



# Article Soil Salinity Mapping of Plowed Agriculture Lands Combining Radar Sentinel-1 and Optical Sentinel-2 with Topographic Data in Machine Learning Models

Diego Tola <sup>1,2,3,\*</sup>, Frédéric Satgé <sup>3,4</sup>, Ramiro Pillco Zolá <sup>4</sup>, Humberto Sainz <sup>2</sup>, Bruno Condori <sup>2</sup>, Roberto Miranda <sup>5</sup>, Elizabeth Yujra <sup>5</sup>, Jorge Molina-Carpio <sup>4</sup>, Renaud Hostache <sup>3</sup> and Raúl Espinoza-Villar <sup>1</sup>

- <sup>1</sup> Programa de Doctorado en Recursos Hídricos (PDRH), Universidad Nacional Agraria La Molina, Lima 15024, Peru; respinoza@lamolina.edu.pe
- <sup>2</sup> Área de Ciencias Agrícolas, Pecuarias y Recursos Naturales (ACAPRN), Universidad Pública de El Alto, La Paz 10077, Bolivia; humbertosainz@gmail.com (H.S.); bcondori@yahoo.com (B.C.)
- <sup>3</sup> ESPACE-DEV, University Montpellier, IRD, University Antilles, University Guyane, University Réunion, 34093 Montpellier, France; frederic.satge@ird.fr (F.S.); renaud.hostache@ird.fr (R.H.)
- <sup>4</sup> Instituto de Hidráulica e Hidrología (IHH), Universidad Mayor de San Andrés, La Paz 10077, Bolivia; rami\_lund99@hotmail.com (R.P.Z.); jamolinacarpio@gmail.com (J.M.-C.)
- Facultad de Agronomía, Universidad Mayor de San Andrés, La Paz 10077, Bolivia;
- robertomicasa@gmail.com (R.M.); eytelizabethyujra@gmail.com (E.Y.)
- \* Correspondence: 20211517@lamolina.edu.pe

**Abstract:** This study assesses the relative performance of Sentinel-1 and -2 and their combination with topographic information for plow agricultural land soil salinity mapping. A learning database made of 255 soil samples' electrical conductivity (EC) along with corresponding radar (R), optical (O), and topographic (T) information derived from Sentinel-2 (S2), Sentinel-1 (S1), and the SRTM digital elevation model, respectively, was used to train four machine learning models (Decision tree—DT, Random Forest—RF, Gradient Boosting—GB, Extreme Gradient Boosting—XGB). Each model was separately trained/validated for four scenarios based on four combinations of R, O, and T (R, O, R+O, R+O+T), with and without feature selection. The Recursive Feature Elimination with k-fold cross validation (RFEcv 10-fold) and the Variance Inflation Factor (VIF) were used for the feature selection process to minimize multicollinearity by selecting the most relevant features. The most reliable salinity estimates are obtained for DT, GB, RF, and XGB, respectively. Conversely, models based on R information led to unreliable soil salinity estimates due to the saturation of the C-band signal in plowed lands.

Keywords: soil salinity mapping; plowed lands; machine learning; Sentinel-1; Sentinel-2

## 1. Introduction

## 1.1. Soil Salinity: An Agriculture Threat

Soil salinization is a problem that affects agriculture worldwide [1], especially soil in dry sub-humid, semi-arid, and arid regions (taking into account the increasing degree of aridity) where salinity has reached alarming levels [2]. The presence of salts leads to a degradation of the soil's physical properties by the dispersion of mineral colloids, as well as generating an osmotic potential in the soil, limiting the easy exchange of water and nutrients with the roots of the plants, which in turn results in a delay in the growth of the plants [3]. This process induces negative effects on crop productivity and quality [4], which can generate significant economic losses [5]. In the current context of a likely increase in temperature and decrease in precipitation in the south-central highlands [6], an increase in irrigation volumes is necessary to meet crop water requirements. However, irrigation without proper leaching and drainage leaves salt precipitates in the soil profile, while



Citation: Tola, D.; Satgé, F.; Pillco Zolá, R.; Sainz, H.; Condori, B.; Miranda, R.; Yujra, E.; Molina-Carpio, J.; Hostache, R.; Espinoza-Villar, R. Soil Salinity Mapping of Plowed Agriculture Lands Combining Radar Sentinel-1 and Optical Sentinel-2 with Topographic Data in Machine Learning Models. *Remote Sens.* **2024**, *16*, 3456. https://doi.org/10.3390/ rs16183456

Academic Editor: Nicolas Baghdadi

Received: 13 July 2024 Revised: 6 September 2024 Accepted: 9 September 2024 Published: 18 September 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). over-irrigation favors the gradual rise of the groundwater level, which in turn contributes to soil salt accumulation [7]. In this context, it is estimated that 50% of global cropland areas will be affected by salinization by 2050 [8], thus representing a major socioeconomic threat that can lead to population outmigration [4,9].

#### 1.2. Soil Salinity Monitoring: Traditional vs. Remote Sensing Techniques

Conventional studies mapping soil properties such as salinity have the advantage of allowing a comprehensible representation of complex soil formation processes through conceptual models [10] and are not restricted by the quantity of soil types that can be mapped with a single conceptual model [11]. However, they require field campaigns that can be (i) slow, (ii) expensive, and (iii) limited in space and time. In this context, salinity data cannot be provided in a timely manner or at the agriculture plot scale to make decisions about planting and/or to understand the influence of agricultural practices on salinity variability [12,13]. To overcome these limitations, digital soil mapping based on remote sensing data emerges as a promising technique to update soil information in a cost-effective manner and with a regional perspective [14,15]. More specifically, soil salinity mapping has been substantially improved across the last decade through the integration of remote sensing data into machine learning (ML) models [2,16]. Freely available, with a global coverage and a high frequency revisit time (5–10 days), Sentinel-2 (S2) optical and/or Sentinel-1 (S1) radar images are generally used as explanatory features in the ML models.

Recent studies have made it possible to map soil salinity on a global scale by combining soil information and remote sensing, which can then be used to estimate the area affected by salinity worldwide, covering about one billion hectares [1,17]. Many studies carried out in China [18], Algeria [19], Iran [20], and Bolivia [21] have reported on the potential of Sentinel-2 optical satellites images (S2) to retrieve soil salinity estimates with a relatively good degree of confidence. However, these techniques are not suitable for regions or time periods affected by clouds. In the presence of clouds, unlike radar sensors, optical sensors (such as S2) cannot observe the ground surface or, therefore, soil salinity content. In this context, several studies carried out in Iran [22], Vietnam [23], India [24,25], and Bolivia [26] reported a good degree of confidence for soil salinity estimates based solely on Sentinel-1 radar images (S1), whereas some authors took advantages of the S1 and S2 complementarity to improve soil salinity estimates [27]. Despite the significant advances in soil salinity mapping brought by the above-mentioned studies, those works are limited to undisturbed natural soils. However, agriculture is developed on plowed soils with different roughness and compaction features than the ones of natural undisturbed soils. As roughness and compaction features affect S2 reflectance and S1 polarization [28–30], soil salinity models calibrated across natural unperturbed land cannot be applied to plow soil. In this context, some authors focused on agriculture plots to assess the potential of S2 alone [31–33] and in combination with S1 [34,35] to estimate soil salinity in cultivated land.

#### 1.3. Which Machine-Learning Model Performs the Best?

Independently of the type of soil (i.e., natural or agricultural), the choice of the machine learning model could affect the reliability of the soil salinity estimates. In natural undisturbed soils, the sensitivity of soil salinity estimates to different machine learning models, such as Artificial Neural Network (ANN), decision tree (DT), Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), Classification and Regression Trees (CART), Partial Least Squares Regression (PLSR), Multi Linear Regression (MLR), and k-Nearest Neighbor (kNN) models, has been extensively reported [18,19,36–39]. In most cases, more reliable soil salinity estimates are obtained when considering the RF or SVM models [21]. For agricultural lands, only one study has reported on the sensitivity of soil salinity estimates to different machine learning models [35]. Carried out in Egypt, this study compared the soil estimates obtained with three machine learning models (i.e., RF, SVM, and Back Propagation Neural Network (BPNN)). The results show that the most reliable soil salinity estimates are obtained with the RF model, as this model is more suitable for mapping heterogeneous regions, resulting in a more accurate representation of the spatial distribution of areas affected by salinity.

## 1.4. Study Objectives

Based on the previously established state of the art, this study aims to assess the potential of S1 and S2 images (i.e., separately and as a whole) for soil salinity estimates in plowed areas. Several machine learning models (i.e., RF, DT, Gradient Boosting, and XGBoost) are considered to expand the current limited knowledge on the sensitivity of agricultural soil salinity estimates to different machine learning models. With 900 km<sup>2</sup> of its area at risk of salinization [40], the Bolivian Altiplano is selected as the study site.

### 2. Materials

## 2.1. Study Area

The study was carried out in the central and southern Altiplano of Bolivia (Figure 1). The region is part of the TDPS endorheic system (Titicaca, Desaguadero, Poopó, Salar) located between the eastern and western Andean Cordilleras at a mean elevation of 4000 m m.a.s.l. with a very flat topography (i.e., mean slope of 4.7°) [41,42]. The region is semi-arid with (i) a mean precipitation of less than 600 mm  $\cdot$  year<sup>-1</sup> concentrated in the austral summer [43], (ii) a mean temperature ranging from 4 °C to 8 °C [44], and (iii) a high evapotranspiration rate estimated at 1700 mm·year<sup>-1</sup> around the Poopó Lake [45,46]. Agriculture is a main source of economic income that increasingly relies on irrigation practices to cope with particularly adverse weather conditions [47]. The study area encompasses three main types of soil: Haplic Xerosols (Xh-northern part), Eutric Fluvisols (Je-surroundings of Poopó Lake), and Haplic Yermosols (Yh-southern part) [18]. According to the guidelines of the WRB System (World Reference Base for Soil), their approximate equivalents could be placed as Xerosols = Association of Leptosols – Durisols – Regosols, and Yermosols = Association of Regosols – Solonetz – Solonchaks [48]. The combination of a semi-arid climatic context and irrigation in this endorheic system favors the accumulation of salts on the soil surface, resulting in extreme saline soil conditions [7,21].

## 2.2. Reference Soil Salinity Data

A total of 255 soil samples (their location based on accessibility and the presence of plowed plots) were collected during the pre-sowing period (April to October), which correspond to the dry period with almost no precipitation events that could affect soil salinity (Figure 1c–e). The sample collections were made during 5 field campaigns lead in 2022 and 2023 (Table 1).

Each soil sample corresponds to a composite sample made of 5 subsamples taken at an average depth of 10 cm at the center and four corners of each sampling site (approximately  $10 \text{ m} \times 10 \text{ m}$  in size) to take into account the potential soil salinity heterogeneity that may occur at the Sentinel images' spatial resolution ( $10 \text{ m} \times 10 \text{ m}$ ).

The soil electrical conductivity (EC) of soil samples was measured in the soil laboratory of the Faculty of Agronomy of the Universidad Mayor de San Andrés (UMSA), as a proxy of soil salinity. To prepare the samples, they were first air-dried, pulverized, homogenized, and then sieved with a 2 mm mesh. The method used to measure the soil EC involved creating a suspension of soil and water using a 1:5 ratio method [49]. In this process, a soil sample weighing  $5 \pm 0.1$  g was placed in a polypropylene centrifuge tube and  $25 \pm 0.5$  mL of deionized water was added. The samples were then agitated on a stirring table for 60 min and allowed to decant for an additional 30 min. The EC ( $25 \degree$ C) was then measured in the supernatant without disturbing the sediment layer, using the Potentiometric Method, with a calibrated multi-parameter device fitted with a conductivity electrode.



**Figure 1.** Location of the study area in Bolivia (**a**) with the location of the 255 soil sample sites at the regional scale (**b**) and within agricultural plots (**c**–**e**). Samples' EC measured in laboratory with FAO soil types (**f**) along with the sample distribution among the different soil types (**g**) and EC classes (**h**).

<b>Fable 1.</b> Soil sampling da	tes and image ac	equisition dates	S1 and S2
----------------------------------	------------------	------------------	-----------

Mission	Date	Number of Samples	Date S1	Date S2
1	17 to 19 July 2022	17	18, 30 July 2022	17, 22, 27 July 2022
2	10 to 12 August 2022	18	11, 23 August 2022	11, 16, 21 August 2022
3	7 to 9 April 2023	57	8, 20 April 2023	8, 13, 18 April 2023
4	10 to 12 July 2023	79	13, 25 July 2023	12, 17, 22 July 2023
5	2 to 4 October 2023	84	5, 17 October 2023	5, 10, 15 October 2023
	Total:	255		

The collected samples correspond to Haplic Yermosol (n = 11), Haplic Xerosol (n = 109) and Eutric Fluvisol (n = 135) (Figure 1f,g). Considering the FAO soil salinity classification [1], most of the soil samples analyzed do not present high levels of salinity (Figure 1f). Experienced farmers favor sowing in low salinity areas rather than high salinity areas to guarantee decent crop production. Soil samples with high salinity are mainly located in Eutric Fluvisol soils because of the presence of a nearby water table [40,50]. Indeed,

regional groundwater is more saline in nature due to factors such as the dissolution of sediments from ancient Quaternary lakes, the recharge of the Desaguadero River, natural contamination due to the erosion of rocks with high concentrations of salts, and contamination due to centuries of mining activity [51,52]. Consequently, groundwater becomes more saline near Poopó Lake and Desaguadero River [52], as identified by the collected samples (Figure 1f).

## 2.3. Sentinel-2 Images and Pre-Processing

Sentinel-2 images (S2) provide reflectance in 12 bands with wavelengths ranging from the visible spectrum to the infrared region. For this study, B1 and B9 bands were not considered due to their low spatial resolution (60 m) and because their applications are mainly associated with water vapor and atmospheric correction (aerosols and water vapor).

S2 images are available prior (level 1C) and after (level 2A) the application of an atmospheric correction. To ensure time and space consistency in S2 reflectance values, the atmospherically corrected level 2A images are used. The "QA60" quality band was used to identify the pixels affected by cloud cover. All pixels identified as "cloudy" were systematically flagged as missing values. To reduce the effect of noise that could occur locally and obtain an average condition of the sampled site, composite images were elaborated for each field campaign [21,53]. The composite images are obtained considering the mean value of the three S2 images obtained during and just after the sampling date (Table 1). Images before the sampling date were not considered to avoid using images recorded when the agricultural plot may not yet have been plowed. S2 pre-processing (i.e., download and composite) was carried out using GEE [54] in the free Google Colab cloud platform (retrieved from https://colab.research.google.com/, accessed on 12 June 2023).

#### 2.4. Sentinel-1 Images and Pre-Processing

The Sentinel-1 images (S1) provide polarization values in VV (Vertical/Vertical) and VH (Vertical/Horizontal) in interferometry mode with a central frequency of 5.405 GHz. S1 images are available as a Single Look Complex product (SLC), providing both phase and amplitude information, and as the post-processed Ground Range Detected product (GRD), providing direct surface characteristics in the form of intensity images. Due to its dual polarization antenna that could discriminate subtle changes in soil salinity content, the GRD product has become a valuable resource for mapping soils affected by salinity and therefore is being considered for the present study [55]. S1 images are available in both ascending and descending orbit. To ensure consistency in S1 observations in space and time, only the descending orbit is used.

S1 images were pre-processed according to four successive steps: (i) noise removal, (ii) radiometric calibration, (iii) topographic correction, and (iv) focal mean filter to reduce speckle effect. As for S2 images, composite images made of the two closest S1 post samples are considered to reduce the effect of noise that could occur locally (Table 1). S1 pre-processing (i.e., download, pre-process, and composite) was carried out using a similar way as in S2, via GEE [54] in the free Google Colab cloud platform (retrieved from https://colab.research.google.com/, accessed on 12 June 2023).

## 2.5. Digital Elevation Model

The Shuttle Radar Topographic Mission Digital Elevation Model (SRTM-DEM) was used in this study because it was previously identified as the most reliable DEM for the Altiplano region [41]. Data to produce the SRTM-DEM were collected over 11 days in February 2000 using dual spaceborne imaging radar (SIR-C) and dual X-band synthetic aperture radar (X-SAR). The SRTM-DEM has a spatial resolution at about 30 m and spans 60°N–56°S.

## 2.6. Machine Learning Models

Four supervised machine learning (ML) models based on decision trees are considered: (i) Random Forest (RF) [56], (ii) Extreme Gradient Boosting (XGB) [57], (iii) decision tree (DT) [58], and (iv) Gradient Boosting (GB) [59]. Decision trees are non-parametric supervised predictive models used for both regression and classification problems. They represent decisions and their possible consequences in a hierarchical manner similar to the structure of a tree. In this way, simple decision rules are inferred from the characteristics of the dataset to create a model that predicts the value of a target variable (i.e., soil salinity) [60].

In the DT model, a single decision tree is constructed by recursively dividing the dataset into smaller subsets based on the features that provide the greatest information gain or impurity reduction [58,60]. In the RF model, multiple decision trees are built using different subsets of the learning database (a technique called bagging). The final prediction corresponds to the average of the predictions of all trees (ensemble) [56]. In the GB model, decision trees are sequentially built so that the new tree attempts to correct the previous tree's errors by minimizing a specific loss function. In this process, each tree is built on the residuals (errors) of the previous tree [59]. Finally, the XGB model is an optimized version of GB, which applies parallel computing techniques to speed up the training and regularization process to avoid overfitting [59].

The DT, RF, and GB models were implemented using the scikit-learn library [61], while XGB was implemented using the xgboost library, which is compatible with scikit-learn functions [57]. The default hyper-parameters assigned by the scikit-learn library in Python were considered for all models (100 decision trees excluding DT (n\_estimator), and considering all features for each decision tree (max\_features)). It is worth mentioning that these hyper-parameters provided good results in ML models [61].

## 3. Methods

## 3.1. Elaboration of the Learning Database

The datasets (i.e., S1, S2, and SRTM) were homogenized to a 10 m spatial resolution using the nearest-neighbor method. Then, S1 polarization (VV and VH), S2 spectral reflectance (B2, B3, B4, B5, B6, B7, B8, B8a, B11 and B12), and SRTM-DEM elevation values were extracted from the pixels where a soil sample was collected. At this stage, the learning database includes 255 EC observations with corresponding S1 polarization and S2 spectral reflectance values and SRTM-DEM elevation. Environmental features, including 24 spectral indices, 5 soil texture indices, and 12 topographic indices, were computed from S2 reflectance, S1 polarization, and SRTM-DEM elevation, respectively, and added to the learning database (Tables 2–4). The topographic indices were computed using the "tagee" library [62]. Those indices were selected as they have shown strong correlation with soil salinity [1,14,25,63–65].

Ν	Index/Acronym	Formula * or Description	Reference
	Sentinel-2 bands	B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12	
1	Normalized Differential Vegetation Index (NDVI)	$NDVI = \frac{NIR-Red}{NIR+Red}$	[66]
2	Soil-Adjusted VegetationIndex (SAVI)	$\mathrm{SAVI} = rac{\mathrm{NIR-Red}}{\mathrm{NIR+Red}+0.5}  imes 1.5$	[67]
3	Enhanced Vegetation Index (EVI)	$\text{EVI} = 2.5  imes rac{\text{NIR}-\text{Red}}{\text{NIR}+6  imes \text{Red}-7.5  imes \text{Blue}+1}$	[68]
4	Normalized Difference Moisture Index (NDMI)	$NDMI = \frac{NIR - SWIR1}{NIR + SWIR1}$	[69]
5	Moisture Stress Index (MSI1)	$MSI1 = \frac{SWIR1}{NIR}$	[70]
6	Moisture Stress Index (MSI2)	$MSI2 = \frac{SWIR1 - NIR}{SWIR1 + NIR} \times \frac{SWIR1 - Red}{SWIR1 + Red}$	Modified from [69]
7	Bare Soil Index (BSI)	$\mathrm{BSI} = \frac{(\mathrm{Red} + \mathrm{SWIR1}) - (\mathrm{NIR} + \mathrm{Blue})}{(\mathrm{Red} + \mathrm{SWIR1}) + (\mathrm{NIR} + \mathrm{Blue})}$	[71]

Table 2. Sentinel-2 features.

N	Index/Acronym	Formula * or Description	Reference
8	Normalized Burn Ratio (NBR)	$NBR = \frac{NIR - SWIR2}{NIR + SWIR2}$	[72]
9	Normalized Burn Ratio 2 (NBR2)	$NBR2 = \frac{SWIR1 - SWIR2}{SWIR1 + SWIR2}$	[73]
10	Normalized Difference Water Index (NDWI)	$NDWI = \frac{NIR-Green}{NIR+Green}$	[74]
11	Normalized Salinity Index (NSI1)	$NSI1 = \frac{SWIR1 - SWIR2}{SWIR1 + NIR}$	[1]
12	Modified Normalized Salinity Index (NSI2)	$NSI2 = \frac{SWIR1 - SWIR2}{SWIR1 - NIR}$	Modified from [1]
13	Salinity Index 1 (SI1)	$SI1 = \sqrt{\text{Red} \times \text{Green}}$	[75]
14	Salinity Index 2 (SI2)	$SI2 = \sqrt{Blue \times Red}$	[14]
15	Salinity Index 3 (SI3)	$SI3 = \sqrt{\text{Red}^2 \times \text{Green}^2}$	[14]
16	Salinity Index 4 (SI4)	$SI4 = \frac{NIR \times SWIR1 - SWIR1^2}{NIR}$	[1]
17	Salinity Index 5 (SI5)	$SI5 = \frac{Blue}{Red}$	[14]
18	Salinity Index 6 (SI6)	$SI6 = \frac{Red \times NIR}{Green}$	[14]
19	Vegetation Soil Salinity Index (VSSI)	$VSSI = 2 \times Green - 5 \times (Red + NIR)$	[76]
20	Normalized Difference Salinity Index (NSI)	$NSI = rac{Red - NIR}{Red + NIR}$	[77]
21	Salinity Ratio (SR)	$SR = \frac{Green-Red}{Blue+Red}$	[76]
22	Canopy Response Salinity Index (CRSI)	$CRSI = \sqrt{\frac{NIR \times Red - Green \times Blue}{NIR \times Red + Green \times Blue}}$	[1]
23	Brightness Index (BI1)	$BI1 = \sqrt{Red^2 + NIR^2}$	[78]
24	Brightness Index (BI2)	$BI2 = \sqrt{Green^2 + Red^2 + NIR^2}$	[75]

## Table 2. Cont.

\* Where Blue = B2; Green = B3; Red = B4; NIR = B8; SWIR1 = B11; SWIR2 = B12.

## Table 3. Sentinel-1 features.

Ν	Index/Acronym	Formula * or Description	Reference
	Sentinel-1 polarization (dB)	VV, VH	
1	Soil texture index 1 (IT1)	IT1 = VV + VH	
2	Soil texture index 1 (IT2)	$IT2 = VV^2 + VH$	-
3	Soil texture index 1 (IT3)	$IT3 = VH^2 - VV$	[25]
4	Soil texture index 1 (IT4)	$IT4 = VV^2 + VH^2$	-
5	Soil texture index 1 (IT5)	$IT5 = \frac{VV^2 + VH^2}{VH}$	-

\* Where VV = Vertical transmission/Vertical reception; VH = Vertical transmission/horizontal reception.

## Table 4. Topographic features.

Ν	Index/Acronym	Formula *	Reference
	DEM	$\mathbf{z} = \mathbf{f}(\mathbf{x}, \mathbf{y})$	[79]
1	Slope	$G = \arctan \sqrt{p^2 + q^2}$	
2	Aspect	$\begin{split} A &= -90[1 - \text{sign}(q)](1 -  \text{sign}(p) ) + 180[1 + \text{sign}(p)] \\ &- \frac{180}{\pi} \text{sign}(p) \text{arcos}\left(\frac{-q}{\sqrt{p^2 + q^2}}\right) \\ &\text{sign}(x) = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x = 0 \\ -1 & \text{for } x < 0 \end{cases} \end{split}$	[62]
3	Hillshade	$H_L = 255(cos(Ze) \times cos(G) + sin(Ze) \times sin(G) \times cos(Az - A)$	

N	Index/Acronym	Formula *	Reference
4	Northness	$A_N = \cos A$	
5	Eastness	$A_E = sinA$	
6	Horizontal Curvature	$k_{h} = \frac{q^{2}r - 2pqs + p^{2}t}{(p^{2} + q^{2})\sqrt{1 + p^{2} + q^{2}}}$	
7	Vertical Curvature	$k_v = \frac{p^2r - 2pqs + q^2t}{(p^2 + q^2)\sqrt{\left(1 + p^2 + q^2\right)^3}}$	
8	Mean Curvature	${ m H}=rac{(1\!+\!q^2)r\!-\!2pqs\!+\!(1\!+\!p^2)t}{2\sqrt{(1\!+\!p^2\!+\!q^2)^3}}$	[62]
9	Gaussian Curvature	$\mathrm{K}=\frac{\mathrm{rt}\mathrm{-s}^2}{(\mathrm{1}\mathrm{+p}^2\mathrm{+q}^2)^2}$	
10	Minimal Curvature	$K_{min} = H - \sqrt{\left(H^2 - K\right)}$	
11	Maximal Curvature	$K_{max} = H + \sqrt{\left(H^2 - K\right)}$	
12	Shape Index	$SI = \frac{2}{\pi} \arctan \frac{H}{(H^2 - K)}$	

#### Table 4. Cont.

\* Where z = elevation (meters); Ze = Zenit (rad); Az = Azimuth (rad); r, t, s, p y q are the partial derivatives obtained by operating on linear distances calculated from the Haversine formula [62].

#### 3.2. Machine Learning Modelling Set-Up

To identify the potential of S1 and S2 features for soil salinity prediction, S1 and S2 features are considered explanatory features (i) separately to assess their respective potential and (ii) together to assess their complementarity. Finally, S1, S2, and topographic features are assessed together to highlight the potential benefit brought by topographic features. In this way, four scenarios are considered for each of the four ML models (i.e., RF, GT, GB, and XGB) (Figure 2).

Multi-collinearity is a well-known problem occurring in ML to decrease models' robustness due to redundancy in selected independent features. To minimize such effects, the Recursive Feature Elimination Cross Validation algorithm (RFEcv) is used to find the subset of independent features with the best score [80]. As a model-specific feature selection method (i.e., wrapping method), the RFEcv is run independently for each model. In this approach, the model is run iteratively, removing one redundant feature at a time until the model performance decreases the least. In this process, the k-fold cross validation method (10-fold) with RMSE as the objective function is used. Then, the Variance Inflation Factor (VIF) is applied to the subset of independent features selected by the RFEcv to reduce the multi-collinearity even further. In this process, only the independent features with a VIF inferior to 10 are selected [81].

To assess the potential benefits of the feature selection process, each machine learning model (i.e., RF, GT, GB, and XGB) was trained and validated with (i) the optimized independent feature subset (i.e., obtained after the successive application of RFEcv and VIF) and (ii) with all independent features (i.e., without features selection). In this process, 70% and 30% of the learning database were used for the training and validation steps, respectively. The models' reliability was evaluated with the coefficient of determination (R2, Equation (1)) and the Root Mean Square Error (RMSE, Equation (2)).

$$R^{2} = \frac{\sum_{i=1}^{n} (P_{i} - O_{i})^{2}}{\sum_{i=1}^{n} (O_{i} - O_{i})^{2}}$$
(1)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - O_i)^2}$$
 (2)



where n represents the number of samples and  $P_i$  and  $O_i$  represent the predicted and observed values of salinity content at the site *i*, respectively.

Figure 2. Salinity modeling process flow chart.

## 4. Results

4.1. Features Selection and Importance

Figure 3 shows the feature importance of the selected independent variables for each ML model and scenario. For scenario-1 (R), only three out of the five texture indices (ITs) were selected for modeling (IT2, IT3, and IT4). According to their equations (Table 4), all ITs are very similar, and therefore, their combination is not expected to bring additional benefits, which can explain why only a few remain after the feature selection process.

For scenario-2 (O), the process of feature selection significantly decreases the quantity of independent variables. The RF and GB models only select six out of twenty-six independent features, while the XGB and DT models selected only five out of twenty-six independent features. As for scenario-1, the indices corresponding to scenario-2 were similar in terms of the considered bands. Half of the indices was based on a combination of NIR and/or Red and/or SWIR bands (Table 2). The MSI2 index is the only one based on the three SWIR, NIR, and Red bands and thus was the most important feature. Therefore, MSI2 was seen as an effective way to synthetize the information brought by these bands. The NDWI is the second most important for all models. This index includes other bands' information (i.e., Green) than the Red, NIR, and SWIR bands. No indices based on the Blue or Vegetation Red Edges (B5, B6, B7, and B8a) bands were selected, which shows that soil salinity content is not sensitive to the spectral range of these bands.



**Figure 3.** Feature contribution for (**a**) scenario-1, (**b**) scenario-2, (**c**) scenario-3, and (**d**) scenario-4. S1, S2, and topography variables are in italic, normal, and bold italic fonts, respectively.

For scenario-3, all selected S2 indices (excepting SI2 for XGB model) are based on Red, NIR, SWIR, and Green bands, as for scenario-2. When combining S1 and S2 products, the most important indices come from S2 (i.e., MSI2 and NDWI), with the indices' importance three times higher than those of the selected S1 features. This shows that in plowed lands, soil salt content is more sensitive to S2 than S1.

For scenario-4, no S1 indices remain, which confirms the low sensitivity of soil salt contents to S1 (as highlighted for scenario-3). The selected features are optical (MSI2 and NDWI) and topographic (Elevation and Slope). Elevation and slope were related to the identification of sites where water accumulates, evaporates, and leaves salts in the soil. Furthermore, lower areas are more exposed to groundwater rise and evaporation that will concentrate soil salts. These processes involve soil humidity dynamics that can explain the high contribution of MSI2, NDWI, and NDVI indices. Indeed, these indices are based on the NIR band (Table 2), which allows for the identification of soil moisture and was previously used across the considered region to separate dry soil from water bodies [46].

In summary, the selected variables show that humidity and drainage pattern play an important role in the soil salinization process.

#### 4.2. Soil Salinity Estimates Prediction

Table 5 shows the metrics obtained in the validation stage for the models with and without feature selection.

Model	Metric	Scen (1	ario-1 R)	Scenario-2 (O)		Scenario-3 (R+O)		Scenario-4 (R+O+T)	
			TT	SR	TT	SR	TT	SR	TT
	R <sup>2</sup>	0.21	0.19	0.57	0.62	0.58	0.63	0.66	0.75 **
RF	RMSE (µS·cm <sup>−1</sup> )	5646	2004	4143	3617	3776	3550	3820	2230 **
	R <sup>2</sup>	0.03	0.07	0.64	0.69	0.63	0.66	0.76	0.76 **
XGB	RMSE (µS·cm <sup>−1</sup> )	7151	5247	3652	3336	3708	3773	3042	2890 **
DT	R <sup>2</sup>	0.07	-0.02	0.54	0.66	0.55	0.56	0.64	0.73 **
	RMSE (µS∙cm <sup>−1</sup> )	6638	6643	4832	3252	4610	3398	4206	2658
GB	R <sup>2</sup>	0.16	0.10	0.56	0.62	0.64	0.73	0.70	0.74 **
	RMSE (µS·cm <sup>−1</sup> )	4513	6212	3897	3604	3507	3026	3421	3027 **

Table 5. Performance of ML models with (SR) and without (TT) feature selection.

Where TT = All features; SR = selected features with RFEcv and VIF. \*\* Performance obtained without radar variables (R).

The feature selection considerably improves the models' reliability: for most scenarios, feature selection (RFEcv + VIF) increases (decreases)  $R^2$  (RMSE). For the XGB model and scenario-1 (scenario-3), the improvement is limited to RMSE ( $R^2$ ). For example, when considering scenario-4, RMSE is reduced by 42%, 5%, 37%, and 12% (RF, XGB, DT, and GB models, respectively) by using feature selection.

All models provided much less reliable soil salinity estimates for scenario-1 than for scenario-2. Indeed, improvements in  $\mathbb{R}^2$  values of 226%, 886%, 3200%, and 520% were observed for RF, XGB, DT, and GB for scenario-2 (O) with respect to scenario-1 (R) (after the feature selection). This shows that estimates of the soil salinity of plowed agricultural lands are more sensitive to optical than radar indices, confirming the lower contribution of radar data highlighted in Figure 3c.

Finally, all models provide the most reliable soil salinity estimates for scenario-4 (O+R+T), indicating that the inclusion (exclusion) of topography (radar) indices lead to significant improvements in the model's reliability (Figure 3d). An increase of 19%, 15%, 30%, and 1% is observed for RF, XGB, DT, and GB models for R<sup>2</sup> when comparing scenario-3 and -4 (Table 5).

## 4.3. Soil Salinity Mapping

Figure 4 shows soil salinity maps derived from the models based on the most reliable scenario for October 2023 (i.e., scenario-4).

Although all models produced similar metrics in the validation stage (Table 5), substantial differences are observed in the salinity spatial distribution predicted by the models. Taking as example an area near Lake Poopó (Figure 4), the XGB model yields more homogenous soil salinity mapping with very high salt concentrations (i.e., EC from 8000 to 15,000  $\mu$ S/cm, Figure 4). On the contrary, the DT shows abrupt changes from high and extreme soil salinities in most of the area to no saline condition (i.e., <750  $\mu$ S/cm) in the northwestern corner.

RF and GB models report a more balanced distribution with dominant high to very high salinity conditions. Both models agree with the higher salinity content in the southwestern part. However, the soil samples have a better agreement with the RF estimates than with the GB ones.





**Figure 4.** Soil salinity maps for the area (**a**) derived from (**b**) RF, (**c**) DT, (**d**) XGB, and (**e**) GB models for scenario-4 with feature selection—October 2023.

## 5. Discussion

Topography plays a fundamental role in soil formation and is therefore used as a key feature to estimate different soil properties, including salinity [1]. Actually, topography provides valuable information to identify regions where water is prone to accumulate (and therefore evaporate) and leave salt precipitates in the soils. Moreover, topography allows for the identification of depression zones which are prone to be closer to the groundwater table and are therefore more exposed to groundwater upraise which leaves salt precipitates in the soil [7]. In this context, numerous studies have already reported on the benefits of including topography indices as an independent feature in ML models for soil salinity mapping [27,82,83]. These findings are consistent with the results obtained in the current investigation. Indeed, a significant improvement in the soil salinity estimates is observed once topographic indices are included as independent features (i.e., scenario-4). By identifying topography as a key factor controlling soil salinity distribution, more reliable models could be generated with a more precise DEM. A study led in the United Kingdom shows that the use of a precise DEM derived from drones considerably improves the ML reliability for soil organic carbon predictions in plow lands [84]. As micro-topography inside plowed areas has already been identified as a key factor in soil salinity spatial distribution [85,86], such a DEM should considerably improve the model's reliability.

Additionally, regions where water accumulates (and therefore where salts are present) are prone to higher soil humidity, which can be detected with spectral indices based on SWIR and NIR bands [87,88]. In this context, the MSI2 and NDWI indices are among the most relevant features for all scenarios and models. At an order of magnitude lower in relevance relative to the previous optical indices, the NDVI was also identified for most models, with this index being more relevant when monitoring saline soils with crops [89]. Numerous studies reported that S1 gathered valuable information on soil humidity estimates [90–93]. Based on these studies, it was expected that S1-derived features should improve soil salinity modeling. However, for all models, scenario-1 (only S1) reaches the lowest performance and S1 features' contribution is insignificant to scenario-3 (S1 + S2) and -4 (S1 + S2 + T). This is due to the fact that the S1 C-band radar signal saturates

in the case of plowed soils because plowed land roughness generally surpasses C-band wavelength (approximately 5.5 cm) [94,95].

Previous studies reporting on S1 sensitivity to soil humidity (and therefore soil salinity) were limited to non-disturbed soils. Indeed, more reliable soil salinity estimates were obtained in non-disturbed soil (low roughness) with S1 variables than with S2 features [26]. These contradictory observations regarding radar backscatter are sensitive to the dielectric constant [23] in non-disturbed and plowed soils, drawing attention to the necessity of providing knowledge of local soil conditions in order to identify the model limitations. Along this line, it is worth mentioning that S1 and S2 are not only sensitive to soil salinity, but also to soil organic carbon [96], nitrogen [97–100], texture [101,102], and pH [103–105]. In this regard, a model trained in a specific region may not apply to another region with different soil parameters (i.e., organic matter, nitrogen and texture). Thus, the proposed approach could be enhanced by training a specific model for every group of soils with a range of similar properties (chemical, physical, and biological) and that are spatially close (i.e., soil associations).

Finally, the model was developed over a specific time period (2022–2023), but trained for the main Altiplano soil type (Figure 1) where agriculture take place. As the soil type is not expected to change over time, the proposed model could be used to retrieve soil salinity maps along the Sentinel (-1 and -2) observation period (i.e., 2015-present). Therefore, this model could be used to assess soil salinity trends, and it links with agriculture practices (i.e., extensive quinoa cultivation in the Altiplano known as quinoa boom [106,107]). Such information could highlight the limit of current agricultural practice regarding its impact on soil salinity which might decrease soil productivity, therefore threatening both food security and socio-economic conditions. Such knowledge will promote investigations on regenerative agriculture effects [108] to define sustainable agriculture practices.

## 6. Conclusions

This study assesses the potential of Sentinel-2 and -1 (S2 and S1) products alone, together, and combined with topographic information for the soil salinity mapping of agricultural plowed lands. For this task, four machine learning (ML) models were used (Decision Tree—DT, random forest—RF, Gradient Boosting—GB, Extreme Gradient Boosting—XGB). The main results can be summarized as follows:

- Plowed land soil salinity estimates cannot be retrieved from S1 as the C-band wavelength is inferior to plowed land roughness, thus leading to the saturation of the signal.
- Accurate soil salinity estimates can be retrieved from S2 features alone ( $R^2 > 0.6$ ), but the addition of topography features considerably improve the model's reliability ( $R^2 > 0.7$ ).
- The most relevant independent variables are the Moisture Saturation Index 2, Normalized Difference Water Index, and Elevation. These variables are related to the identification of water accumulation spots where salt is also expected to accumulate.
- For all ML models, the feature selection process significantly improves the model's reliability, up to 13.6% in the best case.
- The four ML models reach similar global performance. However, high discrepancy is observed when using the models for the mapping of soil salinity in plowed lands in the studied region.

This study highlights the possibility of model soil salinity on plowed land in the Bolivian Altiplano, integrating data from remote sensing and machine learning algorithms. This information could be retrieved before the sowing period in order to provide information to minimize soil salinity risks with appropriate drainage measures. Furthermore, the higher spatial resolution of the elevation/relief information than that provided by the global elevation model (i.e., SRTM) should enhance the micro-topography representation on the agriculture plot scale. Such information would help to identify water accumulation zones inside the agriculture plot and therefore improve soil salinity modeling.

Finally, this study highlights remote sensing potential to assess soil salinization sensitivity to different factors (i.e., climate, irrigation, soil management). This knowledge could support the formulation of modern policies towards the sustainable management of agricultural lands affected by salinity under an integrative and preventive approach to minimize soil degradation and/or favor their rehabilitation.

Author Contributions: Conceptualization, D.T. and F.S.; methodology, D.T. and F.S.; formal analysis, D.T., F.S. and R.P.Z.; investigation, D.T., F.S., R.P.Z., R.H., E.Y., R.E.-V., H.S., B.C., J.M.-C. and R.M.; data curation, D.T., F.S. and E.Y.; writing—original draft preparation, D.T. and F.S.; writing—review and editing, F.S., J.M.-C., R.E.-V., R.H., H.S., B.C., R.P.Z. and R.M.; supervision F.S.; project administration, F.S.; funding acquisition, F.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Agropolis Foundation (project number: 2001-032) in the framework of the project WASACA (Wastewater irrigation: a sustainable agriculture adaptation to climate changes over the Bolivian Altiplano?) and by the Centre National d'Etudes Spatiales (CNES) in the framework of the QUIMONOS project (Quinoa monitoring by satellite). The first author is grateful to the IRD (Institut de Recherche pour le Développement) for its financial support.

**Data Availability Statement:** The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

## References

- Omuto, C.T.; Vargas, R.R.; El Mobarak, A.M.; Mohamed, N.; Viatkin, K.; Yigini, Y. Mapeo de Suelos Afectados por Salinidad—Manual Técnico; FAO: Rome, Italy, 2021; ISBN 978-92-5-133840-7.
- Sahbeni, G.; Ngabire, M.; Musyimi, P.K.; Székely, B. Challenges and Opportunities in Remote Sensing for Soil Salinization Mapping and Monitoring: A Review. *Remote Sens.* 2023, 15, 2540. [CrossRef]
- Porta, J.; López-Acevedo, M.; Poch, R.M. Edafología: Uso y Protección de Suelos, 4th ed.; Mundi-Prensa: Madrid, Spain, 2019; ISBN 978-84-8476-750-3.
- 4. Ondrasek, G.; Rathod, S.; Manohara, K.K.; Gireesh, C.; Anantha, M.S.; Sakhare, A.S.; Parmar, B.; Yadav, B.K.; Bandumula, N.; Raihan, F.; et al. Salt Stress in Plants and Mitigation Approaches. *Plants* **2022**, *11*, 717. [CrossRef]
- 5. Wicke, B.; Smeets, E.; Dornburg, V.; Vashev, B.; Gaiser, T.; Turkenburg, W.; Faaij, A. The Global Technical and Economic Potential of Bioenergy from Salt-Affected Soils. *Energy Environ. Sci.* **2011**, *4*, 2669–2681. [CrossRef]
- 6. Zubieta, R.; Molina-Carpio, J.; Laqui, W.; Sulca, J.; Ilbay, M. Comparative Analysis of Climate Change Impacts on Meteorological, Hydrological, and Agricultural Droughts in the Lake Titicaca Basin. *Water* **2021**, *13*, 175. [CrossRef]
- Stavi, I.; Thevs, N.; Priori, S. Soil Salinity and Sodicity in Drylands: A Review of Causes, Effects, Monitoring, and Restoration Measures. *Front. Environ. Sci.* 2021, 9, 712831. [CrossRef]
- Jamil, A.; Riaz, S.; Ashraf, M.; Foolad, M.R. Gene Expression Profiling of Plants under Salt Stress. Crit. Rev. Plant Sci. 2011, 30, 435–458. [CrossRef]
- Wichelns, D.; Qadir, M. Achieving Sustainable Irrigation Requires Effective Management of Salts, Soil Salinity, and Shallow Groundwater. *Agric. Water Manag.* 2015, 157, 31–38. [CrossRef]
- Kempen, B.; Brus, D.J.; Stoorvogel, J.J.; Heuvelink, G.B.M.; de Vries, F. Efficiency Comparison of Conventional and Digital Soil Mapping for Updating Soil Maps. Soil Sci. Soc. Am. J. 2012, 76, 2097–2115. [CrossRef]
- 11. Brus, D.J.; Bogaert, P.; Heuvelink, G.B.M. Bayesian Maximum Entropy Prediction of Soil Categories Using a Traditional Soil Map as Soft Information. *Eur. J. Soil Sci.* 2008, *59*, 166–177. [CrossRef]
- 12. Zhu, A.X.; Hudson, B.; Burt, J.; Lubich, K.; Simonson, D. Soil Mapping Using GIS, Expert Knowledge, and Fuzzy Logic. *Soil Sci. Soc. Am. J.* 2001, *65*, 1463–1472. [CrossRef]
- 13. McBratney, A.B.; Mendonça Santos, M.L.; Minasny, B. On Digital Soil Mapping. Geoderma 2003, 117, 3–52. [CrossRef]
- 14. Alavipanah, S.K.; Damavandi, A.; Mirzaei, S.; Rezaei, A.; Hamzeh, S.; Matinfar, H.; Teimouri, H.; Javadzarrin, I. Remote Sensing Application in Evaluation of Soil Characteristics in Desert Areas. *Nat. Environ. Change* **2016**, *2*, 1–24.
- 15. Wadoux, A.M.J.-C.; Minasny, B.; McBratney, A.B. Machine Learning for Digital Soil Mapping: Applications, Challenges and Suggested Solutions. *Earth-Sci. Rev.* 2020, 210, 103359. [CrossRef]
- 16. Abdulraheem, M.I.; Zhang, W.; Li, S.; Moshayedi, A.J.; Farooque, A.A.; Hu, J. Advancement of Remote Sensing for Soil Measurements and Applications: A Comprehensive Review. *Sustainability* **2023**, *15*, 15444. [CrossRef]
- 17. Ivushkin, K.; Bartholomeus, H.; Bregt, A.K.; Pulatov, A.; Kempen, B.; de Sousa, L. Global Mapping of Soil Salinity Change. *Remote Sens. Environ.* **2019**, 231, 111260. [CrossRef]
- 18. Yan, Y.; Kayem, K.; Hao, Y.; Shi, Z.; Zhang, C.; Peng, J.; Liu, W.; Zuo, Q.; Ji, W.; Li, B. Mapping the Levels of Soil Salination and Alkalization by Integrating Machining Learning Methods and Soil-Forming Factors. *Remote Sens.* **2022**, *14*, 3020. [CrossRef]

- Yahiaoui, I.; Bradaï, A.; Douaoui, A.; Abdennour, M.A. Performance of Random Forest and Buffer Analysis of Sentinel-2 Data for Modelling Soil Salinity in the Lower-Cheliff Plain (Algeria). *Int. J. Remote Sens.* 2021, 42, 148–171. [CrossRef]
- Gorji, T.; Yildirim, A.; Hamzehpour, N.; Tanik, A.; Sertel, E. Soil Salinity Analysis of Urmia Lake Basin Using Landsat-8 OLI and Sentinel-2A Based Spectral Indices and Electrical Conductivity Measurements. *Ecol. Indic.* 2020, 112, 106173. [CrossRef]
- Sirpa-Poma, J.W.; Satgé, F.; Resongles, E.; Pillco-Zolá, R.; Molina-Carpio, J.; Flores Colque, M.G.; Ormachea, M.; Pacheco Mollinedo, P.; Bonnet, M.-P. Towards the Improvement of Soil Salinity Mapping in a Data-Scarce Context Using Sentinel-2 Images in Machine-Learning Models. *Sensors* 2023, 23, 9328. [CrossRef]
- 22. Taghadosi, M.M.; Hasanlou, M.; Eftekhari, K. Soil Salinity Mapping Using Dual-Polarized SAR Sentinel-1 Imagery. *Int. J. Remote Sens.* 2019, 40, 237–252. [CrossRef]
- 23. Hoa, P.V.; Giang, N.V.; Binh, N.A.; Hai, L.V.H.; Pham, T.-D.; Hasanlou, M.; Tien Bui, D. 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. [CrossRef]
- 24. Periasamy, S.; Ravi, K.P.; Tansey, K. Identification of Saline Landscapes from an Integrated SVM Approach from a Novel 3-D Classification Schema Using Sentinel-1 Dual-Polarized SAR Data. *Remote Sens. Environ.* **2022**, 279, 113144. [CrossRef]
- 25. Periasamy, S.; Ravi, K.P. A Novel Approach to Quantify Soil Salinity by Simulating the Dielectric Loss of SAR in Three-Dimensional Density Space. *Remote Sens. Environ.* **2020**, *251*, 112059. [CrossRef]
- Sirpa-Poma, J.W.; Satgé, F.; Pillco Zolá, R.; Resongles, E.; Perez-Flores, M.; Flores Colque, M.G.; Molina-Carpio, J.; Ramos, O.; Bonnet, M.-P. Complementarity of Sentinel-1 and Sentinel-2 Data for Soil Salinity Monitoring to Support Sustainable Agriculture Practices in the Central Bolivian Altiplano. *Sustainability* 2024, *16*, 6200. [CrossRef]
- 27. Ma, G.; Ding, J.; Han, L.; Zhang, Z.; Ran, S. Digital Mapping of Soil Salinization Based on Sentinel-1 and Sentinel-2 Data Combined with Machine Learning Algorithms. *Reg. Sustain.* **2021**, *2*, 177–188. [CrossRef]
- Zhou, W.; Rao, P.; Jat, M.L.; Singh, B.; Poonia, S.; Bijarniya, D.; Kumar, M.; Singh, L.K.; Schulthess, U.; Singh, R.; et al. Using Sentinel-2 to Track Field-Level Tillage Practices at Regional Scales in Smallholder Systems. *Remote Sens.* 2021, 13, 5108. [CrossRef]
- 29. Zheng, X.; Feng, Z.; Li, L.; Li, B.; Jiang, T.; Li, X.; Li, X.; Chen, S. Simultaneously Estimating Surface Soil Moisture and Roughness of Bare Soils by Combining Optical and Radar Data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *100*, 102345. [CrossRef]
- Ebrahimzadeh, G.; Yaghmaeian Mahabadi, N.; Bayat, H.; MatinFar, H. Estimating Pre-Compression Stress in Agricultural Soils: Integrating Spectral Indices and Soil Properties through Machine Learning. *Comput. Electron. Agric.* 2023, 215, 108393. [CrossRef]
- Ramos, T.B.; Castanheira, N.; Oliveira, A.R.; Paz, A.M.; Darouich, H.; Simionesei, L.; Farzamian, M.; Gonçalves, M.C. Soil Salinity Assessment Using Vegetation Indices Derived from Sentinel-2 Multispectral Data. Application to Lezíria Grande, Portugal. Agric. Water Manag. 2020, 241, 106387. [CrossRef]
- 32. Gopalakrishnan, T.; Kumar, L. Modeling and Mapping of Soil Salinity and Its Impact on Paddy Lands in Jaffna Peninsula, Sri Lanka. *Sustainability* **2020**, *12*, 8317. [CrossRef]
- Wang, N.; Peng, J.; Xue, J.; Zhang, X.; Huang, J.; Biswas, A.; He, Y.; Shi, Z. A Framework for Determining the Total Salt Content of Soil Profiles Using Time-Series Sentinel-2 Images and a Random Forest-Temporal Convolution Network. *Geoderma* 2022, 409, 115656. [CrossRef]
- 34. Wang, N.; Peng, J.; Chen, S.; Huang, J.; Li, H.; Biswas, A.; He, Y.; Shi, Z. Improving Remote Sensing of Salinity on Topsoil with Crop Residues Using Novel Indices of Optical and Microwave Bands. *Geoderma* **2022**, 422, 115935. [CrossRef]
- Mohamed, S.A.; Metwaly, M.M.; Metwalli, M.R.; AbdelRahman, M.A.E.; Badreldin, N. Integrating Active and Passive Remote Sensing Data for Mapping Soil Salinity Using Machine Learning and Feature Selection Approaches in Arid Regions. *Remote Sens.* 2023, 15, 1751. [CrossRef]
- Golestani, M.; Mosleh Ghahfarokhi, Z.; Esfandiarpour-Boroujeni, I.; Shirani, H. Evaluating the Spatiotemporal Variations of Soil Salinity in Sirjan Playa, Iran Using Sentinel-2A and Landsat-8 OLI Imagery. CATENA 2023, 231, 107375. [CrossRef]
- Aksoy, S.; Yildirim, A.; Gorji, T.; Hamzehpour, N.; Tanik, A.; Sertel, E. Assessing the Performance of Machine Learning Algorithms for Soil Salinity Mapping in Google Earth Engine Platform Using Sentinel-2A and Landsat-8 OLI Data. *Adv. Space Res.* 2022, 69, 1072–1086. [CrossRef]
- 38. 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. [CrossRef]
- Naimi, S.; Ayoubi, S.; Zeraatpisheh, M.; Dematte, J.A.M. Ground Observations and Environmental Covariates Integration for Mapping of Soil Salinity: A Machine Learning-Based Approach. *Remote Sens.* 2021, 13, 4825. [CrossRef]
- Hervé, D.; Lezdema, R.; Orsag, V.; Coudrain, A.; Talbi, A.; Ribstein, P.; Marsily, G.D.; Flores, M.; Mita, V.; Chungara, J.; et al. Limitantes y Manejo de Los Suelos Salinos y/o Sodicos En El Altiplano Boliviano. 2002. Available online: https://core.ac.uk/ download/pdf/39844995.pdf (accessed on 5 January 2023).
- Satge, F.; Denezine, M.; Pillco, R.; Timouk, F.; Pinel, S.; Molina, J.; Garnier, J.; Seyler, F.; Bonnet, M.-P. Absolute and Relative Height-Pixel Accuracy of SRTM-GL1 over the South American Andean Plateau. *ISPRS J. Photogramm. Remote Sens.* 2016, 121, 157–166. [CrossRef]
- Satgé, F.; Bonnet, M.P.; Timouk, F.; Calmant, S.; Pillco, R.; Molina, J.; Lavado-Casimiro, W.; Arsen, A.; Crétaux, J.F.; Garnier, J. Accuracy Assessment of SRTM v4 and ASTER GDEM v2 over the Altiplano Watershed Using ICESat/GLAS Data. *Int. J. Remote Sens.* 2015, *36*, 465–488. [CrossRef]

- 43. Miranda, R. Adubação Orgânica em Condições de Irrigação Suplementar e Seu Efeito na Produtividade da Quinua (Chenopodium Quinoa Willd) No Planalto da Bolívia. Tese de Doutorado, Universidade Federal de Santa Maria. 2012. Available online: http://repositorio.ufsm.br/handle/1/3337 (accessed on 20 April 2023).
- 44. Satgé, F.; Pillco, R.; Molina-Carpio, J.; Mollinedo, P.P.; Bonnet, M.-P. Reliability of Gridded Temperature Datasets to Monitor Surface Air Temperature Variability over Bolivia. *Int. J. Climatol.* **2023**, *43*, 6191–6206. [CrossRef]
- 45. Garcia, M.; Raes, D.; Allen, R.; Herbas, C. Dynamics of Reference Evapotranspiration in the Bolivian Highlands (Altiplano). *Agric. For. Meteorol.* **2004**, *125*, 67–82. [CrossRef]
- Satgé, F.; Espinoza, R.; Zolá, R.P.; Roig, H.; Timouk, F.; Molina, J.; Garnier, J.; Calmant, S.; Seyler, F.; Bonnet, M.-P. Role of Climate Variability and Human Activity on Poopó Lake Droughts between 1990 and 2015 Assessed Using Remote Sensing Data. *Remote* Sens. 2017, 9, 218. [CrossRef]
- Canedo, C.; Pillco Zolá, R.; Berndtsson, R. Role of Hydrological Studies for the Development of the TDPS System. Water 2016, 8, 144. [CrossRef]
- World Reference Base for Soil Resources. International Soil Classification System for Naming Soils and Creating Legends for Soil Maps, 4th ed.; International Union of Soil Sciences (IUSS): Viena, Austria, 2022; ISBN 979-8-9862451-1-9.
- 49. FAO. Standard Operating Procedure for Soil Electrical Conductivity, Soil/Water, 1:5; FAO: Rome, Italy, 2021.
- 50. Orsag, V. El Recurso Suelo: Principios Para su Manejo y Conservación, 1st ed.; Zeus: La Paz, Bolivia, 2010; ISBN 978-99954-732-9-7.
- Coudrain-Ribstein, A.; Olive, P.; Quintanilla, J.; Sondag, F.; Cahuaya, D. Salinity and Isotopic Dynamics of the Groundwater Resources on the Bolivian Altiplano. In Proceedings of the 1994 International Symposium on Application of Tracers in Arid Zone Hydrology, Vienna, Italy, 22–26 August 1994; pp. 267–276.
- 52. Garcia, M.E.; Bengtsson, L.; Persson, K.M. On the Distribution of Saline Groundwater in the Poopo Basin, Central Bolivian Highland. *J. Water Manag. Res.* 2010, *66*, 199–203.
- 53. Luo, C.; Zhang, W.; Zhang, X.; Liu, H. Mapping Soil Organic Matter Content Using Sentinel-2 Synthetic Images at Different Time Intervals in Northeast China. *Int. J. Digit. Earth* **2023**, *16*, 1094–1107. [CrossRef]
- 54. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, 202, 18–27. [CrossRef]
- Sahbeni, G.; Székely, B. Salinity Levels Discrimination Using ERS-1/2 and Sentinel-1 SAR Time Series Data in Hortobágyi National Park, Hungary. In Proceedings of the 2022 IEEE Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), Istanbul, Turkey, 7–9 March 2022; pp. 194–197.
- 56. Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- 57. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
- 58. Quinlan, J.R. Induction of Decision Trees. Mach. Learn. 1986, 1, 81–106. [CrossRef]
- 59. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. Ann. Stat. 2001, 29, 1189–1232. [CrossRef]
- Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. Classification and Regression Trees; Taylor & Francis: Abingdon, UK, 1984; ISBN 978-0-412-04841-8.
- 61. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- 62. Safanelli, J.L.; Poppiel, R.R.; Ruiz, L.F.C.; Bonfatti, B.R.; Mello, F.A.d.O.; Rizzo, R.; Demattê, J.A.M. Terrain Analysis in Google Earth Engine: A Method Adapted for High-Performance Global-Scale Analysis. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 400. [CrossRef]
- 63. Farifteh, J.; van der Meer, F.; van der Meijde, M.; Atzberger, C. Spectral Characteristics of Salt-Affected Soils: A Laboratory Experiment. *Geoderma* 2008, 145, 196–206. [CrossRef]
- Amal, A.; Lalit, K. Soil Salinity Mapping and Monitoring in Arid and Semi-Arid Regions Using Remote Sensing Technology: A Review. Adv. Remote Sens. 2013, 2, 373–385. [CrossRef]
- Gholizadeh, A.; Žižala, D.; Saberioon, M.; Borůvka, L. Soil Organic Carbon and Texture Retrieving and Mapping Using Proximal, Airborne and Sentinel-2 Spectral Imaging. *Remote Sens. Environ.* 2018, 218, 89–103. [CrossRef]
- 66. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with Erts. *NASA Spec. Publ.* **1974**, 351, 309.
- 67. Huete, A.R.; Liu, H.Q.; Batchily, K.; van Leeuwen, W. A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [CrossRef]
- 68. 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. [CrossRef]
- Wilson, E.H.; Sader, S.A. Detection of Forest Harvest Type Using Multiple Dates of Landsat TM Imagery. *Remote Sens. Environ.* 2002, 80, 385–396. [CrossRef]
- 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. [CrossRef]
- Chen, W.; Liu, L.; Zhang, C.; Wang, J.; Wang, J.; Pan, Y. Monitoring the Seasonal Bare Soil Areas in Beijing Using Multi-Temporal TM Images. In Proceedings of the IGARSS 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 20–24 September 2004; Volume 5, pp. 3379–3382.

- 72. García, M.J.L.; Caselles, V. Mapping Burns and Natural Reforestation Using Thematic Mapper Data. *Geocarto Int.* **1991**, *6*, 31–37. [CrossRef]
- Key, C.; Benson, N. Landscape Assessment (LA) Sampling and Analysis Methods. In *FIREMON: Fire Effects Monitoring and Inventory System*; U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: Fort Collins, CO, USA, 2005; pp. LA 1–LA 51.
- 74. McFeeters, S.K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
- 75. Douaoui, A.E.K.; Nicolas, H.; Walter, C. Detecting Salinity Hazards within a Semiarid Context by Means of Combining Soil and Remote-Sensing Data. *Geoderma* **2006**, *134*, 217–230. [CrossRef]
- 76. Dehni, A.; Lounis, M. Remote Sensing Techniques for Salt Affected Soil Mapping: Application to the Oran Region of Algeria. *Procedia Eng.* **2012**, *33*, 188–198. [CrossRef]
- 77. 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. [CrossRef]
- 78. Khan, N.; Rastoskuev, V.V.; Shalina, E.; Sato, Y. Mapping Salt-Affected Soils Using Remote Sensing Indicators-A Simple Approach With the Use of GIS IDRISI. In Proceedings of the 22nd Asian Conference on Remote Sensing, Singapore, 5–9 November 2001.
- 79. Lehner, B.; Verdin, K.; Jarvis, A. New Global Hydrography Derived From Spaceborne Elevation Data. *Eos Trans. Am. Geophys. Union* **2008**, *89*, 93–94. [CrossRef]
- 80. Awad, M.; Fraihat, S. Recursive Feature Elimination with Cross-Validation with Decision Tree: Feature Selection Method for Machine Learning-Based Intrusion Detection Systems. *J. Sens. Actuator Netw.* **2023**, *12*, 67. [CrossRef]
- 81. Murray, L.; Nguyen, H.; Lee, Y.-F.; Remmenga, M.; Smith, D. Variance Inflation Factors in Regression Models with Dummy Variables. *Conf. Appl. Stat. Agric.* 2012. [CrossRef]
- Triki Fourati, H.; Bouaziz, M.; Benzina, M.; Bouaziz, S. Detection of Terrain Indices Related to Soil Salinity and Mapping Salt-Affected Soils Using Remote Sensing and Geostatistical Techniques. *Environ. Monit. Assess.* 2017, 189, 177. [CrossRef] [PubMed]
- 83. Shahrayini, E.; Noroozi, A.A. Modeling and Mapping of Soil Salinity and Alkalinity Using Remote Sensing Data and Topographic Factors: A Case Study in Iran. *Environ. Model. Assess.* **2022**, *27*, 901–913. [CrossRef]
- 84. Cutting, B.J.; Atzberger, C.; Gholizadeh, A.; Robinson, D.A.; Mendoza-Ulloa, J.; Marti-Cardona, B. Remote Quantification of Soil Organic Carbon: Role of Topography in the Intra-Field Distribution. *Remote Sens.* **2024**, *16*, 1510. [CrossRef]
- 85. Akça, E.; Aydin, M.; Kapur, S.; Kume, T.; Nagano, T.; Watanabe, T.; Çilek, A.; Zorlu, K. Long-Term Monitoring of Soil Salinity in a Semi-Arid Environment of Turkey. *CATENA* **2020**, *193*, 104614. [CrossRef]
- Zhang, K.; Xia, J.; Su, L.; Gao, F.; Cui, Q.; Xing, X.; Dong, M.; Li, C. Effects of Microtopographic Patterns on Plant Growth and Soil Improvement in Coastal Wetlands of the Yellow River Delta. *Front. Plant Sci.* 2023, 14, 1162013. [CrossRef] [PubMed]
- 87. Yue, J.; Tian, J.; Tian, Q.; Xu, K.; Xu, N. Development of Soil Moisture Indices from Differences in Water Absorption between Shortwave-Infrared Bands. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 216–230. [CrossRef]
- Al-Ali, Z.; Abulibdeh, A.; Al-Awadhi, T.; Mohan, M.; Al Nasiri, N.; Al-Barwani, M.; Al Nabbi, S.; Abdullah, M. Examining the Potential and Effectiveness of Water Indices Using Multispectral Sentinel-2 Data to Detect Soil Moisture as an Indicator of Mudflow Occurrence in Arid Regions. *Int. J. Appl. Earth Obs. Geoinf.* 2024, 130, 103887. [CrossRef]
- Gómez Flores, J.L.; Ramos Rodríguez, M.; González Jiménez, A.; Farzamian, M.; Herencia Galán, J.F.; Salvatierra Bellido, B.; Cermeño Sacristan, P.; Vanderlinden, K. Depth-Specific Soil Electrical Conductivity and NDVI Elucidate Salinity Effects on Crop Development in Reclaimed Marsh Soils. *Remote Sens.* 2022, 14, 3389. [CrossRef]
- Bauer-Marschallinger, B.; Freeman, V.; Cao, S.; Paulik, C.; Schaufler, S.; Stachl, T.; Modanesi, S.; Massari, C.; Ciabatta, L.; Brocca, L.; et al. Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 520–539. [CrossRef]
- Balenzano, A.; Mattia, F.; Satalino, G.; Lovergine, F.P.; Palmisano, D.; Peng, J.; Marzahn, P.; Wegmüller, U.; Cartus, O.; Dabrowska-Zielińska, K.; et al. Sentinel-1 Soil Moisture at 1 Km Resolution: A Validation Study. *Remote Sens. Environ.* 2021, 263, 112554. [CrossRef]
- 92. Bhogapurapu, N.; Dey, S.; Homayouni, S.; Bhattacharya, A.; Rao, Y.S. Field-Scale Soil Moisture Estimation Using Sentinel-1 GRD SAR Data. *Adv. Space Res.* 2022, *70*, 3845–3858. [CrossRef]
- 93. Mohseni, F.; Mirmazloumi, S.M.; Mokhtarzade, M.; Jamali, S.; Homayouni, S. Global Evaluation of SMAP/Sentinel-1 Soil Moisture Products. *Remote Sens.* 2022, 14, 4624. [CrossRef]
- 94. Baghdadi, N.; El Hajj, M.; Choker, M.; Zribi, M.; Bazzi, H.; Vaudour, E.; Gilliot, J.-M.; Ebengo, D.M. Potential of Sentinel-1 Images for Estimating the Soil Roughness over Bare Agricultural Soils. *Water* **2018**, *10*, 131. [CrossRef]
- 95. Liu, Y.; Rao, P.; Zhou, W.; Singh, B.; Srivastava, A.K.; Poonia, S.P.; Berkel, D.V.; Jain, M. Using Sentinel-1, Sentinel-2, and Planet Satellite Data to Map Field-Level Tillage Practices in Smallholder Systems. *PLoS ONE* **2022**, 17, e0277425. [CrossRef] [PubMed]
- Vaudour, E.; Gholizadeh, A.; Castaldi, F.; Saberioon, M.; Borůvka, L.; Urbina-Salazar, D.; Fouad, Y.; Arrouays, D.; Richer-de-Forges, A.C.; Biney, J.; et al. Satellite Imagery to Map Topsoil Organic Carbon Content over Cultivated Areas: An Overview. *Remote Sens.* 2022, 14, 2917. [CrossRef]

- Zhou, T.; Geng, Y.; Chen, J.; Pan, J.; Haase, D.; Lausch, A. High-Resolution Digital Mapping of Soil Organic Carbon and Soil Total Nitrogen Using DEM Derivatives, Sentinel-1 and Sentinel-2 Data Based on Machine Learning Algorithms. *Sci. Total Environ.* 2020, 729, 138244. [CrossRef]
- Zhou, T.; Lv, W.; Geng, Y.; Xiao, S.; Chen, J.; Xu, X.; Pan, J.; Si, B.; Lausch, A. National-Scale Spatial Prediction of Soil Organic Carbon and Total Nitrogen Using Long-Term Optical and Microwave Satellite Observations in Google Earth Engine. *Comput. Electron. Agric.* 2023, 210, 107928. [CrossRef]
- Zhang, Q.; Liu, M.; Zhang, Y.; Mao, D.; Li, F.; Wu, F.; Song, J.; Li, X.; Kou, C.; Li, C.; et al. Comparison of Machine Learning Methods for Predicting Soil Total Nitrogen Content Using Landsat-8, Sentinel-1, and Sentinel-2 Images. *Remote Sens.* 2023, 15, 2907. [CrossRef]
- 100. Li, Z.; Liu, F.; Peng, X.; Hu, B.; Song, X. Synergetic Use of DEM Derivatives, Sentinel-1 and Sentinel-2 Data for Mapping Soil Properties of a Sloped Cropland Based on a Two-Step Ensemble Learning Method. *Sci. Total Environ.* 2023, 866, 161421. [CrossRef] [PubMed]
- 101. Azizi, K.; Garosi, Y.; Ayoubi, S.; Tajik, S. Integration of Sentinel-1/2 and Topographic Attributes to Predict the Spatial Distribution of Soil Texture Fractions in Some Agricultural Soils of Western Iran. *Soil Tillage Res.* **2023**, *229*, 105681. [CrossRef]
- Ngu, N.H.; Thanh, N.N.; Duc, T.T.; Non, D.Q.; Thuy An, N.T.; Chotpantarat, S. Active Learning-Based Random Forest Algorithm Used for Soil Texture Classification Mapping in Central Vietnam. *CATENA* 2024, 234, 107629. [CrossRef]
- 103. Mammadov, E.; Nowosad, J.; Glaesser, C. Estimation and Mapping of Surface Soil Properties in the Caucasus Mountains, Azerbaijan Using High-Resolution Remote Sensing Data. *Geoderma Reg.* **2021**, *26*, e00411. [CrossRef]
- 104. Zhang, L.; Wu, Z.; Sun, X.; Yan, J.; Sun, Y.; Chen, J. Mapping Topsoil pH Using Different Predictive Models and Covariate Sets in Henan Province, Central China. *Ecol. Inform.* 2023, 78, 102290. [CrossRef]
- 105. Lotfollahi, L.; Delavar, M.A.; Biswas, A.; Jamshidi, M.; Fatehi, S.; Taghizadeh-Mehrjardi, R. Spatial Prediction and Uncertainty Estimation of Crucial GlobalSoilMap Properties—A Contextual Study in the Semi-Arid Area of Western Iran. *Geoderma Reg.* 2023, 35, e00713. [CrossRef]
- 106. Sander, L.; Jacobsen, S.-E. The Quinoa Boom of the Southern Bolivian Altiplano: Linking Geomorphology, Erosion and Spatial Production Patterns. In *Proceedings of the Geophysical Research Abstracts*; University of Copenhagen: Vienna, Austria, 2014; Volume 16, p. 1.
- 107. Barrientos, E.; Carevic, F.; Delatorre, J. La Sustentabilidad Del Altiplano Sur de Bolivia y Su Relación Con La Ampliación de Superficies de Cultivo de Quinua. *Idesia Arica* 2017, *35*, 7–15. [CrossRef]
- 108. Bonifacio, A.; Aroni, G.; Villca, M.; Bentley, J.W. Recovering from Quinoa: Regenerative Agricultural Research in Bolivia. *J. Crop Improv.* **2023**, *37*, 687–708. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.