Article

Complementarity of Sentinel-1 and Sentinel-2 Data for Soil Salinity Monitoring to Support Sustainable Agriculture Practices in the Central Bolivian Altiplano

J. W. Sirpa-Poma [1,2,*], F. Satgé [1,2,*], R. Pillco Zolá [2], E. Resongles [3,4], M. Perez-Flores [1], M. G. Flores Colque [4], J. Molina-Carpio [2], O. Ramos [4] and M.-P. Bonnet [1]

1   ESPACE-DEV, Univ Montpellier, IRD, Univ Antilles, Univ Guyane, Univ Réunion, 34093 Montpellier, France; may.perez.flores@gmail.com (M.P.-F.); marie-paule.bonnet@ird.fr (M.-P.B.)
2   Instituto de Hidráulica e Hidrología, Universidad Mayor de San Andrés, La Paz, Bolivia; rami_lund99@hotmail.com (R.P.Z.); jamolinacarpio@gmail.com (J.M.-C.)
3   HSM, CNRS, IRD, Univ Montpellier, 34093 Montpellier, France; eleonore.resongles@ird.fr
4   Instituto de Investigaciones Químicas, Universidad Mayor de San Andrés, La Paz, Bolivia; marizolflores14@gmail.com (M.G.F.C.); oswalram2@gmail.com (O.R.)
*   Correspondence: jsirpap@fcpn.edu.bo (J.W.S.-P.); frederic.satge@ird.fr (F.S.)

Abstract: Soil salinization will affect 50% of global cropland areas by 2050 and represents a major threat to agricultural production and food sovereignty. As soil salinity monitoring is costly and time consuming, many regions of the world undertake very limited soil salinity observation (in space and time), preventing the accurate assessment of soil salinity hazards. In this context, this study assesses the relative performance of Sentinel-1 radar and Sentinel-2 optical images, and the combination of the two, for monitoring changes in soil salinity at high spatial and temporal resolution, which is essential to evaluate the mitigation measures required for the sustainable adaptation of agriculture practices. For this purpose, an improved learning database made of 863 soil electrical conductivity (i.e., soil salinity) observations is considered for the training/validation step of a Random Forest (RF) model. The RF model is successively trained with (1) only Sentinel-1, (2) only Sentinel-2 and (3) both Sentinel-1 and -2 features using the Genetic Algorithm (GA) to reduce multi-collinearity in the independent variables. Using k-fold cross validation (3-fold), overall accuracy (OA) values of 0.83, 0.88 and 0.95 are obtained when considering only Sentinel-2, only Sentinel-1 and both Sentinel-1 and -2 features as independent variables. Therefore, these results highlight the clear complementarity of radar (i.e., Sentinel-1) and optical (i.e., Sentinel-2) images to improve soil salinity mapping, with OA increases of approximately 10% and 7% when compared to Sentinel-2 and Sentinel-1 alone. Finally, pre-sowing soil salinity maps over a five-year period (2019–2023) are presented to highlight the benefit of the proposed procedure to support the sustainable management of agricultural lands in the context of soil salinization on a regional scale.

Keywords: soil salinity; monitoring; machine learning; remote sensing; sentinel-1; sentinel-2

1. Introduction

1.1. Soil Salinity: A Major Threat Requiring Close Monitoring

At present, soil salinization affects 33% and 20% of irrigated and croplands, respectively [1–3]. In the current general context of global change (i.e., increase in temperature and increase in temperature) an increase in the use of irrigation is necessary to meet crop growing requirements. However, irrigation without proper leaching and drainage leaves salt precipitates in the soil, while over-irrigation favors a gradual rise in the groundwater level, which contribute to soil salt accumulation [4]. In this context, 50% of global cropland areas will be affected by salinization by 2050 [5]. As soil salinity reduces agriculture productivity and quality [6] this projection represents a major socioeconomic threat that can lead to food insecurity and population outmigration [6,7].

The first step to mitigating the soil salinization process and its impacts relies on our ability to monitor soil salinity changes in space and time. However, traditional soil salinity measurements are based on soil field sampling followed by laboratory analysis, which is neither practical nor applicable for large-scale multi-temporal monitoring [8,9]. As a consequence, information on soil salinity is generally only available for specific sites and dates, preventing the study of the role of agricultural activity and climate variability on changes in soil salinity.

1.2. Remote Sensing and Machine Learning Opportunities for Monitoring Soil Salinity

Numerous studies have addressed the potential of using satellite images to retrieve soil salinity estimates at the regional scale [10–13]. Actually, satellite images have gradually replaced the conventional method for monitoring regional soil salinity [14,15]. Freely available, with 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 for this purpose. Soil salinity estimates are quantified using the interaction between soil reflectance (i.e., Sentinel-2) or polarization (i.e., Sentinel-1) and soil salinity content [16,17].

With the rapid growth of artificial intelligence, this relation is generally established through machine learning models such as Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART) [18], Artificial Neural Network (ANN), Decision Tree (DT) [19], Convolutional Neural Network (CNN) [20], Multi Linear Regression (MLR) [21], k-Nearest Neighbor (KNN) [22] or Partial Least Squares Regression (PLSR) [23]. The differences in the machine learning models' structures implies different degrees of reliability in the soil salinity estimates derived from the models. In this context, numerous studies have assessed the sensitivity of soil salinity estimates to different machine learning models [18,20]. Generally, RF and SVM models provide the most reliable soil salinity estimates [18,20,22,24].

1.3. Complementarity of Optical and Radar Images

The variables used as inputs also affect the soil salinity estimate reliability. In this context, several studies have assessed models' sensitivity to different sets of variables derived from S2 optical or S1 radar images for soil salinity mapping. S2 variables, such as single-band values, salinity and vegetation indices, were used as inputs in machine learning models to estimate soil salinity in different regions such as China [20], Algeria [17] and Iran [21]. Despite the reliable soil salinity maps obtained in these studies, the presence of clouds that prevent S2 soil surface observations for certain regions and/or time periods is a major constraint, and different approaches are used to tackle this problem [25,26]. As the radar signal is not affected by cloud cover, the potential of S1 variables such as polarization and texture indices was assessed for soil salinity mapping in different countries such as Iran [27], Vietnam [28] and India [29,30]. S1 radar has the advantage of being sensitive to soil humidity [31–34], which allows us to highlight regions where salt is expected to accumulate due to water accumulation and evaporation [13]. Finally, as S2 and S1 provide observations in different spectral ranges, some authors combine both S2 and S1 to take advantage of both satellite properties [35,36].

1.4. Study Objectives

Whatever the variables used as inputs in the modeling process (S2 and/or S1), similar degrees of confidence are observed in the derived soil salinity maps. However, comparisons are biased by (i) the differences in the study site properties (i.e., soil, climate and topography features), (ii) the machine learning models considered and (iii) the number of samples used for their training. In this context, a stable machine learning set-up based on the RF model is applied to assess soil salinity estimate sensitivity to (i) only S1 variables, (ii) only S2 variables and (ii) both S1 and S2 variables to highlight the respective benefits of S1 and S2 as well as their synergistic use for soil salinity mapping. Finally, annual soil salinity maps corresponding to the pre-sowing period (June to September) are produced for the 2019–2023

period (5 years) in order to highlight the benefit of applying the proposed method for soil salinity monitoring in order to guide mitigation actions to tackle the adverse effects of soil salinization on agricultural productivity. The study site is located in the Bolivian Altiplano, a remote endoreic region that suffers from the major problem of soil salinization, threatening the sustainability of agriculture activity [13].

2. Materials

2.1. Study Area

Lake Poopó is located in the central Bolivian Altiplano region at an average elevation of 3700 m.a.s.l (Figure 1). The climate around Poopó Lake is arid, with the annual potential evapotranspiration rate and precipitation estimated at 1700 mm [37] and 400 mm [38], respectively. The only outflow of Lake Poopó is the Lakajawira River at the southern end of the lake, which rarely flows towards the Coipasa Salar. In the last 50 years, this river only flowed once in 1986 [39]. Therefore, Lake Poopó is now considered to be the terminal point of the endorheic TDPS system, with TDPS standing for Lake Titicaca, Desaguadero River, Lake Poopó and Coipasa Salar (meaning salt pan). The endorheic and arid context, along with the very flat topography and the proximity of the groundwater level, favor soil surface salt accumulation. As a result, the Lake Poopó region holds extremely saline soils [4,13], and the salinization process is expected to worsen over time due to the ongoing desertification induced by climate change and human activity (mining, irrigation) [40,41]. Currently, irrigation plays a vital role in maintaining agricultural activity (Figure 1e). Irrigation contributes (i) to the regional desertification [40,41] and (ii) to soil salt accumulation through the gradual rise in the groundwater level and the irrigation water evaporation leaving salt precipitates.
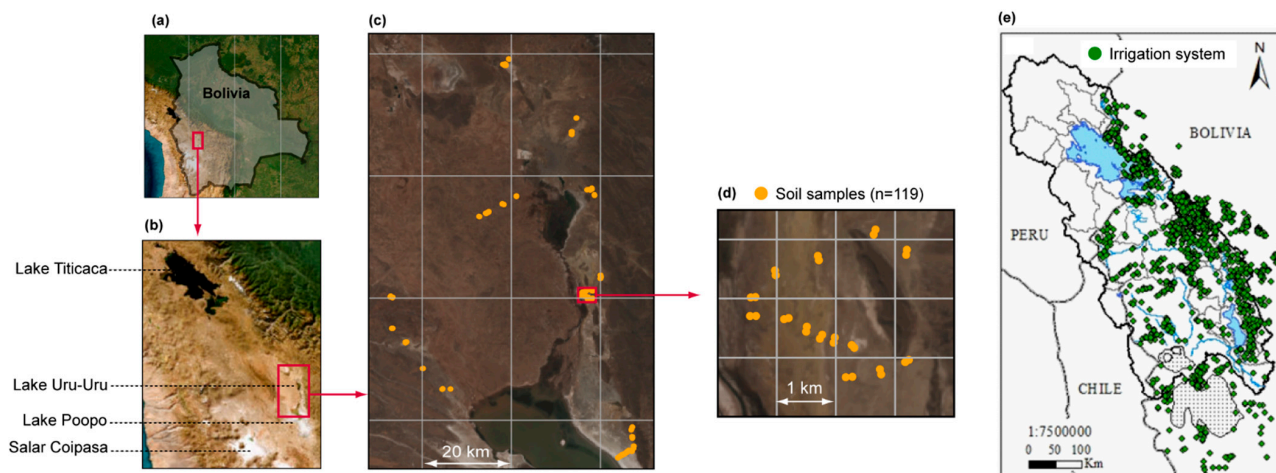

Figure 1. (**a**) Study area located in Bolivia, (**b**) the Titicaca–Desaguadero–Poopó–Salar endorheic system, (**c**,**d**) the location of the 119 soil samples around Lake Poopó, and (**e**) the irrigation system location (obtained from [42]). Subplot (**c**,**d**) shows the mean Sentinel RGB composition obtained between 6 and 11 August in 2022.

2.2. Soil Sample Collection and Analysis

Soil samples (n = 119) were collected during two field campaigns carried out in July and August 2022 (Figure 1c). In July, soil samples were collected on days 17, 18 and 19 (n = 73) whereas in August, soil samples were collected on days 10, 11 and 12 (n = 42).

Each sample was made up from 5 subsamples, each corresponding to a piece of soil of 20 × 20 cm with a depth of 10 cm, extracted at the corners and the center of a mesh of 10 × 10 m cells to take into account potential soil salinity heterogeneity [24].

Soil electrical conductivity (EC) of the 119 samples was measured in the IIQ-UMSA laboratory (Instituto de Investigaciones Químicas, Universidad Mayor de San Andrés, La Paz, Bolivia) as a proxy for soil salinity using the standardized method described in [13].

2.3. Sentinel-2 Images and Pre-Processing

Sentinel-2 images (S2) contain reflectance values in different bands at resolutions of 10, 20 and 60 m in the visible-to-infra-red spectral range. S2 images are available before and after the application of an atmospheric correction process. As the product level 2A (Bottom of Atmosphere—BOA) already includes an atmospheric correction process, it was used in this study. S2 images were obtained from the Copernicus website.

Since soil salinity is not expected to change over short periods of time under stable, dry weather conditions, the reflectance values from S2 images that cover the study area, captured on 6 and 11 August, were averaged (i) to minimize potential artifacts in S2 images and (ii) to ensure temporal matching between field and S2 observations.

Before the averaging process, S2 images were mosaicked and resampled to 20 m using the nearest neighbor resampling method available in the Sentinel Application Platform (SNAP) software version 10.0. The choice of 20 m allowed us to minimize resampling interference on S2 reflectance as there were more bands of interest at 20 m (B5, B6, B7, B8A, B11, B12) than at 10 m (B2, B3, B4, B8).

It is worth mentioning that the metadata of each of the S2 images used indicate cloud coverage of less than 5%, with an absence of cloud over the studied area.

2.4. Sentinel-1 SAR Images and Pre-Processing

Sentinel-1 (S1) operate at the C-band with a central frequency of 5.405 GHz and an incidence angle of 20–45 degrees, acquiring images in dual polarization mode: vertical (VV) and horizontal (VH). The Ground Range Detected (GRD) product was chosen over the Single-Look Complex (SLC). While the SLC provides both the phase and amplitude information of the signal, GRD provides intensity images. These images are easier to interpret as they directly show surface characteristics. The GRD products were obtained from the Earth Data website. S1 images were obtained before, during and after field campaigns 1 and 2, captured on 12 and 24 July, and 5 and 17 August, respectively. S1 images were pre-processed using SNAP software following 5 steps [30,43]:

- Application of the Orbit File to correct the image geometry;
- Application of a radiometric calibration to convert the backscatter intensity into a normalized radar cross-section;
- Application of the refined Lee filter to reduce the influence of incoherent noise (speckle) and to preserve the edges, linear features, point targets and texture information on the SAR imagery [43];
- Application of Range–Doppler correction to correct the inherent geometrical distortions and to geocode the image using a digital elevation model (DEM);
- Application of incidence angle correction for backscattering coefficients [30].

The result was a set of four geo-referenced intensity images converted to Sigma Nought backscatter coefficients ranging from 0 to 1, where Sigma Nought is a normalized dimensionless number that compares the observed intensity to that expected from an area of one square meter. Finally, the four processed image polarizations (VV and VH) were averaged in order to (i) minimize potential artifacts in S1 images and (ii) ensure temporal matching between field and S1 observations.

3. Methods

3.1. Elaboration of the Learning Database

Although the pre-processed S1 and S2 images had the same spatial resolution of 20 m, it was necessary to align the geographic grid to ensure the spatial coincidence of the centroids of the S1 and S2 pixels. This process was carried out using the collocate tool available in the SNAP software, which resamples the S1 image using the S2 image as a reference. Then, S1 polarization (VV and VH) and S2 spectral reflectance (B2, B3, B4, B5, B6, B7, B8, B8a, B11 and B12) were extracted from the pixels where a soil sample was collected. At this stage, the learning database included 119 EC observations with corresponding S1 polarization and S2 spectral reflectance values.

The efficiency of machine learning models for soil salinity mapping is sensitive to the training set size, with larger training sets resulting in better model performance [13,44,45]. In this context, the method presented by Sirpa-Poma [13] was used to increase the learning database observations and therefore modeling reliability (Figure 2). This method consists of adding the pixel information (S1 polarization and S2 spectral reflectance) of the eight neighboring pixels of each of the sampled pixels to the learning database. In this process, the EC observed at the central pixel is attributed to the eight neighboring pixel, whereas S1 polarization and S2 spectral reflectance are the ones observed at the respective locations of the eight neighboring pixels [13]. To guarantee consistency in the extrapolation of the observed EC to neighboring pixels, EC classes have to be considered rather than numeric values [13]. A total of nine EC classes (0–2; 2–4; 4–6; 6–8; 8–10; 10–12; 12–14; 14–16; >16 mS/cm) were defined and used as a rearrangement of the Food and Agriculture Organization (FAO) classification. This method increased the original database size from 119 to 863 observations (Figure 2). According to the described method, there should be 873 (97 × 9) observations in the learning database. Nevertheless, when two samples were very close to each other, one of them was discarded to avoid sharing neighboring pixels.
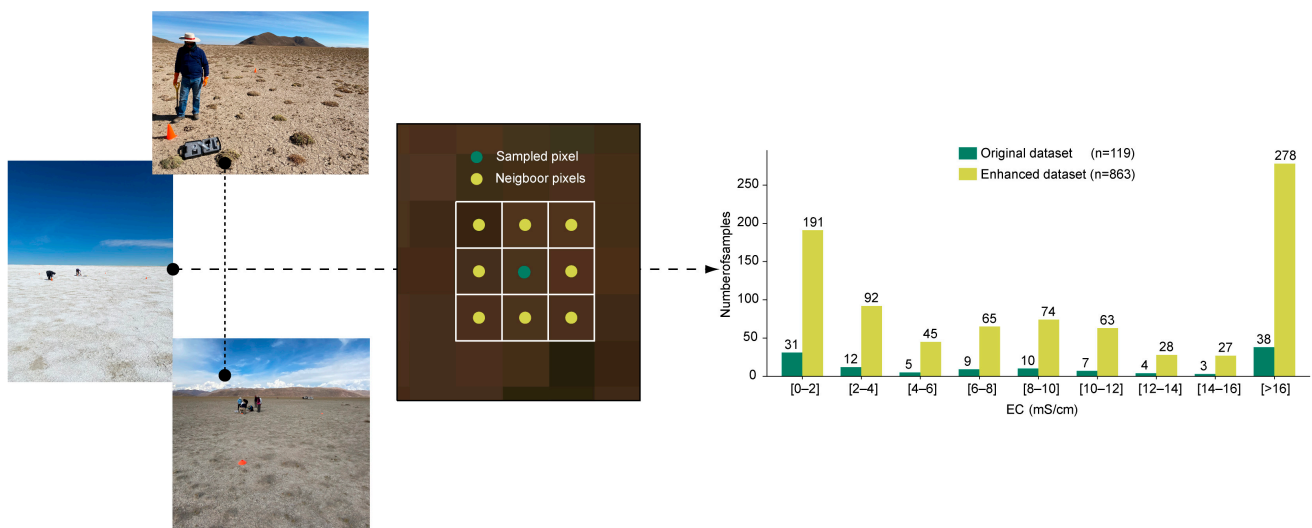

Figure 2. Method used to increase the learning database (adapted from [13]).

Finally, different polarization/textures and spectral indices commonly used for soil salinity mapping were derived from S1 and S2 information and added to the learning database [35,36,46]. From the S2 reflectance values, 6 vegetation indices and 10 salinity indices along with the Normalized Difference Snow Index (NDSI) and the Tasseled Cap Wetness (TCW) were considered (Table 1), whereas 8 polarization indices and 20 texture indices were derived from S1 polarizations (Table 2). Note that S1 texture indices were derived from the Gray Level Co-occurrence Matrix (GLCM) tool available in SNAP software.

Table 1. S2 features used in the modeling process.

Index/Acronym	Definition	Reference
Sentinel-2 bands	B2-Blue, B3-Green, B4-Red, B5-Rededge1, B6-Rededge2, B7-Rededge3, B8-NIR, B8a-Rededge4, B11-SWIR1, B12-SWIR2	
Normalized Difference Snow Index (NDSI)	$(B4 - B8A)/(B4 + B8A)$	[8]
Tasseled cap wetness (TCW)	$0.1509 \times B2 + 0.1973 \times B3$ $+0.3272 \times B4$ $+0.3406 \times B8$ $-0.7112 \times B11$ $-0.4573 \times B12$	[47]

Table 1. *Cont.*

Index/Acronym	Definition	Reference
Salinity Index 1 (SI)	$\sqrt{B2 \times B4}$	[48,49]
Salinity Index 2 (SI1)	$\sqrt{B3 \times B4}$	[48,49]
Salinity Index 3 (SI2)	$\sqrt{B3^2 + B4^2 + B8A^2}$	[48,49]
Salinity Index 4 (SI3)	$\sqrt{B3^2 + B4^2}$	[48,49]
Salinity Index 5 (S)	$B4/B8A$	[50]
Salinity Index (S1)	$B2/B4$	[8]
Salinity Index (S2)	$(B2 - B4)/(B2 + B4)$	[8]
Salinity Index (S3)	$B3 \times B4/B2$	[8]
Salinity Index (S5)	$B2 \times B4/B3$	[8]
Salinity Index (S6)	$B4 \times B8A/B3$	[8]
Normalized Difference Vegetation Index (NDVI)	$(B8A - B4)/(B4 + B8A)$	[51]
Normalized Difference Vegetation Index red-edge 1 (NDVIre1)	$(B8A - B5)/(B5 + B8A)$	[52]
Normalized Difference Vegetation Index red-edge 2 (NDVIre2)	$(B8A - B6)/(B6 + B8A)$	[52]
Normalized Difference Vegetation Index red-edge 2 (NDVIre3)	$(B8A - B7)/(B7 + B8A)$	[53]
Renormalized Difference Vegetation Index (RDVI)	$(B8A - B4)/\left(\sqrt{B8A + B4}\right)$	[54]
Weighted Difference Vegetation Index (WDVI)	$B8A - 0.5 \times B4$	[55,56]

Table 2. S1 features used in the modeling process.

Index/Acronym	Definition	Reference
Sentinel-1 Polarization	VV and VH	[57]
VVaddVH	VV + VH	[57]
VV2addVH	$VV^2 + VH$	[57]
VH2diffVV	$VH^2 - VV$	[57]
VV2addVH2	$VV^2 + VH^2$	[57]
VV2addVH2divVH	$\left(VV^2 + VH^2\right)/VH$	[16]
10logVV	$10logVV$	[16]
10logVH	$10logVH$	[16]
10logVVaddVH	$10(logVV + logVH)$	[16]
VV_GLCM	VV_Contrast, VV_Dissimilarity, VV_Homogeneit, VV_AngularSecondMoment, VV_Energy, VV_Entropy, VV_MaximumProbability, VV_Correlation, VV_Mean and VV_StandardDeviation, VV_Variance	
VH_GLCM	VH_Contrast, VH_Dissimilarity, VH_Homogeneity, VH_AngularSecondMoment, VH_Energy, VH_Entropy, VH_MaximumProbability, VH_Correlation, VH_Mean and VH_StandardDeviation, VH_Variance	

3.2. Machine Learning Modeling Set-Up

In order to highlight S1 and S2 ability for soil salinity mapping, S1 and S2 features were used separately as independent variables in the RF model. Note that the RF model was selected for the present study as it provides more reliable soil salinity estimates than other machine learning models [13,18,20,22,24]. Then, the S1 and S2 variables were used together as independent variables in order to assess their complementarity for soil salinity mapping.

Multi-collinearity in the independent variables drives machine learning models to be unstable due to the redundancy of the input variables [22,24]. Among feature selection tools, the Genetic Algorithm (GA) and/or VIF are generally used for the removal of irrelevant and redundant features, in order to improve the performance of machine learning

models [24,58]. A recent study showed that the use of GA led to more reliable soil salinity estimates than VIF [13], and therefore, GA is used in this study.

Additionally, the consideration of an unbalanced learning dataset (i.e., different numbers of observations for the considered classes) may lead to unstable model prediction. Actually, the models' efficiency may increase (decrease) for classes with more (less) observations [59,60]. As observed in Figure 2, the learning database is unbalanced. Indeed, classes [0–2] and [>16] have the highest numbers of observations (n = 191 and n = 278, respectively), whereas classes [12–14] and [14–16] have the lowest numbers of observations (n = 28 and n = 27, respectively). In this context, whereas observations for one class may be much less frequent than for the others, non-equal weights for classes are especially useful in classifiers when addressing class imbalance by assigning different weights to each class [60]. This means that classes that are less represented in the dataset will receive a higher weight, allowing the model to pay more attention to these minority classes. When the class-weight parameter in RF is set to 'balanced', the weights are adjusted inversely proportionally to the frequency of each class, and the dataset is now considered to be balanced.

In the above-described context, six scenarios are considered (Figure 3):

- Scenario-1: applying GA to S2 variables with unbalanced dataset;
- Scenario-2: applying GA to S2 variables with balanced dataset;
- Scenario-3: applying GA to S1 variables with unbalanced dataset;
- Scenario-4: applying GA to S1 variables with balanced dataset;
- Scenario-5: applying GA to S1 and S2 variables with unbalanced dataset;
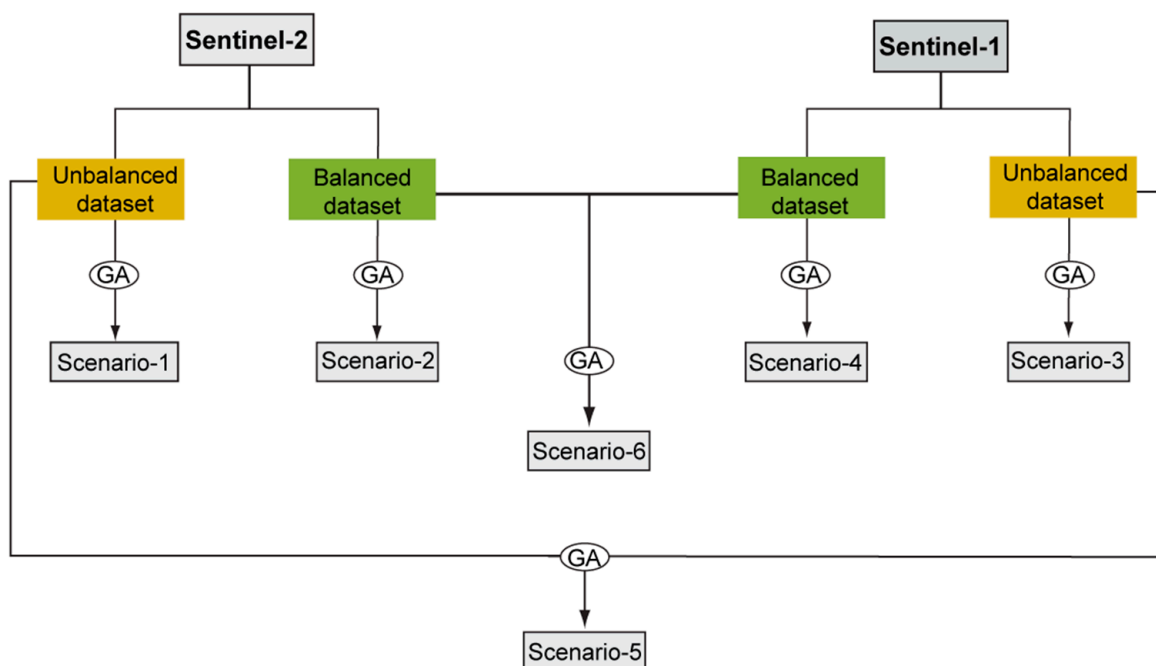- Scenario-6: applying GA to S1 and S2 variables with balanced dataset.

Figure 3. Flow-chart of the modeling set-up with the five scenarios.

For each scenario, the learning database was randomly split into a training set (70%) and a validation set (30%). It is noteworthy that the k-fold cross-validation method (3-fold) was used during the training step along with the GA.

The relevance of the features selected by the GA was assessed through the feature and permutation importance scores. The feature importance (FI) indicates the relative contribution of each feature for the model training (ranging from 0 to 1), whereas the permutation importance (PI) indicates the trained model's sensitivity to each feature (ranging from 0 to 1). The PI was obtained through a two-step process. First, the trained model was run with the validation dataset to determine the model's overall accuracy, (OA) and secondly, the

values of each variable were successively randomly shuffled before running the model with the modified validation dataset. The OAs obtained before and after the random shuffling were compared (subtraction) to obtain the PI. Therefore, the PI shows the drop in model accuracy in the case of inaccurate values for each one of the variables independently.

Finally, the model accuracy was assessed considering the confusion matrix and associated recall, precision, F1 score, and overall accuracy statistical indicators (Equations (1)–(4)). All statistical indicators are in the range 0 to 1, where 1 is the best value for a classifier.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{F1score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{3}$$

$$\text{OA} = \frac{TP + TN}{TN + TP + FN + FP} \tag{4}$$

where TP, FP, FN and TN are true positive, false positive, false negative, and true negative respectively.

3.3. Soil Salinity Spatio-Temporal Monitoring Using Remote Sensing Data: A Relevant Tool for the Sustainable Adaptation of Agriculture?

As soil salinity is a well know factor affecting agriculture yields [5], this section proposes elaborating soil salinity maps for the 2019–2023 pre-sowing period (one for each year) using the most efficient machine learning set-up previously highlighted (i.e., scenario-6). The maps are elaborated with S2 and S1 composite images. Composite images consist of the average observation made by all available images during a specific period. This approach aims to ensure that all pixels include at least one observation. However, the spectral response of a specific pixel is variable over time (i.e., clouds, vegetation growing, soil humidity variability due to climate and/or irrigation, land use, land cover modification due to harvest and plowing, etc.). Therefore, spatial inconsistency in the composite image can be present as the available spectral information dates may be different according to the considered pixel. To minimize such inconsistency, S1 and S2 composite images were derived from images registered between June and September. This period corresponds to the pre-sowing and dry seasons with (i) stable climate conditions (i.e., almost no precipitation) and (ii) no agriculture practices (i.e., sowing, plowing, harvesting, irrigation). It is worth mentioning that the pixels identified as cloud covered by the QA_PIXEL band do not take into account in the averaging process. In this condition, stable S1 and S2 signals are expected across the considered region and period.

For the final mapping, pixels identified as covered by water by the Automated Water Extraction Index (AWEI) were flagged to compare the exact same area for all the years under consideration. Indeed, for years with a very small extent of Lake Poopó, extreme saline conditions were detected at the lake shore pixels, whereas for years with a large Lake Poopó extent, these pixels were "inundated" and no salt concentration could be retrieved. By masking water-covered pixels, the overestimation of soil salinity condition was avoided for years with a small lake area when compared to years with a large lake extent.

4. Results

4.1. Feature Selection

Figure 4 shows the 10 most contributing features for the training and validation steps, as expressed by the FI and PI, respectively.
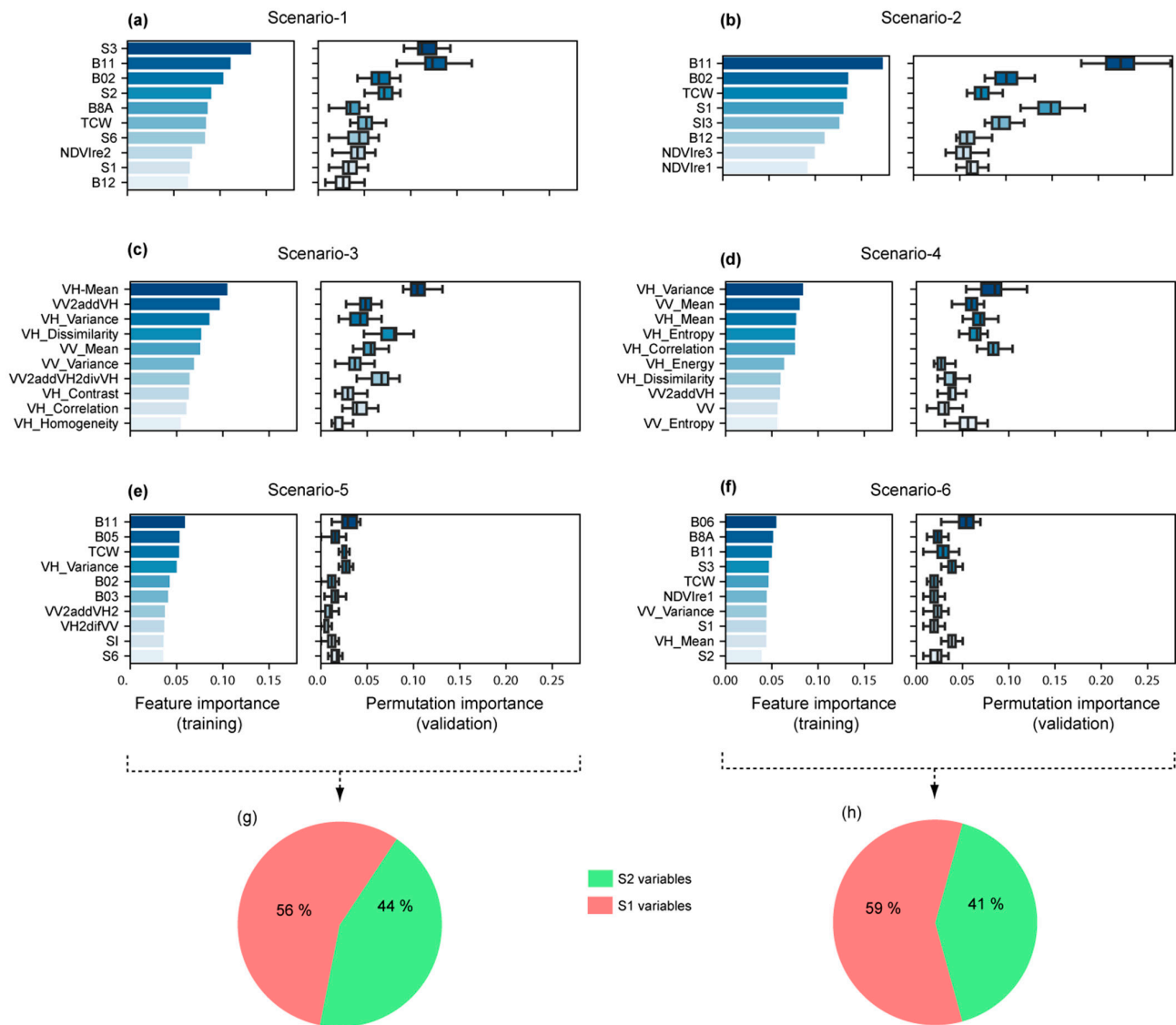
Figure 4. Feature contribution in the training and validation steps of each scenario expressed in the forms of feature importance and permutation importance (PI), respectively (**a**–**f**), along with S1 and S2 variable ratios selected by GA for scenarios-5 and -6 (**g**,**h**).

Five out of eight variables from scenario-2 (B02, B11, B12, TCW, S1) are in the top 10 most contributing variables of scenario-1 for the training step (i.e., feature importance). Additionally, scenarios-1 and -2 (i.e., applying the GA to S2 variables with the balanced and unbalanced datasets, respectively) show the highest sensitivity to B11 in the validation step, with PI values of 0.13 and 0.23, respectively. Therefore, when considering only S2 variables, balanced and unbalanced datasets are sensitive to similar variables.

Scenarios-3 and -4 (i.e., applying GA to S1 variables with the balanced and unbalanced datasets, respectively) share 60% of variables contributing the most to the training step (VH_Mean, VV2addVH, VH_Variance, VH_dissimilary, VV_Mean, VH_Correlation). Additionally, in the validation step, scenario-3 (scenario-4) is more (highly) sensitive to VH_Mean, with PI values of 0.11 and 0.07 for scenarios-3 and -4, respectively. As for scenarios-1 and -2, the balanced and unbalanced datasets are sensitive to similar variables.

Scenarios-5 and -6 (i.e., applying GA to both S1 and S2 variables with the balanced and unbalanced datasets, respectively) only share 20% of their top 10 most contributing variables for the model training (B11, TCW). However, for both scenarios, the FI of all variables is very low (<0.05), suggesting low sensitivity to individual variables. The same

occurs for the validation step with very low PI values (<0.05) for all the variables. Therefore, when considering both S1 and S2 variables, the reliability of the considered scenarios (balanced and unbalanced) is more sensitive to variable complementarity (than individual variables), indicating a complementarity of optical (S2) and radar (S1) properties for soil salinity estimates. This behavior is observable in Figure 4g,h, with 56% and 44% (59% and 41%) of the variables selected by the GA derived from S1 and S2, respectively, for scenario-5 (scenario-6).

4.2. Benefits of the Combination of Sentinel-1 and Sentinel-2

Figure 5 shows the results obtained in the validation step for the different scenarios. For both the unbalanced and balanced datasets, the use of S1 and S2 variables together as predictive features led to the most reliable soil salinity estimates. For example, when considering a balanced dataset, the four statistical indices yield higher values for scenario-6 (i.e., S1 + S2) than for scenarios-2 (S2) and -4 (S1). The OA of 0.95 obtained in scenario-6 corresponds to increases of 8% and 14.5% when compared to the OAs reached by scenarios-4 (0.88) and -2 (0.83). This observation is consistent with a study carried out in China that reached more accurate soil salinity estimates by combining S2 and S1 features as predictive variables than by considering only S2 features [36].
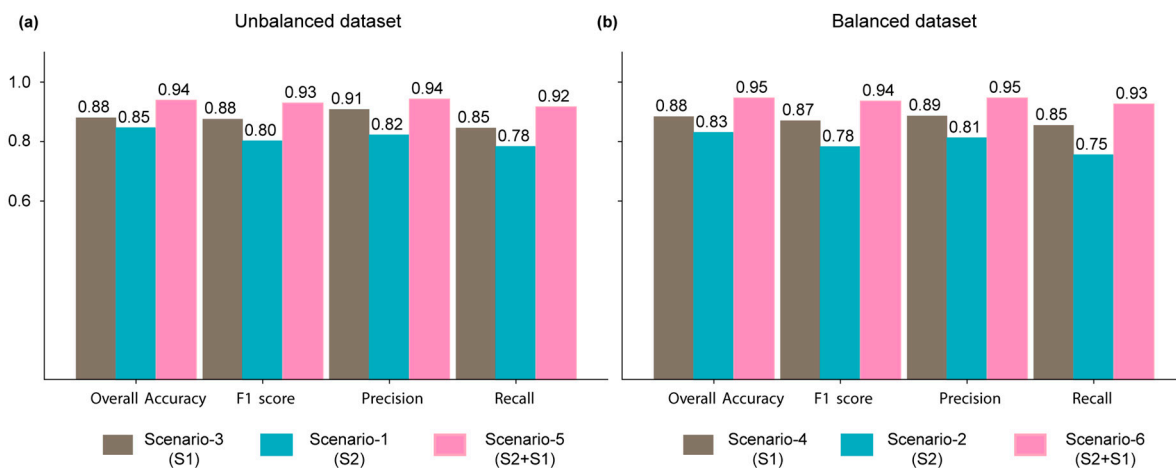

Figure 5. Models' performance for the validation step expressed in terms of OA, F1-score, precision and recall with (**a**) unbalanced and (**b**) balanced datasets.

Interestingly, the soil salinity estimates derived from S1 features only are systematically more reliable than the estimates obtained from S2 features alone. Actually, when considering an unbalanced dataset (balanced dataset), scenario-3 (scenario-4) reaches higher OA, F1-score, precision and recall values than scenario-1 (scenario-2). Therefore, even for regions or periods not affected by frequent cloud cover preventing the use of S2 images, soil salinity estimates are better retrieved from S1 images.

Finally, the use of balanced and unbalanced datasets led to an overall similar level of accuracy for the soil salinity estimates regardless of the considered scenarios (Figure 5).

Figure 6 shows the confusion matrix obtained for scenarios-5 and -6 (using both S1 and S2) to provide more insight in the most efficient scenarios. The results from both scenarios are very similar for all classes, except for classes (6–8) and (8–10), which show the highest discrepancies (Figure 5). Scenario-6 presents higher OA for the (6–8) class, whereas scenario-5 does for the (8–10) class. Actually, scenario-5 tends to underestimate the EC values for the (6–8) class, whereas scenario-6 does for the (8–10) class (Figure 6).
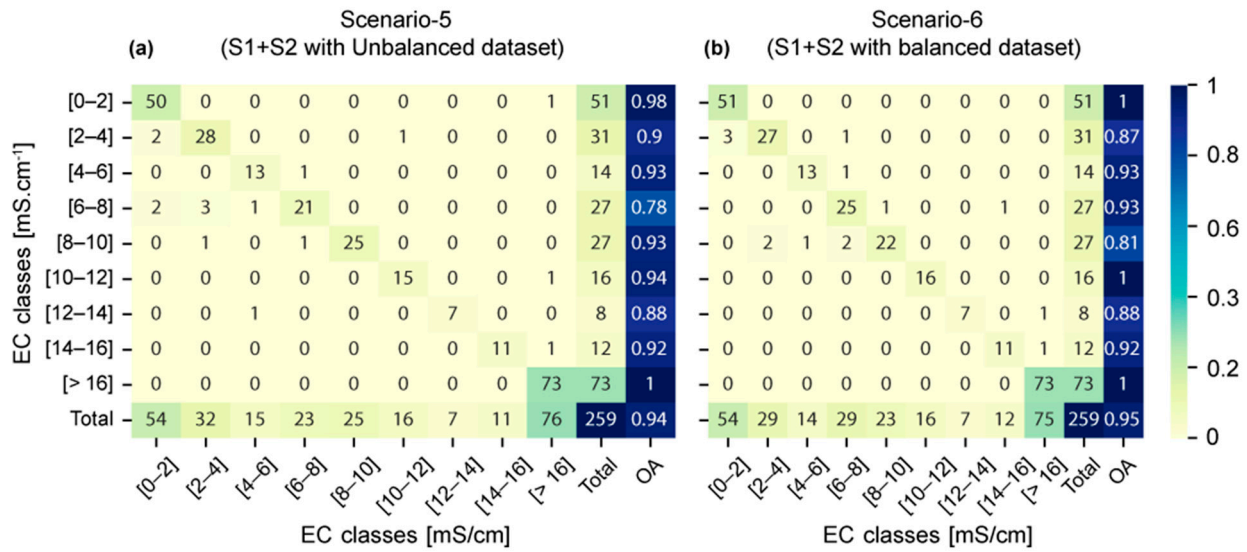
Figure 6. Confusion matrices with OA scores obtained for the most efficient scenarios (S1 + S2) using (**a**) unbalanced and (**b**) balanced datasets.

4.3. Sensitivity of Soil Salinity Maps to Unbalanced and Balanced Datasets

Figure 7 shows the soil salinity maps obtained with the most efficient scenarios (scenarios-5 and -6) along with the values derived from the field sample. A previous filter based on AWEI was used to mask pixels concerned with water bodies, with −0.1 as the threshold [40]. Overall, the field observations match the soil salinity maps derived from both scenarios.
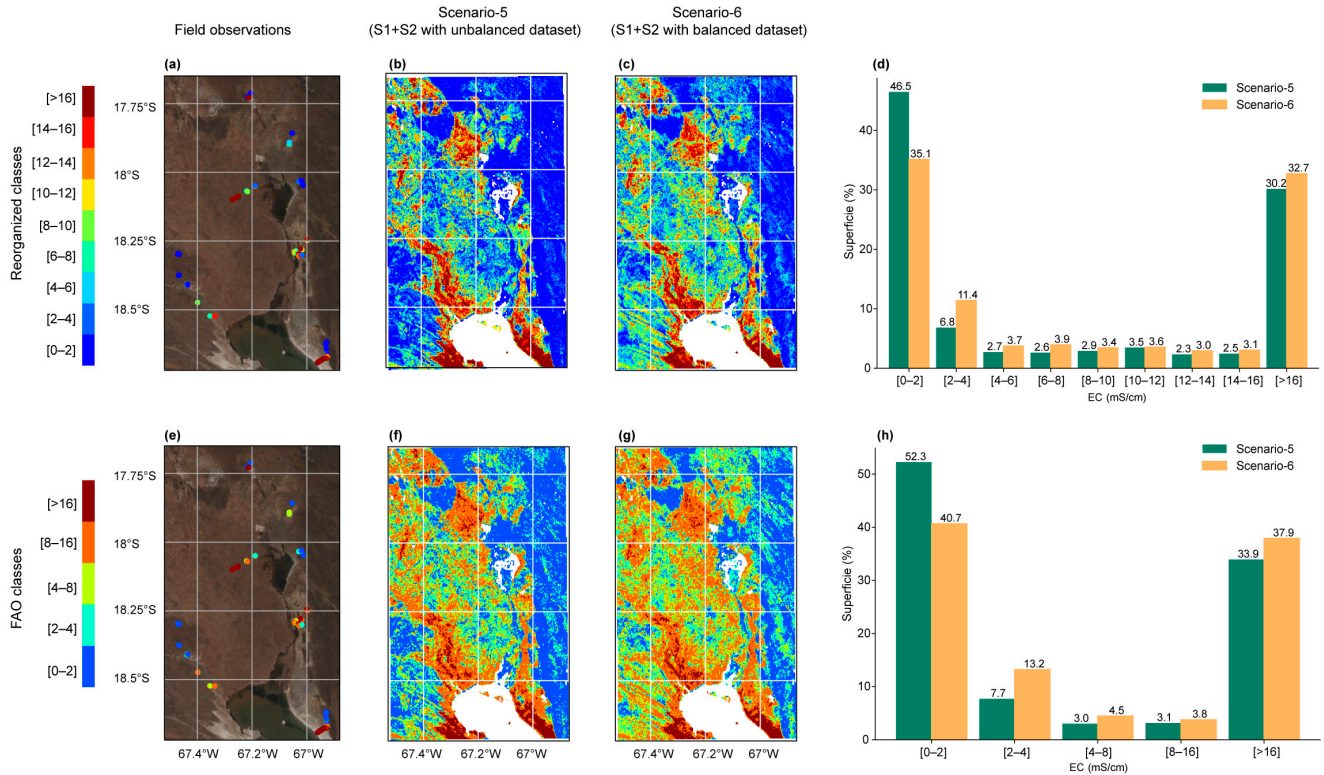

Figure 7. (**a**–**e**) Soil salinity derived from field samples. (**b**,**c**) Salinity map obtained with scenarios-5 and -6 using the reorganized EC classes, with (**d**) the corresponding superficial extent of each class. (**f**,**g**) Salinity map obtained with scenarios-5 and -6 using the FAO EC classes, with (**h**) the corresponding superficial extent of each class.

The soil salinity maps derived from scenarios-5 and -6 show a similar spatial pattern (Figure 7). Lake Poopó shores (east, west and north) are the regions with the highest soil salinity because of the lake shrinkage/evaporation occurring during the dry season (April to September) that concentrates salt at the top soil layer. A similar process occurs along the Desaguadero River's west branch (northwest of Lake Poopó) and for the river's east branch connecting Lake Uru-Uru to Lake Poopó (northeast of Lake Poopó). As the eastern part of the study area is steep and mountainous, favoring water runoff instead of water storage, it shows the lowest soil salinity values. This sub-region is also colder, and thus, potential evapotranspiration is lower, which, in turn, reduces soil humidity losses and salt concentration.

Despite similar soil salinity distribution patterns, the soil salinity estimates from scenario-6 are slightly higher than the estimates derived from scenario-5 (Figure 7d). Indeed, except for the [0–2] EC class, larger areas are observed for all classes of scenario-6 (Figure 7d–h). This difference is more noticeable for the low salinity classes of [2–4], [4–6] and [6–8], where area increases of 68%, 37% and 50% are observed for scenario-6 with respect to scenario-5. In the unbalanced dataset, classes [0–2] and [>16] have the highest number of observations (n = 182 and 278, respectively) (Figure 1e). Even if class [>16] has more observations (n = 278) than class [0–2] (n = 182), class [>16] covers a much wider soil salinity range (from 16.25 to 48.7 mS.cm^{-1}) than class [0–2]. Therefore, with a smaller range of soil salinity, class [0–2] should be considered the class with the most observations. In this sense, a bias towards low salinity estimates (i.e., [0–2] class) is expected from scenario-5. On the other hand, scenario-6, which is trained with a balanced dataset, is not expected to provide soil salinity estimates biased towards any soil salinity concentration class. This difference between scenarios-5 and -6 could explain why scenario-5 underestimates soil salinity when compared to scenario-6.

4.4. Spatio-Temporal Monitoring of Regional Soil Salinity: An Opportunity for Decision Making in Sustainable Agriculture?

Figure 8 shows soil salinity maps obtained for the pre-sowing period between 2019 and 2023 (5 years) using scenario-6. Considerable space and time variation occurs in the regional distribution of soil salinity. For example, on the north-west shore of Lake Poopó, much more pixels with a strongly saline condition (i.e., EC > 8 mS/cm, unsuitable for agriculture activity) are found in 2020, 2022 and 2023 than for the 2019 and 2021 years. Similarly, in the northern Lake Poopó region, much more pixels with a non-saline condition (i.e., EC < 2 mS/cm, favorable for optimum crop development) are found in 2019, 2022 and 2023 than for the 2020 and 2021 years.

In this context, soil salinity maps obtained just before the sowing period could allow us to select lands with minimum salinity concern in order to optimize (minimize) crop yield (losses). In regions such as the Altiplano, this would represent a considerable step towards sustainable development, as agriculture represents a major economic activity. Based on freely available information, this tool provides an efficient alternative to the traditional monitoring system, which is generally out of reach for many remote regions, especially in an unfavorable socio-economic context. Finally, these maps could be overlaid with additional information, such as crop rotation, irrigation practices and climate data, to support the sustainable management of agricultural land and to minimize the potential effects of soil salinity hazards.
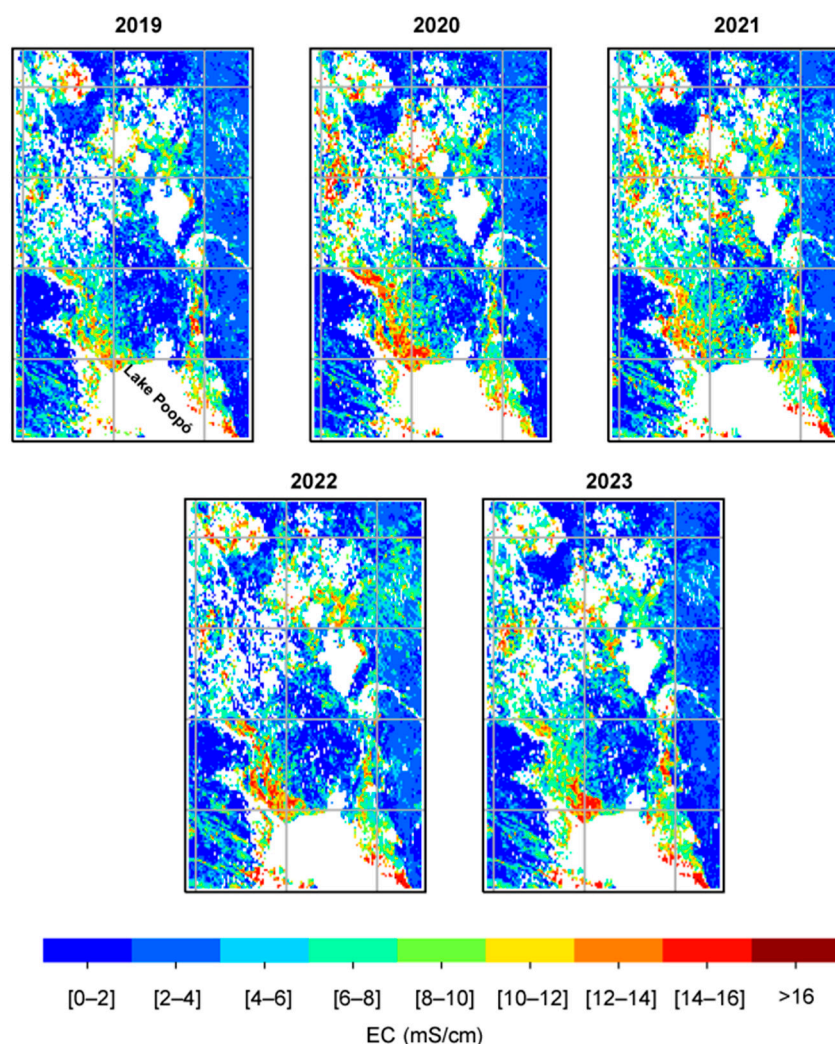
Figure 8. Pre-sowing soil salinity maps derived from scenario-6 for the 2019–2023 period.

5. Discussion

This study focuses on the sensitivity of soil salinity to S2 and S1 in order to evaluate the possibility of regional monitoring based on satellite images. However, additional features controlling soil salinity variations in space and time could have been considered to improve the model outputs.

For example, soil salinity is expected to follow the drainage pattern, as regions where water accumulates are more prone to the concentration of soluble salts at the soil surface due to the water evaporation process [61–65]. In this context, some studies have reported on the benefits of including topography features (i.e., elevation, slope and aspect) derived from freely available digital elevation models as independent variables to improve soil salinity modeling [66,67].

Similarly, soil humidity is another useful parameter to highlight regions where water (and therefore salt) accumulates. In this context, the inclusion of soil moisture information could help to improve soil salinity estimates and spatial distribution determination. Numerous studies have reported on the possibility of retrieving soil humidity estimates from S1 information [31–34]. In this context, soil humidity sensitivity to S1 features can explain the higher performance of scenarios based on S1 features (scenarios-3 and -4) compared to the scenarios based on S2 features (scenarios-1 and -2). Interestingly, the highest performance observed for scenarios-5 and -6 (S1 + S2) can be explained by the inclusion of additional S2 indices related to soil humidity. Indeed, a recent study used S2 soil salinity indices asindependent variables to retrieve soil humidity [68]. In this context,

soil humidity's sensitivity to S2 soil salinity indices improve the soil salinity estimates of scenarios-5 and -6. Finally, both scenarios ranked the TCW indice as the third most contributing variable (Figure 3). This index is also highly related to soil humidity and therefore to soil salinity. As a general conclusion, the combination of S1 and S2 features enhances the model's ability to identify regions likely to accumulate greater quantities of water and therefore salt, which improves the predictions of the models. Furthermore, the consideration of additional topography features (i.e., elevation, slope and aspect) should improve the models' predictions even more.

It is worth mentioning here that S2 and S1 are not only sensitive to soil salinity. Numerous studies have already shown the possibility of retrieving other soil parameters from S1 and S2, such as organic matter [69–72], nitrogen [73–76] and texture [77,78]. In this respect, a model trained across a specific region (i.e., Altiplano) may not be efficient across another region. Indeed, each region would need to assess specific soil parameters (i.e., organic matter, nitrogen and texture). Although the proposed model is trained and adapted to a specific study site (i.e., Altiplano), the method developed in this study can be applied to other regions of the world through a calibration stage specific to the studied site.

6. Conclusions

This study assesses the potential of Sentinel-2 and -1 (S2 and S1) alone and the potential of their combined use for soil salinity mapping. The method used is based on a machine learning set-up considering both unbalanced and balanced training datasets to assess soil salinity sensitivity to the training set characteristics. The main results regarding the proposed procedure can be summarized as follows:

- Higher overall accuracy (OA) is obtained from the combination of S2 and S1 (0.95) than from S1 (0.88) and S2 (0.83) alone, which confirms the complementarity of optical and radar information for soil salinity mapping.
- The use of S1 features alone as regressors leads to more accurate soil salinity estimates than the use of S2 features alone. For regions and/or periods with consistent cloud covers preventing soil surface observations, reliable soil salinity estimates can be obtained with S1 images alone.
- The model trained with an unbalanced dataset tends to provide soil salinity estimates biased towards the EC class with the (relative) highest number of observations (i.e., [0–2] class). Therefore, balanced datasets should be considered rather than unbalanced datasets to obtain consistent soil salinity estimates.

The proposed and validated procedure can be used in the following ways to generate annual soil salinity maps to highlight its potential for supporting decisions on sustainable agriculture:

- In an operative way to provide farmers with pre-sowing soil salinity maps to help with sustainable land selection (land with a low salinity concentration) to improve crop yields.
- In a retrospective way to understand soil salinization dynamics and its links with climate and/or agriculture practices (i.e., irrigation, crop rotation, etc.).

Author Contributions: J.W.S.-P.: Conceptualization, Data Curation, Methodology, Validation, Visualization, Writing—Original Draft. F.S.: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Validation, Visualization, Writing—Original Draft. R.P.Z.: Data Curation, Writing—Review and Editing. E.R.: Conceptualization, Data Curation, Formal Analysis, Resources, Writing—Review and Editing. M.P.-F.: Data curation, Writing—Review and Editing. M.G.F.C.: Data Curation, Writing—Review and Editing. J.M.-C.: Writing—Review and Editing. O.R.: Resources, Writing—Review and Editing. M.-P.B.: Funding Acquisition, Writing—Review and Editing. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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 conflict of interest.

References

1. Shrivastava, P.; Kumar, R. Soil Salinity: A Serious Environmental Issue and Plant Growth Promoting Bacteria as One of the Tools for Its Alleviation. *Saudi J. Biol. Sci.* **2015**, *22*, 123–131. [CrossRef] [PubMed]
2. Singh, A.; Kumar, U.; Seitz, F. Remote Sensing of Storage Fluctuations of Poorly Gauged Reservoirs and State Space Model (SSM)-Based Estimation. *Remote Sens.* **2015**, *7*, 17113–17134. [CrossRef]
3. Thenkabail, P.S. Global Croplands and Their Importance for Water and Food Security in the Twenty-First Century: Towards an Ever Green Revolution That Combines a Second Green Revolution with a Blue Revolution. *Remote Sens.* **2010**, *2*, 2305–2312. [CrossRef]
4. 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]
5. 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]
6. 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] [PubMed]
7. 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]
8. 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. [CrossRef]
9. Mulder, V.L.; de Bruin, S.; Schaepman, M.E.; Mayr, T.R. The Use of Remote Sensing in Soil and Terrain Mapping—A Review. *Geoderma* **2011**, *162*, 1–19. [CrossRef]
10. Erkin, N.; Zhu, L.; Gu, H.; Tusiyiti, A. Method for Predicting Soil Salinity Concentrations in Croplands Based on Machine Learning and Remote Sensing Techniques. *J. Appl. Remote Sens.* **2019**, *13*, 034520. [CrossRef]
11. Farahmand, N.; Sadeghi, V. Estimating Soil Salinity in the Dried Lake Bed of Urmia Lake Using Optical Sentinel-2 Images and Nonlinear Regression Models. *J. Indian Soc. Remote Sens.* **2020**, *48*, 675–687. [CrossRef]
12. Guo, L.; Zhao, S.; Gao, J.; Zhang, H.; Zou, Y.; Xiao, X. A Novel Workflow for Crop Type Mapping with a Time Series of Synthetic Aperture Radar and Optical Images in the Google Earth Engine. *Remote Sens.* **2022**, *14*, 5458. [CrossRef]
13. 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] [PubMed]
14. Davis, E.; Wang, C.; Dow, K. Comparing Sentinel-2 MSI and Landsat 8 OLI in Soil Salinity Detection: A Case Study of Agricultural Lands in Coastal North Carolina. *Int. J. Remote Sens.* **2019**, *40*, 6134–6153. [CrossRef]
15. Elnaggar, A.A.; Noller, J.S. Application of Remote-Sensing Data and Decision-Tree Analysis to Mapping Salt-Affected Soils over Large Areas. *Remote Sens.* **2010**, *2*, 151–165. [CrossRef]
16. Zhang, Z.; Du, Y.; Lao, C.; Yang, N.; Zhou, Y.; Yang, Y. Inversion Model of Soil Salt Content in Different Depths Based on Radar Remote Sensing. *Nongye Jixie Xuebao/Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 243–251. [CrossRef]
17. 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]
18. 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]
19. 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]
20. 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]
21. 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]
22. 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]

23. 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]

24. 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]

25. Singh, P.; Komodakis, N. Cloud-GAN: Cloud Removal for Sentinel-2 Imagery Using a Cyclic Consistent Generative Adversarial Networks. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; Volume 2018-July.

26. Ebel, P.; Meraner, A.; Schmitt, M.; Zhu, X.X. Multisensor Data Fusion for Cloud Removal in Global and All-Season Sentinel-2 Imagery. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 5866–5878. [CrossRef]

27. 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]

28. 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]

29. 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]

30. 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]

31. Huang, S.; Ding, J.; Zou, J.; Liu, B.; Zhang, J.; Chen, W. Soil Moisture Retrival Based on Sentinel-1 Imagery under Sparse Vegetation Coverage. *Sensors* **2019**, *19*, 589. [CrossRef]

32. Gao, Q.; Zribi, M.; Escorihuela, M.J.; Baghdadi, N. Synergetic Use of Sentinel-1 and Sentinel-2 Data for Soil Moisture Mapping at 100 m Resolution. *Sensors* **2017**, *17*, 1966. [CrossRef] [PubMed]

33. Alexakis, D.D.; Mexis, F.-D.K.; Vozinaki, A.-E.K.; Daliakopoulos, I.N.; Tsanis, I.K. Soil Moisture Content Estimation Based on Sentinel-1 and Auxiliary Earth Observation Products. A Hydrological Approach. *Sensors* **2017**, *17*, 1455. [CrossRef] [PubMed]

34. Attarzadeh, R.; Amini, J.; Notarnicola, C.; Greifeneder, F. Synergetic Use of Sentinel-1 and Sentinel-2 Data for Soil Moisture Mapping at Plot Scale. *Remote Sens.* **2018**, *10*, 1285. [CrossRef]

35. 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]

36. 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]

37. 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]

38. Satgé, F.; Bonnet, M.-P.; Gosset, M.; Molina, J.; Hernan Yuque Lima, W.; Pillco Zolá, R.; Timouk, F.; Garnier, J. Assessment of Satellite Rainfall Products over the Andean Plateau. *Atmos. Res.* **2016**, *167*, 1–14. [CrossRef]

39. Pillco Zolá, R.; Bengtsson, L. Long-Term and Extreme Water Level Variations of the Shallow Lake Poopó, Bolivia. *Hydrol. Sci. J.* **2006**, *51*, 98–114. [CrossRef]

40. 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]

41. Satgé, F.; Hussain, Y.; Xavier, A.; Zolá, R.P.; Salles, L.; Timouk, F.; Seyler, F.; Garnier, J.; Frappart, F.; Bonnet, M.-P. Unraveling the Impacts of Droughts and Agricultural Intensification on the Altiplano Water Resources. *Agric. For. Meteorol.* **2019**, *279*, 107710. [CrossRef]

42. Canedo, C.; Zolá, R.P.; Berndtsson, R. Role of Hydrological Studies for the Development of the TDPS System. *Water* **2016**, *8*, 144. [CrossRef]

43. Qiu, J.; Crow, W.T.; Wagner, W.; Zhao, T. Effect of Vegetation Index Choice on Soil Moisture Retrievals via the Synergistic Use of Synthetic Aperture Radar and Optical Remote Sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *80*, 47–57. [CrossRef]

44. Foody, G.M.; Mathur, A.; Sanchez-Hernandez, C.; Boyd, D.S. Training Set Size Requirements for the Classification of a Specific Class. *Remote Sens. Environ.* **2006**, *104*, 1–14. [CrossRef]

45. Ramezan, C.A.; Warner, T.A.; Maxwell, A.E.; Price, B.S. Effects of Training Set Size on Supervised Machine-Learning Land-Cover Classification of Large-Area High-Resolution Remotely Sensed Data. *Remote Sens.* **2021**, *13*, 368. [CrossRef]

46. Yin, H.; Chen, C.; He, Y.; Jia, J.; Chen, Y.; Du, R.; Xiang, R.; Zhang, X.; Zhang, Z. Synergistic Estimation of Soil Salinity Based on Sentinel-1 Image Texture and Sentinel-2 Salinity Spectral Indices. *J. Appl. Remote Sens.* **2023**, *17*, 018502. [CrossRef]

47. Baig, M.H.A.; Zhang, L.; Shuai, T.; Tong, Q. Derivation of a Tasselled Cap Transformation Based on Landsat 8 At-Satellite Reflectance. *Remote Sens. Lett.* **2014**, *5*, 423–431. [CrossRef]

48. Rahmati, M.; Mohammadi-Oskooei, M.; Neyshabouri, M.R.; Fakheri-Fard, A.; Ahmadi, A.; Walker, J. ETM+ Data Applicability for Remote Sensing of Soil Salinity in Lighvan Watershed, Northwest of Iran. *Curr. Opin. Agric.* **2014**, *3*, 10.

49. El Kader Douaoui, A.; 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]

50. 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*, 41262. [CrossRef]

51. Cho, K.H.; Beon, M.-S.; Jeong, J.-C. Dynamics of Soil Salinity and Vegetation in a Reclaimed Area in Saemangeum, Republic of Korea. *Geoderma* **2018**, *321*, 42–51. [CrossRef]

52. Qiu, Y.; Chen, C.; Han, J.; Wang, X.; Wei, S.; Zhang, Z. Satellite Remote Sensing Estimation Model of Soil Salinity in Jiefangzha Irrigation under Vegetation Coverage. *Water Sav. Irrig.* **2019**, *44*, 108–112.

53. Fernández-Manso, A.; Fernández-Manso, O.; Quintano, C. SENTINEL-2A Red-Edge Spectral Indices Suitability for Discriminating Burn Severity. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 170–175. [CrossRef]

54. Fan, X.; Pedroli, B.; Liu, G.; Liu, Q.; Liu, H.; Shu, L. Soil Salinity Development in the Yellow River Delta in Relation to Groundwater Dynamics. *Land Degrad. Dev.* **2012**, *23*, 175–189. [CrossRef]

55. Clevers, J.G.P.W. Application of the WDVI in Estimating LAI at the Generative Stage of Barley. *ISPRS J. Photogramm. Remote Sens.* **1991**, *46*, 37–47. [CrossRef]

56. Clevers, J.G.P.W. The Derivation of a Simplified Reflectance Model for the Estimation of Leaf Area Index. *Remote Sens. Environ.* **1988**, *25*, 53–69. [CrossRef]

57. Ma, C. Quantitative Retrieval of Soil Salt Content Based on Sentinel-1 Dual Polarization Radar Image. *Nongye Gongcheng Xuebao/Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 153–158. [CrossRef]

58. Taghadosi, M.M.; Hasanlou, M.; Eftekhari, K. Retrieval of Soil Salinity from Sentinel-2 Multispectral Imagery. *Eur. J. Remote Sens.* **2019**, *52*, 138–154. [CrossRef]

59. More, A.S.; Rana, D.P. Review of Random Forest Classification Techniques to Resolve Data Imbalance. In Proceedings of the Proceedings—1st International Conference on Intelligent Systems and Information Management, ICISIM 2017, Aurangabad, India, 5–6 October 2017; Volume 2017-January.

60. Shahhosseini, M.; Hu, G. Improved Weighted Random Forest for Classification Problems. In Proceedings of the 4th International Online Conference on Intelligent Decision Science (IDS 2020), Istanbul, Turkey, 7–8 August 2020; Springer: Cham, Switzerland, 2021.

61. Gerardo, R.; de Lima, I.P. Sentinel-2 Satellite Imagery-Based Assessment of Soil Salinity in Irrigated Rice Fields in Portugal. *Agriculture* **2022**, *12*, 1490. [CrossRef]

62. Ren, D.; Wei, B.; Xu, X.; Engel, B.; Li, G.; Huang, Q.; Xiong, Y.; Huang, G. Analyzing Spatiotemporal Characteristics of Soil Salinity in Arid Irrigated Agro-Ecosystems Using Integrated Approaches. *Geoderma* **2019**, *356*, 113935. [CrossRef]

63. Duan, Z.; Wang, X.; Sun, L. Monitoring and Mapping of Soil Salinity on the Exposed Seabed of the Aral Sea, Central Asia. *Water* **2022**, *14*, 1438. [CrossRef]

64. Wang, J.; Ding, J.; Yu, D.; Teng, D.; He, B.; Chen, X.; Ge, X.; Zhang, Z.; Wang, Y.; Yang, X.; et al. Machine Learning-Based Detection of Soil Salinity in an Arid Desert Region, Northwest China: A Comparison between Landsat-8 OLI and Sentinel-2 MSI. *Sci. Total Environ.* **2020**, *707*, 136092. [CrossRef] [PubMed]

65. Abdel-Kader, F.H. Digital Soil Mapping at Pilot Sites in the Northwest Coast of Egypt: A Multinomial Logistic Regression Approach. *Egypt. J. Remote Sens. Space Sci.* **2011**, *14*, 29–40. [CrossRef]

66. Abulaiti, A.; Nurmemet, I.; Muhetaer, N.; Xiao, S.; Zhao, J. Monitoring of Soil Salinization in the Keriya Oasis Based on Deep Learning with PALSAR-2 and Landsat-8 Datasets. *Sustainability* **2022**, *14*, 2666. [CrossRef]

67. Ge, X.; Ding, J.; Teng, D.; Wang, J.; Huo, T.; Jin, X.; Wang, J.; He, B.; Han, L. Updated Soil Salinity with Fine Spatial Resolution and High Accuracy: The Synergy of Sentinel-2 MSI, Environmental Covariates and Hybrid Machine Learning Approaches. *CATENA* **2022**, *212*, 106054. [CrossRef]

68. Bandak, S.; Movahedi Naeini, S.A.R.; Komaki, C.B.; Verrelst, J.; Kakooei, M.; Mahmoodi, M.A. Satellite-Based Estimation of Soil Moisture Content in Croplands: A Case Study in Golestan Province, North of Iran. *Remote Sens.* **2023**, *15*, 2155. [CrossRef]

69. Shafizadeh-Moghadam, H.; Minaei, F.; Talebi-khiyavi, H.; Xu, T.; Homaee, M. Synergetic Use of Multi-Temporal Sentinel-1, Sentinel-2, NDVI, and Topographic Factors for Estimating Soil Organic Carbon. *CATENA* **2022**, *212*, 106077. [CrossRef]

70. Zhang, Y.; Kou, C.; Liu, M.; Man, W.; Li, F.; Lu, C.; Song, J.; Song, T.; Zhang, Q.; Li, X.; et al. Estimation of Coastal Wetland Soil Organic Carbon Content in Western Bohai Bay Using Remote Sensing, Climate, and Topographic Data. *Remote Sens.* **2023**, *15*, 4241. [CrossRef]

71. dos Santos, E.P.; Moreira, M.C.; Fernandes-Filho, E.I.; Demattê, J.A.M.; Dionizio, E.A.; da Silva, D.D.; Cruz, R.R.P.; Moura-Bueno, J.M.; dos Santos, U.J.; Costa, M.H. Sentinel-1 Imagery Used for Estimation of Soil Organic Carbon by Dual-Polarization SAR Vegetation Indices. *Remote Sens.* **2023**, *15*, 5464. [CrossRef]

72. Castaldi, F.; Hueni, A.; Chabrillat, S.; Ward, K.; Buttafuoco, G.; Bomans, B.; Vreys, K.; Brell, M.; van Wesemael, B. Evaluating the Capability of the Sentinel 2 Data for Soil Organic Carbon Prediction in Croplands. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 267–282. [CrossRef]

73. 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]

74. 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]

75. 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]
76. 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]
77. 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]
78. 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]