splitting process, which improves the efficiency of model construction, optimizes the objective function, and ensures prediction accuracy [94]. As one of boosting algorithms, the GB model places more weight on more important factors in the training–testing process [95]. Therefore, this model shows a minimal reduction in model accuracy. Many studies concluded that the accuracy of the predictions of boosting algorithms depended mainly on the relevance of input parameters [96–98]. Our results agree with the findings of [2], in which the authors compared three machine learning models (RF, SVM, and GB) to predict soil salinity parameters. These authors found that GB and RF models showed good prediction ability for some salinity parameters, with a slightly higher performance of the GB model. The same research also showed that all models underestimated salinity variables for high salinity values, which may result in more challenging predictions in soils with higher potential salinity values However, [99] still recommended GB models for extreme values. Moreover, the authors of [100] found that the boosting model performed best, but the simulation results of the RF model were also acceptable.

On the other hand, the accuracy of the random forest model was high. Generally, decision-tree-based algorithms (e.g., RF) require no assumptions about the data distribution, adapt to outliers by isolating them in small regions of the feature space [101], have no hidden layers in their structure, and use tree algorithms for estimation on the basis of pattern recognition [102]. Our finding confirmed previous studies that indicated superior estimation capacity of tree-based models [103,104]. Indeed, due to the nonlinear

nature of many environmental phenomena, more flexible models with non-linear structures will yield better results [105].

From this perspective, among the tested machine learning algorithms, the results revealed that RF performed better than support vector regression (SVR), with higher accuracy and a lower root-mean-square error. Our results agree with those of [106]. Indeed, in their research, [106] combined a dataset consisting of Landsat 5 Thematic Mapper and ALOS L-band radar data acquired at the same time as field-measured salinity and used the support vector regression and random forest algorithms for salinity prediction. The results revealed that RF performed better than SVR. The SVR model showed low performance in learning with values of EC greater than 7 mS/cm [26]. These samples caused the low degree-of-fit of the model, and consequently, the model lacked the sensitivity required to predict values with high EC.

Generally, decision-tree-based algorithms (e.g., RF) perform better, particularly in comparison to data-intelligence algorithms with hidden layers in their structures [102]. Hence, GPR, GB, and RF are recommended for soil salinity estimation and mapping.

Recently, the increasing accessibility of the new generation of hyperspectral satellite sensors and the operational enhancements in the retrieval methods are forging the opportunity to attain up-to-date knowledge about the variability in salinity characteristics throughout the Earth system. With the upcoming stream of imaging spectroscopy data with the new, already-launched PRISMA and EnMAP hyperspectral satellites, there is a rising need to differentiate between methodologies, tools, and software applications exploiting the spectral possibilities to extract relevant information on an operational basis to quantify electrical conductivity variables. The quantitative estimation of soil traits considerably relies on the retrieval approach. Moreover, according to the literature, hybrid algorithms, in most cases, are more flexible and can enhance the prediction power of standalone models [107–110].

5. Conclusions

The results of this research provide a further contribution to the ability of Sentinel-2 spectral data to predict soil and leaf electrical conductivity in an agricultural area in central Tunisia. The performance of twenty-five regression algorithms for estimating electrical conductivity at the soil and vegetation levels was compared and quantitively assessed. Our results indicate that the combined use of machine learning algorithms and Sentinel-2 whole multispectral band set improved the accuracy of the electrical conductivity estimation.

An additional key goal of this research was to investigate the potential of recently developed retrieval regression methodologies for quantifying electrical conductivity not only at the soil level but also at the vegetation (leaf) level.

For future perspectives, the higher spectral resolution of newly launched hyperspectral satellites (e.g., PRISMA and EnMAP) will be a great tool to exploit narrow absorption spectral features in the Vis-NIR and SWIR spectral ranges associated with salinity estimation. Therefore, combining data from hyperspectral sensors with the most advanced machine learning algorithms could represent an innovative and robust strategy for improving the mapping and monitoring of soil and vegetation salinity properties.

References

1. Litalien, A.; Zeeb, B. Curing the Earth: A Review of Anthropogenic Soil Salinization and Plant-Based Strategies for Sustainable Mitigation. *Sci. Total Environ.* **2020**, *698*, 134235. https://doi.org/10.1016/J.SCITOTENV.2019.134235.

2. Xiao, Y.; Zhao, G.; Li, T.; Zhou, X.; Li, J. Soil Salinization of Cultivated Land in Shandong Province, China—Dynamics during the Past 40 Years. *Land Degrad Dev.* **2019**, *30*, 426–436. https://doi.org/10.1002/LDR.3236.

3. Wang, J.; Peng, J.; Li, H.; Yin, C.; Liu, W.; Wang, T.; Zhang, H. Soil Salinity Mapping Using Machine Learning Algorithms with the Sentinel-2 MSI in Arid Areas, China. *Remote Sens.* **2021**, *13*, 305. https://doi.org/10.3390/RS13020305.

4. Zhang, X.; Huang, B. Prediction of Soil Salinity with Soil-Reflected Spectra: A Comparison of Two Regression Methods. *Sci. Rep.* **2019**, *9*, 5067. https://doi.org/10.1038/s41598-019-41470-0.

5. Hafez, E.M.; Omara, A.E.D.; Alhumaydhi, F.A.; El-Esawi, M.A. Minimizing Hazard Impacts of Soil Salinity and Water Stress on Wheat Plants by Soil Application of Vermicompost and Biochar. *Physiol. Plant.* **2021**, *172*, 587–602. https://doi.org/10.1111/PPL.13261.

6. Deng, G.; Yao, X.; Jiang, H.; Cao, Y.; Wen, Y.; Wang, W.; Zhao, S.; He, C. Study on the Ecological Operation and Watershed Management of Urban Rivers in Northern China. *Water* **2020**, *12*, 914. https://doi.org/10.3390/W12030914.

7. Ma, L.; Yang, S.; Simayi, Z.; Gu, Q.; Li, J.; Yang, X.; Ding, J. Modeling Variations in Soil Salinity in the Oasis of Junggar Basin, China. *Land Degrad. Dev.* **2018**, *29*, 551–562. https://doi.org/10.1002/LDR.2890.

8. Vermeulen, D.; van Niekerk, A. Machine Learning Performance for Predicting Soil Salinity Using Different Combinations of Geomorphometric Covariates. *Geoderma* **2017**, *299*, 1–12. https://doi.org/10.1016/J.GEODERMA.2017.03.013.

9. Zhang, Y.; Biswas, A.; Adamchuk, V.I. Implementation of a Sigmoid Depth Function to Describe Change of Soil PH with Depth. *Geoderma* **2017**, *289*, 1–10. https://doi.org/10.1016/J.GEODERMA.2016.11.022.

10. Shoshany, M.; Goldshleger, N.; Chudnovsky, A. Monitoring of Agricultural Soil Degradation by Remote-Sensing Methods: A Review. *Int. J. Remote. Sens.* **2013**, *34*, 6152–6181. https://doi.org/10.1080/01431161.2013.793872.

11. Xiao, C.; Ji, Q.; Chen, J.; Zhang, F.; Li, Y.; Fan, J.; Hou, X.; Yan, F.; Wang, H. Prediction of Soil Salinity Parameters Using Machine Learning Models in an Arid Region of Northwest China. *Comput. Electron. Agric.* **2023**, *204*, 107512. https://doi.org/10.1016/j.compag.2022.107512.

12. Hossain, M.S.; Rahman, G.K.M.M.; Solaiman, A.R.M.; Alam, M.S.; Rahman, M.M.; Mia, M.A.B. Estimating Electrical Conductivity for Soil Salinity Monitoring Using Various Soil-Water Ratios Depending on Soil Texture. *Commun. Soil Sci. Plant Anal.* **2020**, *51*, 635–644. https://doi.org/10.1080/00103624.2020.1729378.

13. Corwin, D.L.; Yemoto, K. Salinity: Electrical Conductivity and Total Dissolved Solids. *Methods Soil Anal.* **2017**, *2*, 1442–1461,. https://doi.org/10.2136/msa2015.0039.

14. Corwin, D.L.; Lesch, S.M. Characterizing Soil Spatial Variability with Apparent Soil Electrical Conductivity: I. Survey Protocols. *Comput. Electron. Agric.* **2005**, *46*, 103–133. https://doi.org/10.1016/J.COMPAG.2004.11.002.

15. Chabrillat, S.; Ben-Dor, E.; Cierniewski, J.; Gomez, C.; Schmid, T.; van Wesemael, B. Imaging Spectroscopy for Soil Mapping and Monitoring; *Surveys in Geophysics*, **2019**; 40, 361-399, doi:10.1007/s10712-019-09524-0.

16. Ward, K.J.; Chabrillat, S.; Brell, M.; Castaldi, F.; Spengler, D.; Foerster, S. Mapping Soil Organic Carbon for Airborne and Simulated Enmap Imagery Using the Lucas Soil Database and a Local Plsr. *Remote Sens.* **2020**, *12*, 3451. https://doi.org/10.3390/RS12203451.

17. Omran, E.S.E. Rapid Prediction of Soil Mineralogy Using Imaging Spectroscopy. *Eur. Soil Sci.* **2017**, *50*, 597–612. https://doi.org/10.1134/S106422931705012X.

18. Zeng, W.; Xu, C.; Huang, J.; Wu, J.; Tuller, M. Predicting Near-Surface Moisture Content of Saline Soils from Near-Infrared Reflectance Spectra with a Modified Gaussian Model. *Soil Sci. Soc. Am. J.* **2016**, *80*, 1496–1506. https://doi.org/10.2136/SSSAJ2016.06.0188.

19. Gomez, C.; Oltra-Carrió, R.; Bacha, S.; Lagacherie, P.; Briottet, X. Evaluating the Sensitivity of Clay Content Prediction to Atmospheric Effects and Degradation of Image Spatial Resolution Using Hyperspectral VNIR/SWIR Imagery. *Remote Sens. Environ.* **2015**, *164*, 1–15. https://doi.org/10.1016/J.RSE.2015.02.019.

20. Sadeghi, M.; Babaeian, E.; Tuller, M.; Jones, S.B. Particle Size Effects on Soil Reflectance Explained by an Analytical Radiative Transfer Model. *Remote Sens. Environ.* **2018**, *210*, 375–386. https://doi.org/10.1016/J.RSE.2018.03.028.

21. Yuan, J.; Wang, X.; Yan, C.; Chen, S.; Wang, S.; Zhang, J.; Xu, Z.; Ju, X.; Ding, N.; Dong, Y.; et al. Wavelength Selection for Estimating Soil Organic Matter Contents through the Radiative Transfer Model. *IEEE Access* **2020**, *8*, 176286–176293. https://doi.org/10.1109/ACCESS.2020.3026813.

22. Sidike, A.; Zhao, S.; Wen, Y. Estimating Soil Salinity in Pingluo County of China Using QuickBird Data and Soil Reflectance Spectra. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 156–175. https://doi.org/10.1016/J.JAG.2013.06.002.

23. Nawar, S.; Buddenbaum, H.; Hill, J. Estimation of Soil Salinity Using Three Quantitative Methods Based on Visible and Near-Infrared Reflectance Spectroscopy: A Case Study from Egypt. *Arab. J. Geosci.* **2015**, *8*, 5127–5140. https://doi.org/10.1007/S12517-014-1580-Y.

24. Ding, J.; Yu, D. Monitoring and Evaluating Spatial Variability of Soil Salinity in Dry and Wet Seasons in the Werigan–Kuqa Oasis, China, Using Remote Sensing and Electromagnetic Induction Instruments. *Geoderma* **2014**, *235–236*, 316–322. https://doi.org/10.1016/J.GEODERMA.2014.07.028.

25. Xu, H.; Chen, C.; Zheng, H.; Luo, G.; Yang, L.; Wang, W.; Wu, S.; Ding, J. AGA-SVR-Based Selection of Feature Subsets and Optimization of Parameter in Regional Soil Salinization Monitoring. *Int. J. Remote Sens.* **2020**, *41*, 4470–4495. https://doi.org/10.1080/01431161.2020.1718239.

26. Hoa, P.V.; Giang, N.V.; Binh, N.A.; Hai, L.V.H.; Pham, T.-D.; Hasanlou, M.; Bui, D.T. Soil Salinity Mapping Using SAR Sentinel-1 Data and Advanced Machine Learning Algorithms: A Case Study at Ben Tre Province of the Mekong River Delta (Vietnam). *Remote Sens.* **2019**, *11*, 128. https://doi.org/10.3390/RS11020128.

27. Mzid, N. Assessment of Remote-Based Vegetation Indices to Face Agriculture Challenges in The Mediterranean. 2021, 1-134.

28. DGACTA. Examen et Évaluation de La Situation Actuelle de La Salinisation Des Sols et Préparation d'un Plan d'action de Lutte Contre Ce Fléau Dans Les Périmètres Irrigués En Tunisie. Phase 2: Ebauche Du Plan D'action. *DGACTA,* Ministère de l'agriculture et des ressources hydrauliques (Tunisie), **2007**.

29. Hachicha, M. Les Sols Salés et Leur Mise En Valeur En Tunisie. *Sécher. Montroug.* **2007**, *18*, 45–50. https://doi.org/10.1684/sec.2007.0063.

30. Zarai, B.; Walter, C.; Michot, D.; Montoroi, J.P.; Hachicha, M. Integrating Multiple Electromagnetic Data to Map Spatiotemporal Variability of Soil Salinity in Kairouan Region, Central Tunisia. J Arid Land 2022, 14, 186–202, doi:10.1007/S40333-022-0052-6/METRICS.

31. Ibrahimi, K.; Khader, N.; Adouni, L. Soil Salinity Assessment and Characterization in Abandoned Farmlands of Metouia Oasis, South Tunisia. Environmental Sciences Proceedings 2022, Vol. 16, Page 3 2022, 16, 3, doi:10.3390/ENVIRONSCIPROC2022016003.

32. Todorovic, M. An Excel-Based Tool for Real-Time Irrigation Management At Field Scale. In Proceedings of the International Symposium on Water and Land Management for Sustainable Irrigated Agriculture"; 1992; pp. 1–11.

33. Allen, R.G.; Pereira, L.S.; Raes, D. Crop Evapotranspiration (Guidelines for Computing Crop Water Requirements). FAO: Rome, Italy. Available online: https://appgeodb.nancy.inra.fr/biljou/pdf/Allen_FAO1998.pdf (accessed on 17 September 2020).

34. Ayers, R.S.; Westcot, D.W. Water Quality for Agriculture. FAO: Rome, Italy, 1985. Available online: https://www.waterboards.ca.gov/water_issues/programs/tmdl/records/state_board/1985/ref2648.pdf (accessed on 18 September 2020).

35. Ismail, O.M. Use of Electrical Conductivity as a Tool for Determining Damage Index of Some Mango Cultivars. *Int. J. Plant Soil Sci.* **2014**, *3*, 448–456. https://doi.org/10.9734/ijpss/2014/8200.

36. Khan, N.M.; Rastoskuev, V.V.; Sato, Y.; Shiozawa, S. Assessment of Hydrosaline Land Degradation by Using a Simple Approach of Remote Sensing Indicators. *Agric. Water Manag.* **2005**, *77*, 96–109.

37. Abbas, M.A.; Khan, S. Using Remote Sensing Techniques for Appraisal of Irrigated Soil Salinity. In Proceedings of the International Congress on Modelling and Simulation (MODSIM), Christchurch, New Zealand, Modelling and Simulation Society of Australia and New Zealand: Brighton, UK, 2007; pp. 2632–2638.

38. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; de Colstoun, E.B.; McMurtrey, J.E. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. https://doi.org/10.1016/S0034-4257(00)00113-9.

39. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W.; Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with Erts. *NASSP* **1974**, *351*, 309.

40. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.

41. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2.

42. Saleska, S.R.; Didan, K.; Huete, A.R.; da Rocha, H.R. Amazon Forests Green-up during 2005 Drought. *Science* **2007**, *318*, 612. https://doi.org/10.1126/science.1146663.

43. Wu, W. The Generalized Difference Vegetation Index (GDVI) for Dryland Characterization. *Remote Sens.* **2014**, *6*, 1211–1233. https://doi.org/10.3390/rs6021211.

44. Wang, L.; Qu, J.J. NMDI: A Normalized Multi-Band Drought Index for Monitoring Soil and Vegetation Moisture with Satellite Remote Sensing. *Geophys. Res. Lett.* **2007**, *34*. https://doi.org/10.1029/2007GL031021.

45. Fensholt, R.; Sandholt, I. Derivation of a Shortwave Infrared Water Stress Index from MODIS Near- and Shortwave Infrared Data in a Semiarid Environment. *Remote Sens. Environ.* **2003**, *87*, 111–121. https://doi.org/10.1016/j.rse.2003.07.002.

46. Ceccato, P.; Flasse, S.; Tarantola, S.; Jacquemoud, S.; Grégoire, J.-M. Detecting Vegetation Leaf Water Content Using Reflectance in the Optical Domain. *Remote Sens. Environ.* **2001**, *77*, 22–33. https://doi.org/10.1016/S0034-4257(01)00191-2.

47. Hunt, E.R.; Rock, B.N. Detection of Changes in Leaf Water Content Using Near- and Middle-Infrared Reflectances. *Remote Sens. Environ.* **1989**, *30*, 43–54. https://doi.org/10.1016/0034-4257(89)90046-1.

48. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated Narrow-Band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. https://doi.org/10.1016/S0034-4257(02)00018-4.

49. Vincini, M.; Frazzi, E.; D'Alessio, P. A Broad-Band Leaf Chlorophyll Vegetation Index at the Canopy Scale. *Precis. Agric.* **2008**, *9*, 303–319. https://doi.org/10.1007/s11119-008-9075-z.

50. Hernandez-Clemente, R.; Navarro-Cerrillo, R.M.; Zarco-Tejada, P.J. Deriving Predictive Relationships of Carotenoid Content at the Canopy Level in a Conifer Forest Using Hyperspectral Imagery and Model Simulation. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5206–5217. https://doi.org/10.1109/TGRS.2013.2287304.

51. IDB—Index: TCARI/OSAVI. Available online: https://www.indexdatabase.de/db/i-single.php?id=191 (accessed on 29 January 2023).

52. Clevers, J.G.P.W.; Kooistra, L.; van den Brande, M.M.M. Using Sentinel-2 Data for Retrieving LAI and Leaf and Canopy Chlorophyll Content of a Potato Crop. *Remote Sens.* **2017**, *9*, 405. https://doi.org/10.3390/RS9050405.

53. Kamenova, I.; Dimitrov, P. Evaluation of Sentinel-2 Vegetation Indices for Prediction of LAI, FAPAR and FCover of Winter Wheat in Bulgaria. *Eur. J. Remote. Sens.* **2020**, *54*, 89–108. https://doi.org/10.1080/22797254.2020.1839359.

54. Zarco-Tejada, P.J.; Hornero, A.; Beck, P.S.A.; Kattenborn, T.; Kempeneers, P.; Hernández-Clemente, R. Chlorophyll Content Estimation in an Open-Canopy Conifer Forest with Sentinel-2A and Hyperspectral Imagery in the Context of Forest Decline. *Remote Sens. Environ.* **2019**, *223*, 320–335. https://doi.org/10.1016/j.rse.2019.01.031.

55. Huang, Y.; Qiu, B.; Chen, C.; Zhu, X.; Wu, W.; Jiang, F.; Lin, D.; Peng, Y. Automated Soybean Mapping Based on Canopy Water Content and Chlorophyll Content Using Sentinel-2 Images. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *109*, 102801. https://doi.org/10.1016/J.JAG.2022.102801.

56. IDB—Index: Modified Chlorophyll Absorption in Reflectance Index. Available online: https://www.indexdatabase.de/db/i-single.php?id=41 (accessed on 29 January 2023).

57. Maleki, M.; Arriga, N.; Barrios, J.M.; Wieneke, S.; Liu, Q.; Peñuelas, J.; Janssens, I.A.; Balzarolo, M. Estimation of Gross Primary Productivity (GPP) Phenology of a Short-Rotation Plantation Using Remotely Sensed Indices Derived from Sentinel-2 Images. *Remote Sens.* **2020**, *12*, 2104. https://doi.org/10.3390/RS12132104.

58. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the Capabilities of Sentinel-2 for Quantitative Estimation of Biophysical Variables in Vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92. https://doi.org/10.1016/J.ISPRSJPRS.2013.04.007.

59. Nguyen, N.; Binh, N.; Nguyen, B. Potential of Drought Monitoring Using Sentinel-2 Data; GIS-IDEAS, 2016. Available online: https://www.humg.edu.vn (accessed on 18 January 2021).

60. Wang, L.; Qu, J.J.; Hao, X. Forest Fire Detection Using the Normalized Multi-Band Drought Index (NMDI) with Satellite Measurements. *Agric. For. Meteorol.* **2008**, *148*, 1767–1776. https://doi.org/10.1016/J.AGRFORMET.2008.06.005.

61. Liu, Y.; Qian, J.; Yue, H. Combined Sentinel-1A with Sentinel-2A to Estimate Soil Moisture in Farmland. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1292–1310. https://doi.org/10.1109/JSTARS.2020.3043628.

62. Markov, B. Comparison of Remote Sensing-Based Indexes for Monitoring Drought Impact on Forest Ecosystems. Annual of Sofia University "St. Kliment Ohridski" Faculty of Geology and Geography Book 2. 2018, 11, 237-246. Available online: file:///C:/Users/MDPI/Downloads/15_Ann_Tom_111_geography_237-246%20(2).pdf (accessed on 19 August 2021).

63. Chen, Y.; Qiu, Y.; Zhang, Z.; Zhang, J.; Chen, C.; Han, J.; Liu, D. Estimating Salt Content of Vegetated Soil at Different Depths with Sentinel-2 Data. *PeerJ* **2020**, *8*, e10585. https://doi.org/10.7717/PEERJ.10585.

64. Meyer, L.H.; Heurich, M.; Beudert, B.; Premier, J.; Pflugmacher, D. Comparison of Landsat-8 and Sentinel-2 Data for Estimation of Leaf Area Index in Temperate Forests. *Remote Sens.* **2019**, *11*, 1160. https://doi.org/10.3390/RS11101160.

65. Hamrouni, Y.; Paillassa, E.; Chéret, V.; Monteil, C.; Sheeren, D. Sentinel-2 Poplar Index for Operational Mapping of Poplar Plantations over Large Areas. *Remote Sens.* **2022**, *14*, 3975. https://doi.org/10.3390/RS14163975.

66. Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K.-I. Crop Classification from Sentinel-2-Derived Vegetation Indices Using Ensemble Learning. *J. Appl. Remote Sens.* **2018**, *12*, 026019. https://doi.org/10.1117/1.JRS.12.026019.

67. Kobayashi, N.; Tani, H.; Wang, X.; Sonobe, R. Crop Classification Using Spectral Indices Derived from Sentinel-2A Imagery. *J. Inf. Telecommun.* **2019**, *4*, 67–90. https://doi.org/10.1080/24751839.2019.1694765.

68. IDB—Index: Normalized Difference 860/1640. Available online: https://www.indexdatabase.de/db/i-single.php?id=219 (accessed on 29 January 2023).

69. Elhag, M.; Bahrawi, J.A. Soil Salinity Mapping and Hydrological Drought Indices Assessment in Arid Environments Based on Remote Sensing Techniques. *Geosci. Instrum. Methods Data Syst.* **2017**, *6*, 149–158. https://doi.org/10.5194/GI-6-149-2017.

70. IDB—Index: Simple Ratio 1600/820 Moisture Stress Index. Available online: https://www.indexdatabase.de/db/i-single.php?id=48 (accessed on 29 January 2023).

71. Verrelst, J.; Romijn, E.; Kooistra, L. Mapping Vegetation Density in a Heterogeneous River Floodplain Ecosystem Using Pointable CHRIS/PROBA Data. *Remote Sens.* **2012**, *4*, 2866–2889. https://doi.org/10.3390/RS4092866.

72. Rivera, J.P.; Verrelst, J.; Delegido, J.; Veroustraete, F.; Moreno, J. On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization. *Remote Sens.* **2014**, *6*, 4927–4951. https://doi.org/10.3390/RS6064927.

73. Pablo, J.; Caicedo, R.; Verrelst, J.; Muñoz-Marí, J.; Moreno, J.; Camps-Valls, G. Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens*. **2014**, *7*, 1249–1259. https://doi.org/10.1109/JSTARS.2014.2298752.

74. Rivera, J.P.; Verrelst, J.; Leonenko, G.; Moreno, J. Multiple Cost Functions and Regularization Options for Improved Retrieval of Leaf Chlorophyll Content and LAI through Inversion of the PROSAIL Model. *Remote Sens.* **2013**, *5*, 3280–3304. https://doi.org/10.3390/RS5073280.

75. Metternicht, G.I.; Zinck, J.A. Remote Sensing of Soil Salinity: Potentials and Constraints. *Remote Sens. Environ.* **2003**, *85*, 1–20.

76. Allbed, A.; Kumar, L.; Aldakheel, Y.Y. Assessing Soil Salinity Using Soil Salinity and Vegetation Indices Derived from IKONOS High-Spatial Resolution Imageries: Applications in a Date Palm Dominated Region. *Geoderma* **2014**, *230–231*, 1–8. https://doi.org/10.1016/j.geoderma.2014.03.025.

77. Celleri, C.; Zapperi, G.; Trilla, G.G.; Pratolongo, P. Assessing the Capability of Broadband Indices Derived from Landsat 8 Operational Land Imager to Monitor above Ground Biomass and Salinity in Semiarid Saline Environments of the Bahía Blanca Estuary, Argentina. *Int. J. Remote Sens.* **2019**, *40*, 4817–4838. https://doi.org/10.1080/01431161.2019.1574992.

78. Allbed, A.; Kumar, L. Soil Salinity Mapping and Monitoring in Arid and Semi-Arid Regions Using Remote Sensing Technology: A Review. *Adv. Remote. Sens.* **2013**, *2*, 373–385. https://doi.org/10.4236/ars.2013.24040.

79. Pakparvar, M.; Gabriels, D.; Aarabi, K.; Edraki, M.; Raes, D.; Cornelis, W. Incorporating Legacy Soil Data to Minimize Errors in Salinity Change Detection: A Case Study of Darab Plain, Iran. *Int. J. Remote Sens.* **2012**, *33*, 6215–6238. https://doi.org/10.1080/01431161.2012.676688.

80. Verrelst, J.; Malenovský, Z.; van der Tol, C.; Camps-Valls, G.; Gastellu-Etchegorry, J.-P.; Lewis, P.; North, P.; Moreno, J. Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surv. Geophys.* **2019**, *40*, 589–629. https://doi.org/10.1007/s10712-018-9478-y.

81. Verrelst, J.; Camps-Valls, G.; Muñoz-Marí, J.; Rivera, J.P.; Veroustraete, F.; Clevers, J.G.P.W.; Moreno, J. Optical Remote Sensing and the Retrieval of Terrestrial Vegetation Bio-Geophysical Properties—A Review. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 273–290. https://doi.org/10.1016/j.isprsjprs.2015.05.005.

82. Pignatti, S.; Acito, N.; Amato, U.; Casa, R.; Castaldi, F.; Coluzzi, R.; de Bonis, R.; Diani, M.; Imbrenda, V.; Laneve, G.; et al. Environmental Products Overview of the Italian Hyperspectral Prisma Mission: The SAP4PRISMA Project. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA; pp. 3997–4000.

83. Guanter, L.; Kaufmann, H.; Segl, K.; Foerster, S.; Rogass, C.; Chabrillat, S.; Kuester, T.; Hollstein, A.; Rossner, G.; Chlebek, C.; et al. The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. *Remote Sens.* **2015**, *7*, 8830–8857. https://doi.org/10.3390/rs70708830.

84. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine Learning Regression Algorithms for Biophysical Parameter Retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* **2012**, *118*, 127–139. https://doi.org/10.1016/J.RSE.2011.11.002.

85. Knudby, A.; LeDrew, E.; Brenning, A. Predictive Mapping of Reef Fish Species Richness, Diversity and Biomass in Zanzibar Using IKONOS Imagery and Machine-Learning Techniques. *Remote Sens. Environ.* **2010**, *114*, 1230–1241. https://doi.org/10.1016/J.RSE.2010.01.007.

86. Jiapaer, G.; Liang, S.; Yi, Q.; Liu, J. Vegetation Dynamics and Responses to Recent Climate Change in Xinjiang Using Leaf Area Index as an Indicator. *Ecol. Indic.* **2015**, *58*, 64–76. https://doi.org/10.1016/J.ECOLIND.2015.05.036.

87. Houborg, R.; McCabe, M.F. Daily Retrieval of NDVI and LAI at 3 m Resolution via the Fusion of CubeSat, Landsat, and MODIS Data. *Remote Sens.* **2018**, *10*, 890. https://doi.org/10.3390/RS10060890.

88. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.

89. Aghababaei, M.; Ebrahimi, A.; Naghipour, A.A.; Asadi, E.; Pérez-Suay, A.; Morata, M.; Garcia, J.L.; Caicedo, J.P.R.; Verrelst, J. Introducing ARTMO’s Machine-Learning Classification Algorithms Toolbox: Application to Plant-Type Detection in a Semi-Steppe Iranian Landscape. *Remote Sens.* **2022**, *14*, 4452. https://doi.org/10.3390/RS14184452.

90. Caballero, G.; Pezzola, A.; Winschel, C.; Casella, A.; Angonova, P.S.; Rivera-Caicedo, J.P.; Berger, K.; Verrelst, J.; Delegido, J. Seasonal Mapping of Irrigated Winter Wheat Traits in Argentina with a Hybrid Retrieval Workflow Using Sentinel-2 Imagery. *Remote Sens.* **2022**, *14*, 4531. https://doi.org/10.3390/RS14184531.

91. Belda, S.; Pipia, L.; Morcillo-Pallarés, P.; Rivera-Caicedo, J.P.; Amin, E.; de Grave, C.; Verrelst, J. DATimeS: A Machine Learning Time Series GUI Toolbox for Gap-Filling and Vegetation Phenology Trends Detection. *Environ. Model. Softw.* **2020**, *127*, 104666. https://doi.org/10.1016/J.ENVSOFT.2020.104666.

92. Pipia, L.; Muñoz-Marí, J.; Amin, E.; Belda, S.; Camps-Valls, G.; Verrelst, J. Fusing Optical and SAR Time Series for LAI Gap Filling with Multioutput Gaussian Processes. *Remote Sens. Environ.* **2019**, *235*, 111452. https://doi.org/10.1016/J.RSE.2019.111452.

93. Nguyen-Sy, T.; To, Q.-D.; Vu, M.-N.; Nguyen, T.-D.; Nguyen, T.-T. Predicting the Electrical Conductivity of Brine-Saturated Rocks Using Machine Learning Methods. *J. Appl. Geophys.* **2021**, *184*, 104238. https://doi.org/10.1016/J.JAPPGEO.2020.104238.

94. Feng, Y.; Liu, L.; Shu, J. A Link Quality Prediction Method for Wireless Sensor Networks Based on Xgboost. *IEEE Access* **2019**, *7*, 155229–155241. https://doi.org/10.1109/ACCESS.2019.2949612.

95. Yan, S.; Wu, L.; Fan, J.; Zhang, F.; Zou, Y.; Wu, Y. A Novel Hybrid WOA-XGB Model for Estimating Daily Reference Evapotranspiration Using Local and External Meteorological Data: Applications in Arid and Humid Regions of China. *Agric. Water Manag.* **2021**, *244*, 106594. https://doi.org/10.1016/J.AGWAT.2020.106594.

96. Liu, P.; Wang, J.; Sangaiah, A.K.; Xie, Y.; Yin, X. Analysis and Prediction of Water Quality Using LSTM Deep Neural Networks in IoT Environment. *Sustainability* **2019**, *11*, 2058. https://doi.org/10.3390/SU11072058.

97. Chen, K.; Chen, H.; Zhou, C.; Huang, Y.; Qi, X.; Shen, R.; Liu, F.; Zuo, M.; Zou, X.; Wang, J.; et al. Comparative Analysis of Surface Water Quality Prediction Performance and Identification of Key Water Parameters Using Different Machine Learning Models Based on Big Data. *Water Res.* **2020**, *171*, 115454. https://doi.org/10.1016/J.WATRES.2019.115454.

98. Yaseen, Z.M. An Insight into Machine Learning Models Era in Simulating Soil, Water Bodies and Adsorption Heavy Metals: Review, Challenges and Solutions. *Chemosphere* **2021**, *277*, 130126. https://doi.org/10.1016/J.CHEMOSPHERE.2021.130126.

99. Gharaibeh, M.A.; Albalasmeh, A.A.; Pratt, C.; el Hanandeh, A. Estimation of Exchangeable Sodium Percentage from Sodium Adsorption Ratio of Salt-Affected Soils Using Traditional and Dilution Extracts, Saturation Percentage, Electrical Conductivity, and Generalized Regression Neural Networks. *Catena* **2021**, *205*, 105466. https://doi.org/10.1016/J.CATENA.2021.105466.

100. el Bilali, A.; Taleb, A.; Brouziyne, Y. Groundwater Quality Forecasting Using Machine Learning Algorithms for Irrigation Purposes. *Agric. Water Manag.* **2021**, *245*, 106625. https://doi.org/10.1016/J.AGWAT.2020.106625.

101. Loh, W.-Y. Classification and Regression Trees. *WIREs Data Min. Knowl. Discov.* **2011**, *1*, 14–23. https://doi.org/10.1002/WIDM.8.

102. Kisi, O.; Dailr, A.H.; Cimen, M.; Shiri, J. Suspended Sediment Modeling Using Genetic Programming and Soft Computing Techniques. *J. Hydrol.* **2012**, *450–451*, 48–58. https://doi.org/10.1016/J.JHYDROL.2012.05.031.

103. Choubin, B.; Darabi, H.; Rahmati, O.; Sajedi-Hosseini, F.; Kløve, B. River Suspended Sediment Modelling Using the CART Model: A Comparative Study of Machine Learning Techniques. *Sci. Total Environ.* **2018**, *615*, 272–281. https://doi.org/10.1016/J.SCITOTENV.2017.09.293.

104. Melesse, A.M.; Khosravi, K.; Tiefenbacher, J.P.; Heddam, S.; Kim, S.; Mosavi, A.; Pham, B.T. River Water Salinity Prediction Using Hybrid Machine Learning Models. *Water* **2020**, *12*, 2951. https://doi.org/10.3390/w12102951.

105. Johns, K.A.; Emslie, M.J.; Hoey, A.S.; Osborne, K.; Jonker, M.J.; Cheal, A.J. Macroalgal Feedbacks and Substrate Properties Maintain a Coral Reef Regime Shift. *Ecosphere* **2018**, *9*, e02349. https://doi.org/10.1002/ECS2.2349.

106. Wu, W.; Zucca, C.; Muhaimeed, A.S.; Al-Shafie, W.M.; Al-Quraishi, A.M.F.; Nangia, V.; Zhu, M.; Liu, G. Soil Salinity Prediction and Mapping by Machine Learning Regression in Central Mesopotamia, Iraq. *Land Degrad. Dev.* **2018**, *29*, 4005–4014. https://doi.org/10.1002/ldr.3148.

107. Ghorbani, M.A.; Aalami, M.T.; Naghipour, L. Use of Artificial Neural Networks for Electrical Conductivity Modeling in Asi River. *Appl. Water Sci.* **2017**, *7*, 1761–1772. https://doi.org/10.1007/S13201-015-0349-0.

108. Khozani, Z.S.; Khosravi, K.; Pham, B.T.; Kløve, B.; Mohtar, W.H.M.W.; Yaseen, Z.M. Determination of Compound Channel Apparent Shear Stress: Application of Novel Data Mining Models. *J. Hydroinfor.* **2019**, *21*, 798–811. https://doi.org/10.2166/HYDRO.2019.037.

109. Pham, B.T.; Prakash, I.; Khosravi, K.; Chapi, K.; Trinh, P.T.; Ngo, T.Q.; Hosseini, S.V.; Bui, D.T. A Comparison of Support Vector Machines and Bayesian Algorithms for Landslide Susceptibility Modelling. *Geocarto Int.* **2018**, *34*, 1385–1407. https://doi.org/10.1080/10106049.2018.1489422.

110. Khosravi, K.; Mao, L.; Kisi, O.; Yaseen, Z.M.; Shahid, S. Quantifying Hourly Suspended Sediment Load Using Data Mining Models: Case Study of a Glacierized Andean Catchment in Chile. *J. Hydrol.* **2018**, *567*, 165–179. https://doi.org/10.1016/J.JHYDROL.2018.10.015.