

Mapping of Top-Soil Salinity Zoning in the Coastal Area of Ben Tre Province, Vietnam

Thu Thi Minh Nguyen¹, Ngan Truong Nguyen², Du Xuan Nguyen¹, Cong Thanh Tran³, Doan Quang Tri^{4*}

Submitted: 07/11/2022

Accepted: 11/02/2023

Highlight

- The construction of a model to predict topsoil salinity based on satellite image is currently being studied; one of the popular is multivariable linear regression.
- Applying Bayesian statistical method to build a linear regression model between the on-site topsoil salinity and the channels extracted from Landsat 8 satellite images in the coastal area of Ben Tre province, Vietnam.
- The result of the model is statistically significant, the model efficiency reaches 67%, the RMSE and IOA validation parameters meet the requirements: 1.31 and 0.63 respectively.

Abstract: The study is aimed to construction of top-soil salinity maps for three coastal districts of Ben Tre province, which is one of the Mekong Delta areas that have been seriously affected by saline intrusion in recent years. A total of 124 top-soil samples were collected and analyzed during the dry season in 2021, of which 59 and 65 samples were taken in January and March respectively. Multivariable linear regression model was then built based on measured salinity, by Electrical Conductivity parameter, and spectral bands extracted from Landsat 8 satellite images for the respective days. The model was built based on the Bayesian statistical method, processed by the R statistical program (version 4.0.3), which big difference between Bayesian and classical statistical methods is that this method does not need to evaluate the distribution of the inputs, therefore helping to optimize the quality and quantity of the collected dataset. The results of the model showed that the model efficiency reaches 67%, and the RMSE and IOA validation parameters meet the requirements: 1.31 and 0.63 respectively. The model also showed that the infrared and panchromatic spectral bands are the two spectral bands which are capable of predicting the salinity of the topsoil in this area. Therefore, the results of this study contribute to a potential method in building an effective and low-cost model to assess surface soil salinity in the study area.

Keywords: Top-soil salinity, EC, Landsat 8, Bayesian statistic, multivariable linear regression.

1. Introduction

One of the frequent consequences of drought and sea-level rise is saltwater intrusion from the sea into irrigation systems of the coastal areas, leading to increasing soil salinization and thereby reduce the agricultural production (Shahid et al., 2010, Judkins and Myint, 2012). In Vietnam, saline intrusion in alluvial areas is recently considered a serious problem and forecasted to severely reduce rice and aquaculture production, together with mixed farming systems deficiency, particular in Ben Tre province (CGIAR, 2016). According to Vietnam National Administration of Science and Technology Information, it is assessed that the saline intrusion in 2016 caused the

most severe consequences in the past 100 years, which damages of about 150,000 billion VND, and affects of 170,000 hectares of agricultural crops. In the dry season of 2020 (from January to May), drought and saltwater intrusion took place very seriously in five provinces of the Mekong Delta, of which the most serious is Ben Tre province. Although the monitoring system has continuously constructed and measured the salinity data of surface water, Ben Tre does not have a salinity monitoring system for arable land.

Therefore, predicting soil salinity has become vital and been recently studied, especially in the applying Remote Sensing images. One methodology is to predict the topsoil salinity by developing a regression model between laboratory Electrical conductivity (EC) and the wavelengths channels extracted from the satellite image. The extracted bands can be used to predict soil moisture (Ngo et al., 2019), or soil salinity which are usually salinity index (SI), vegetation index (VI) or other

¹ Sai Gon University, Ho Chi Minh City, Viet Nam

² University of Technology, Ho Chi Minh City, Viet Nam

³ University of Science, Vietnam National University Ho Chi Minh City, Viet Nam

⁴ Vietnam Journal of Hydrometeorology, Viet Nam Meteorological and Hydrological Administration, Hanoi, Viet Nam

*Corresponding author: doanquangtriktvt@gmail.com

characteristic indexes. Studies have shown that the red, near-infrared, and infrared bands were capable of predicting EC with high accuracy (El hafyani et al., 2019; Bannari and Al-Ali, 2020; Fourati et al., 2017; Allbed et al., 2014; Qian et al., 2019). For arid soils, VI indices may not be feasible (Wang et al., 2013; Wang et al., 2020). They however showed possible correlation to soil EC in urban green or agricultural areas (Nouri et al., 2018). For agricultural land with the vegetation cover, studies have shown a negative correlation between SI and VI values (Solangi et al., 2019; Adeeb et al., 2021; Guo et al., 2019; Kurbatova et al., 2021; Sahana et al., 2020). Besides, different models of correlation among the salinity indices, EC and environmental indicators such as climate conditions, land-use information, soil compositions, and topographical data are widely studied. Applied methods include Partial Least Square (Nawar et al., 2014; Fan et al., 2015; Bhat et al., 2015); Principal Component Analysis (Lhissou et al., 2014); or Machine Learning such as Random Forest (RF) algorithm (Hu et al., 2019; Ding et al., 2020; Sultanov et al., 2018). The relationship among SI, soil moisture and EC value was also used as a basic to build a method to determine salinity, such as Tasseled Cap transformation-derived wetness (Hasab et al., 2020; Wang et al., 2019, Gabriel & Soe, 2012; Thiam et al., 2021). Among these methods, multivariable regression approaches were extensively applied (Samra and Ali, 2018; Fourati et al., 2017; Yang et al., 2015; Yahiaoui et al., 2015; Wang et al., 2020; Taghadosi and Hasanlou, 2021; Taghadosi et al., 2019), especially multiple linear regression models (Bouaziz et al., 2011; Gorji et al., 2017; Ijaz et al., 2020; Shahrayini and Noroozi, 2022). In the Mekong Delta provinces, studies on building models to predict saline intrusion in soil using EC values of topsoil and spectral indices from satellite images have also been of great interest in recent years. Study of Hoa et al. (2019) in Ben Tre province used EC field data of 63 soil samples, and Sentinel-1 SAR C-band image remote sensing data; the results showed that Gaussian Processes method has the best fit with the salinity data. Recently, study of Nguyen et al. (2020) in Tra Vinh province used a linear regression model for the SI and VI indexes of Landsat 8 images (OLI) and the EC index. The results showed a good correlation between the NIR near infrared band, SI and EC indexes.

Although different types of salinity index and models have been applied in many studies to predict topsoil salinity, there has not been a certainly accurate model, for the reasons of low salinity, unstable vegetation indices, and unconsidered effects of perennials, or different quality of satellite images. The major problem is not finding a direct relationship between the salinity of the topsoil and the root-zone soil, together with the soil salinity modelling is affected by a number of agronomic, hydrological,

biological, topographic, meteorological, anthropological and multi-scale topographic factors (Akramkhanov et al., 2012; Scudiero et al., 2016, Matinfar et al., 2013). Besides, regarding to the multiple linear regression models, Ordinary Least Square and Maximum Likelihood Estimation are the two most applied methods to estimate regression parameters (Raftery, 1995; Lopez Puga et al., 2015; Fornacon-Wood et al., 2022). These methods require strict assumptions such as error independent, identically normal distribution of the dataset, and the residual for the model validation (Raftery, 1995; Diana and Heruna, 2018; Fornacon-Wood et al., 2022). Furthermore, in developing regression models the most appropriate independent variables need to be chosen in a single suitable model, which can lead to misleading results with classic variable selection procedures and an uncertainty underestimation due to only selecting one single model (Raftery, 1995; Lopez Puga et al., 2015). For these reasons, in this study a different approach to building a regression model between topsoil salinity and satellite image channels was applied, which is the Bayesian methodology. This method has been considered as a better approach for developing multiple linear regression models when compared to traditional methods (Raftery, 1995; Fei et al., 2017; Lopez Puga et al., 2015; Permai and Tanty, 2018, Fornacon-Wood et al., 2022). Being carried out by the Bayes' theorem, this approach considers parameters as random variables and therefore they are observed as probability distributions. The method combines the prior knowledge about a parameter with the likelihood distribution of the parameter in order to converge to the posterior distribution. Independent variables selection, hypothesis testing, model selection, and model uncertainty estimation are simultaneously performed. This method also gives quantitative outcomes for easily comparing to other non-nested models and determining the best model.

In the study, therefore, we applied Bayesian statistical method and GIS-based technique to different dataset in the three coastal districts of Ben Tre province, which are Ba Tri, Binh Dai and Thanh Phu. The purpose of this study is presented as follow: (1) Establish an integrated correlation regression model between two datasets: on-site topsoil salinity data and extracted information channel bands of the Landsat 8 images; (2) Establish top-soil salinity zoning maps of the study area.

2. Materials and Methods

2.1. Description of study area

This research was carried out in three districts: Ba Tri, Thanh Phu and Binh Dai in Ben Tre, a regional province in the Mekong Delta, Vietnam. This area has a total area of 1,166.8 km², and the average population density is 407

people per km² (Figure 1). Except for Ba Tri town in Ba Tri district which is a small urban area with a higher population density than other areas, the remaining areas

are mostly coastal agricultural land, which mainly exploits aquaculture and salt production, perennial orchards, along with mangrove areas in Thanh Phu district.

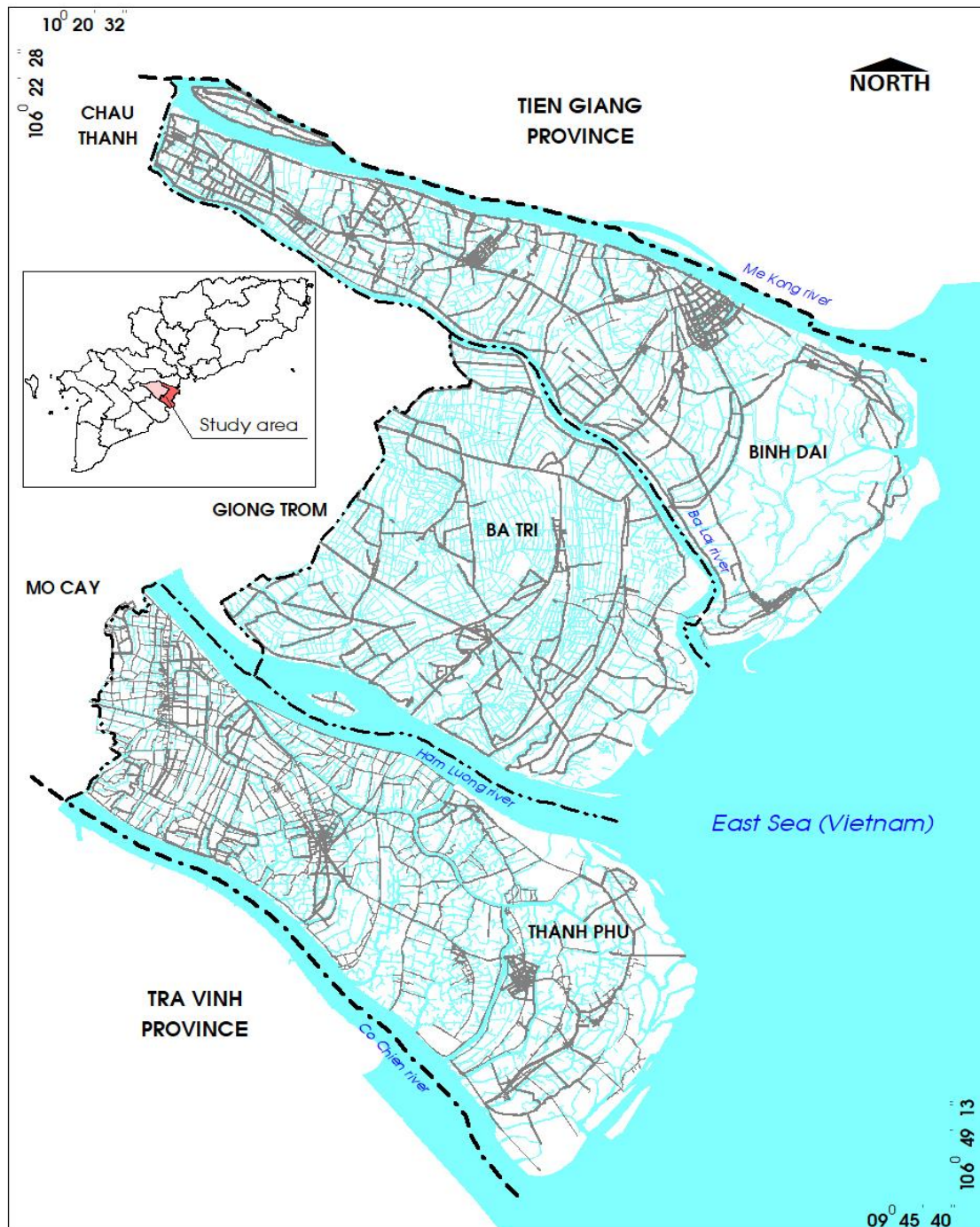


Fig. 1. Study area.

2.2. Methods

Conducting steps of carrying out the study are presented in Figure 2.

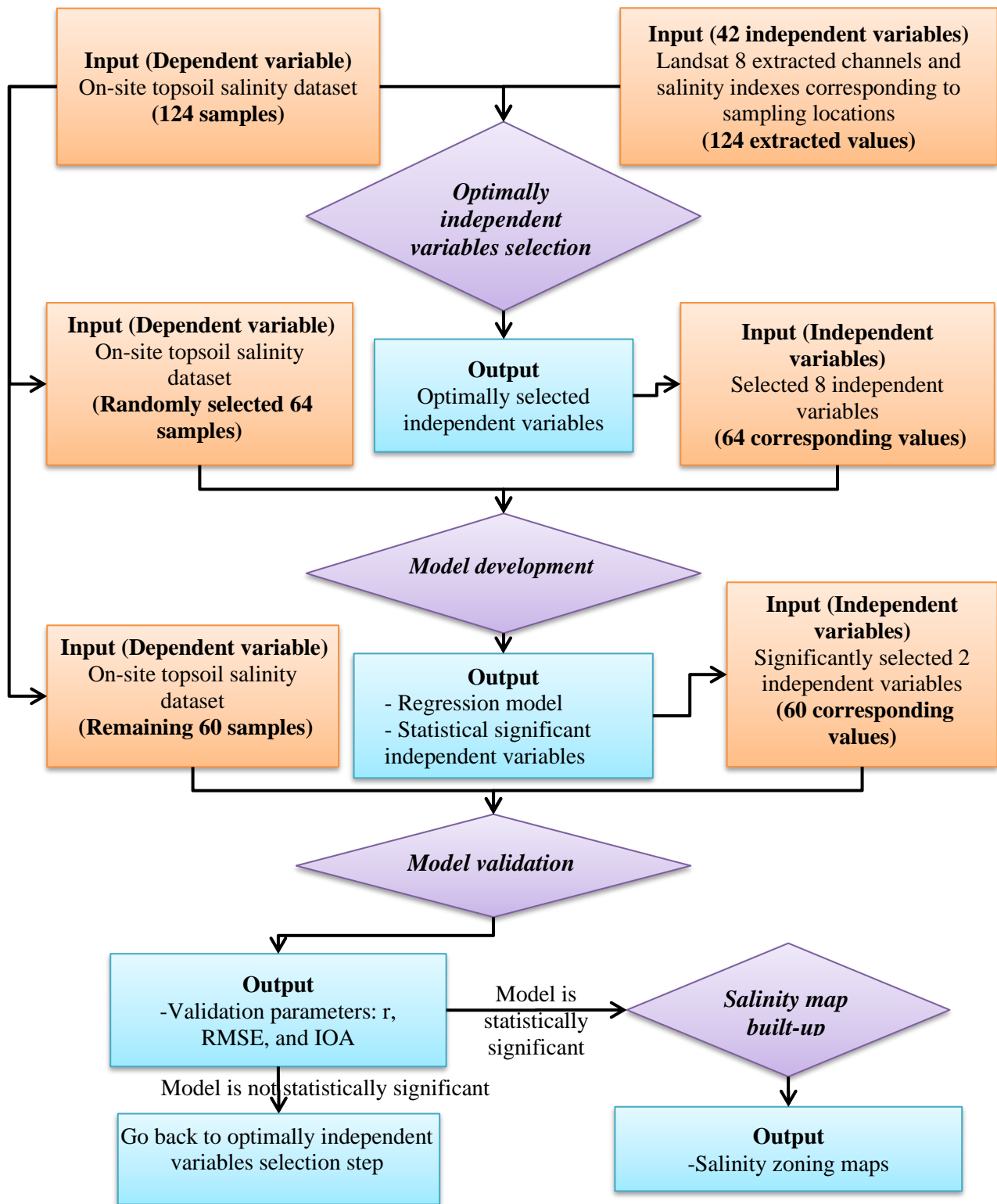


Fig. 2. Study structure flowchart.

2.2.1 Materials

Two main datasets were used in this study: image extracted information and topsoil EC salinity.

a) Image dataset

The base map is provided by the Department of Natural Resources and Environment of Ben Tre Province. The

study also applied remote sensing dataset of Landsat 8 images downloaded from the USGS website, which is a free source of image data with spatial resolution of 30m × 30m. The images were collected in January, and March of 2021, which must have a cloud cover of less than 10%. The process of extracting this information is presented in Figure 3.

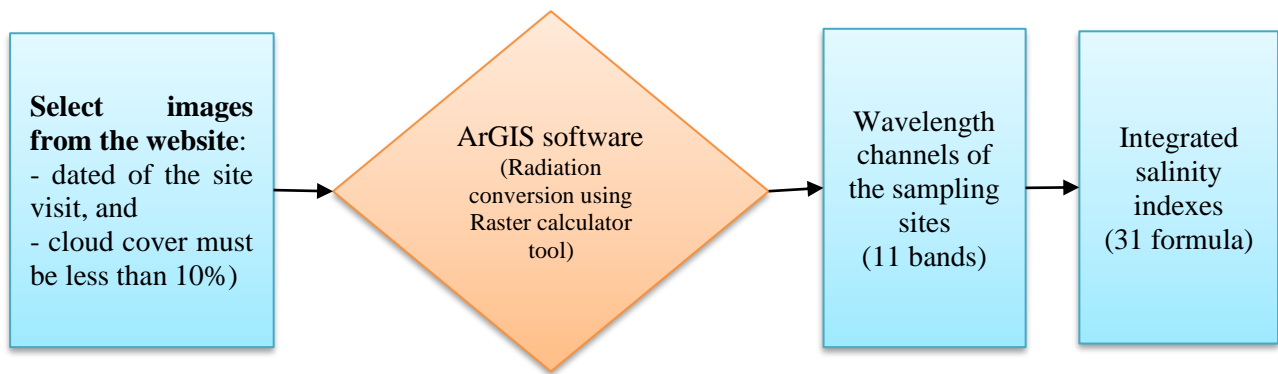


Fig. 3. Images' process of extracting wavelength bands and salinity indexes.

We used specialized remote sensing software to extract a total of 11 image channels from selected Landsat 8 images, shown in Table 1. Based on these channels, we

then calculated additional 31 salinity indices as a basis for evaluation, presented in Table 2. In total, 42 independent variables are considered in the model development.

Table 1. Landsat 8 extracted channels.

Bands	Label	Wavelength (micrometers)	Resolution (meters)
Band 1 - Coastal aerosol (B1)	B1	0.43-0.45	30
Band 2 - Blue	B2	0.45-0.51	30
Band 3 - Green	B3	0.53-0.59	30
Band 4 - Red	B4	0.64-0.67	30
Band 5 - Near Infrared (NIR)	B5	0.85-0.88	30
Band 6 - SWIR 1	B6	1.57-1.65	30
Band 7 - SWIR 2	B7	2.11-2.29	30
Band 8 - Panchromatic	B8	0.50-0.68	15
Band 9 - Cirrus	B9	1.36-1.38	30
Band 10 - Thermal Infrared (TIRS) 1	B10	10.6-11.19	100
Band 11 - Thermal Infrared (TIRS) 2	B11	11.50-12.51	100

<https://www.usgs.gov/>

Table 2. Calculated top-soil salinity indexes.

No.	Index lable	Calculation	Reference
1	NDVI	$(B5-B4)/(B5+B4)$	Meyer et al., 2019
2	VSSI	$(2*B3)-5*(B4+B5)$	Ghazavi et al., 2018
3	Alberdo	$(0.356*B2)+(0.13*B4)+(0.375*B5)+(0.085*B6)$	Sahbeni, 2021
4	CRSI	$\sqrt{((B5 * B4) - (B3 * B4))/((B5 * B4) + (B3 * B4))}$	Sahbeni, 2021
5	BI	$\sqrt{(B4)^2 + (B5)^2}$	Ghazavi et al., 2018
6	SI	$B4*B3/B2$	Ghazavi et al., 2018
7	SI1	$\sqrt{B4 * B3}$	Ghazavi et al., 2018

No.	Index lable	Calculation	Reference
8	SB4	$\sqrt{B4 * B5}$	Ghazavi et al., 2018
9	SI3	$\sqrt{(B3)^2 + (B4)^2 + (B5)^2}$	Guo et al., 2019
10	SI4	$\sqrt{(B4)^2 + (B3)^2}$	Guo et al., 2019
11	SI5	B2/B4	Ghazavi et al., 2018
12	RSI	B4/B5	Sahbeni, 2021
13	DVI	B5-B4	Meyer et al., 2019
14	Int1	1/2*(B3+B4)	Sahbeni, 2021
15	Int2	1/2*(B4+B5+B3)	Sahbeni, 2021
16	SR	(B4-B5)/(B3+B5)	Sahbeni, 2021
17	GNDVI	(B5-B3)/(B5+B3)	Guo et al., 2019
18	SAVI	(1+L)*B5 - B4/L +B5+B4 (L = 0.5)	Meyer et al., 2019
19	MSAVI	$\frac{(2B5 + 1) - \sqrt{(2B5 + 1)^2 - 8(B5 - B4)}}{2}$	Guo et al., 2019
20	VSSI	2*B3 - 5*(B4+B5)	Sahbeni, 2021
21	RVI	B4/B3	Yahiaoui et al., 2015
22	BI2	$\sqrt{(B4)^2 + (B2)^2}$	Sahbeni, 2021
23	ENDVI	(B5+B7-B4)/(B5+B7+B4)	Guo et al., 2019
24	Aster SI	(B6-B7)/(B6+B7)	Wang et al., 2020
25	LSWI	(B5-B6)/(B5+B6)	Gao, 1996
26	TCT-Bright	B2*0.3029+B3*0.2786+B4*0.4733+B5*0.5599+B6*0.508 +B7*0.1872	Kauth & Thomas, 1976
27	TCT-Green	B2*0.3029+B3*0.2786+B4*0.4733+B5*0.5599+B6*0.508 +B7*0.1872	Kauth & Thomas, 1976
28	TCT-Wet	B2*0.1511+B3*0.1973+B4*0.3283+B5*0.3407+B6*(-0.7117)+B7*(-0.4559)	Kauth & Thomas, 1976
29	TCT4	B2*(-0.8239)+B3*0.0849+B4*0.4396+B5*(-0.058)+B6*0.2013+B7*(-0.2773)	Kauth & Thomas, 1976
30	TCT5	B2*(-0.3294)+B3*0.0557+B4*0.1056+B5*0.1855+B6*(-0.4349) +B7*0.8085	Kauth & Thomas, 1976
31	TCT6	B2*0.1079+B3*(-0.9023)+B4*0.4119++B5*0.0575+B6*(-0.0259) +B7*0.0252	Kauth & Thomas, 1976

b) *EC Salinity dataset*

In this research, we applied the Electrical conductivity of the topsoil to assess the salinity of the soil, with the unit being dS/m (1dS/m = 0.64‰). The topsoil sample in this study was taken from the topsoil to about 5-20 cm below.

Saline soils are those with the conductivity EC > 4 dS/m at 25°C (Richards 1954), equivalent to a dissolved salt concentration of about 2.56‰. In the study, we collected 124 samples, equally distributed into 3 districts. Sampling is divided into two phases: 59 samples in January and 65 samples in March 2021 (Figure 4).

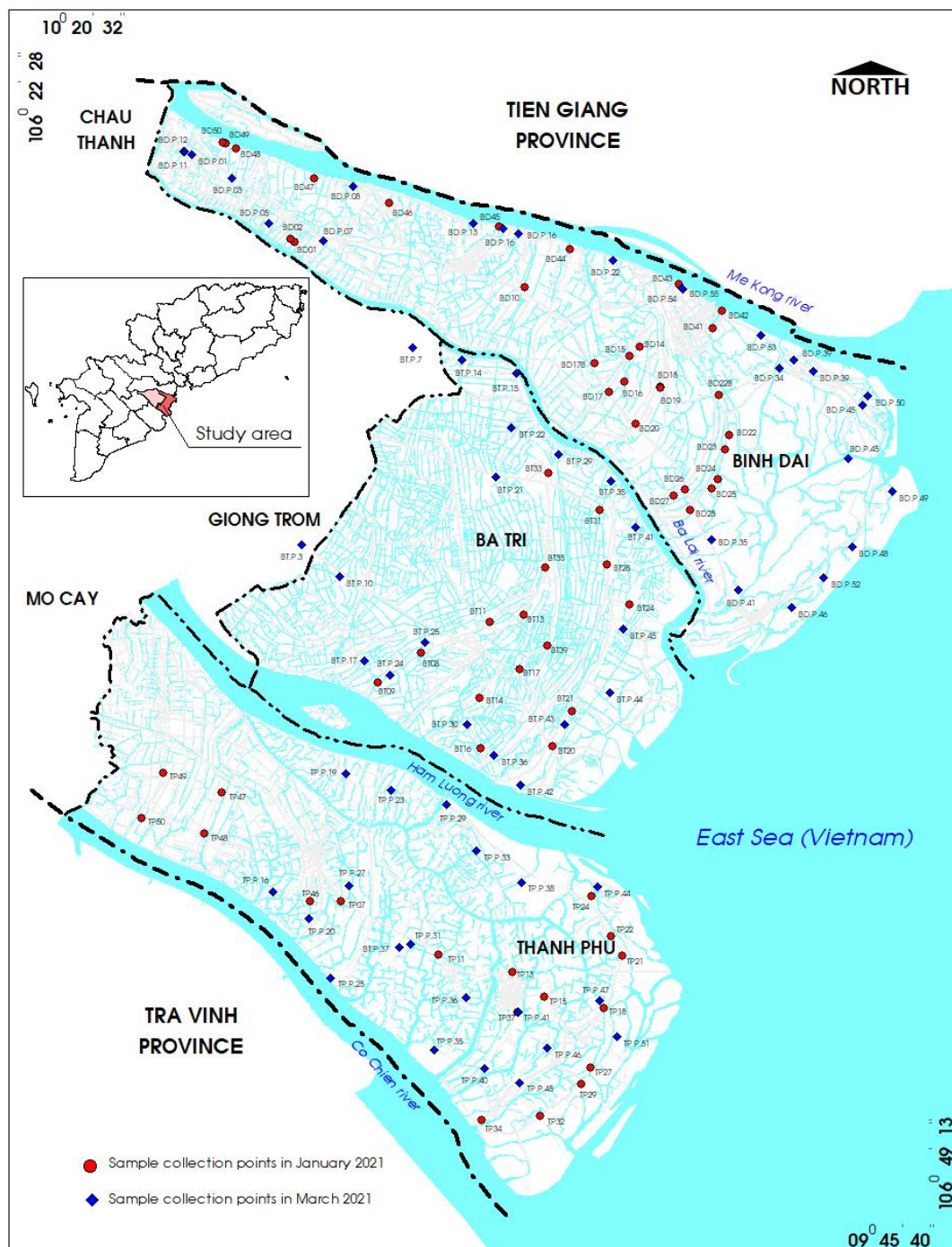


Fig. 4. Sampling sites in January and March 2021.

Soil samples after being collected at the field were preserved and brought to the laboratory to determine the EC value. According to Vietnam Standard TCVN 6650:2000 on soil quality, the process includes the following steps: Soil samples are removed from surface impurities, dried in natural conditions or by a dryer, and then finely grounded. The sample was then weighed 20.00 grams into a shaker bottle; add to 100 ml of deionized pure water at $20^{\circ}\text{C} \pm 1^{\circ}\text{C}$; cap the bottle and place it in the shaker in a horizontal position; shake 30 minutes; then filter directly through a filter paper. The filtered solution is then tested EC by a specialized digital refractometer. The blank sample is processed in the same way, which the tested value should not exceed 1 mS/m.

2.3. Methods

2.3.1. Bayesian multiple linear regression models

This research applied the Bayesian approach to estimate regression model parameters by multiplying the prior distribution and likelihood function to converge posterior distribution (Rubio and Genton, 2016).

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}$$

The estimation process in this research based on Bayesian Model Average approach (BMA) which clarifying the inevitable model uncertainty in variable selection procedures by averaging over the best models in a combination of candidate models according to

approximate posterior model probability (Fragoso et al., 2018). All statistical analyses were performed in R (version 4.0.3) by using the “BMA” package (version 3.18.15). The process of conducting the model is presented as follow.

Optimally independent variables selection

A total of 42 indexes from satellite-based data of soil salinity (independent variables) (Table 1 and 2) and 124 sampling values (dependent variable) of two-phases ground-measured data of soil salinity were used to derive optimally correlated variables by using the BMA approach. The best model with optimally correlated variables were simply selected by the one with highest posterior probability.

Model development

Based on the results of the optimally selected variables, the inputs were randomly and automatically divided into development and validation datasets with a ratio 1:1. The data splitting process was performed in R (version 4.0.3) by using the “caret” package (version 1.0-90). The

development dataset was continuously applied to build a multiple linear regression model with 64 sampling values (dependent variable) and the corresponding selected variables (independent variables). To evaluate important variables in the model, the “relaimpo” package (version 2.2-6) was applied.






Model validation

The validation dataset (about 50% of the total dataset, 60 samples) was used to demonstrate a validation of the model by testing the correlation between prediction values from the model and observed values from the validation data set. The accuracy of the model was evaluated by using Correlation coefficient (r), Root mean squared error (RMSE), and Index of Agreement (IOA).

2.3.2. Salinity zoning map

Based on the selected salinity indexes in the linear regression model, salinity values are assigned to each pixel of the map. The construction of the salinity zoning map is divided into five colour zones, presented in Table 3.

Table 3. Assigned colours for salinity zoning maps.

Colour	Salinity level	Equivalent EC (dS/m)	Equivalent salinity (part per thousand)	Effect on crops
	Non saline	0 - 2	0 – 1.36	Not detect saline effects
	Slightly saline	2 - 4	1.36 - 2.56	Restricted yields of sensitive crops
	Moderately saline	4 – 8	2.56 – 5.12	Restricted yields of many crops
	Strongly saline	8 – 16	5.12 – 10.24	Only tolerant crops satisfactorily
	Very strongly saline	>16	>10.24	Only very few tolerant crops satisfactorily

<https://www.fao.org/>

3. Results and Discussion

3.1. Descriptive statistics of the tested soil salinity

Statistical parameters describing the salinity values converted from EC samples are presented in Table 4. In phase 1, the average salinity value was 1.169, equivalent to very low salinity levels. The average salinity value of

the second phase obtained in March shows that the salinity level in the topsoil is higher than that of the first phase and is approximately equivalent to brackish water, compared with the water salinity. The standard deviations of the two corresponding phases are otherwise high, at 1.227 and 2.019, respectively; together with high skewness and kurtosis values indicating an asymmetric distribution.

Table 4. Descriptive statistics of the tested soil salinity (unit: part per thousand).

	Phase 1 (January 2021)	Phase 2 (March 2021)	Both two phases
Number of observation	59	65	124
Minimum	0.001	0.427	0.001

	Phase 1 (January 2021)	Phase 2 (March 2021)	Both two phases
Maximum	5.562	9.856	9.856
1st Quartile	0.462	0.664	0.547
3rd Quartile	1.344	3.776	2.656
Mean	1.169	2.209	1.714
Median	0.704	1.024	0.871
Sum	68.959	143.566	212.525
SE Mean	0.160	0.250	0.158
Variance	1.505	4.075	3.102
Stdev	1.227	2.019	1.761
Skewness	2.224	1.234	1.690
Kurtosis	4.713	1.345	2.905

The data show some unusually high salinity level, and therefore they are possibly outliers, presented in Figure 5. Abnormally high values can be either coastal sampling points or riparian soils. The mean score also deviates from the central position of the sample series, showing the tendency to eccentricity of the data.

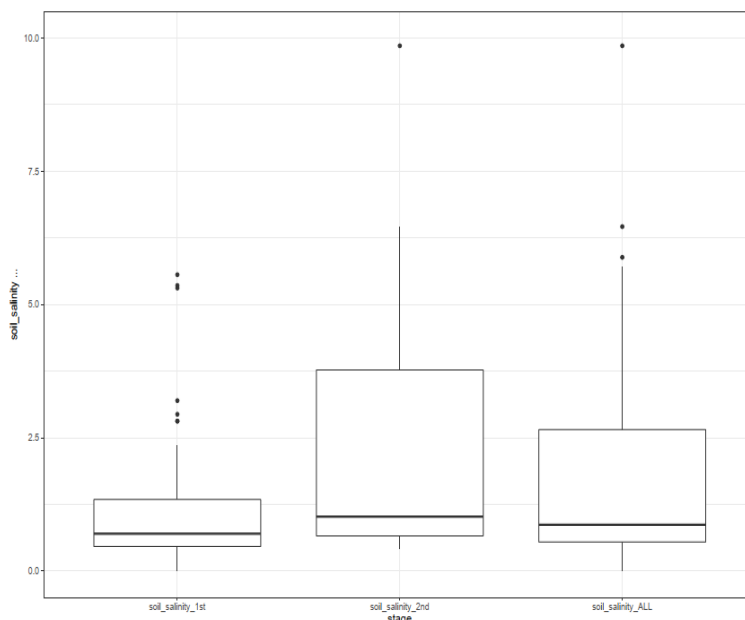


Fig. 5. Box plots of the tested soil salinity.

3.2 Bayesian multiple linear regression models

3.2.1. Optimally correlated variables selection

Among 42 proposed variables, eleven were chosen as potential variables, including B1, B5, B7, B8, B9, B10, B11, NDVI, Alberdo, CRSI, and BI. The importance of potential explanatory variables is illustrated in Figure 6.

From these variables, four models were established and assessed, shown in Table 5. Model 1 is best fit with the highest posterior probability (post prob = 72%). Therefore, eight variables B1, B5, B7, B8, B9, B10, B11, and NDVI were selected as optimal correlated variables in the next stage.

Table 5. Optimal correlated variables.

	p!=0	EV	SD	model 1	model 2	model 3	model 4
Intercept	100	13.044	5.637	12.566	13.905	15.493	14.508
B1	100	-14.3817	22.929	-4.183	-36.984	-58.974	-34.832

	p!=0	EV	SD	model 1	model 2	model 3	model 4
B5	100	-6.6722	25.459	5.545	-30.262	-54.459	-57.205
B7	100	-17.0232	7.993	-13.821	-25.343	-28.757	-20.39
B8	100	-2.9504	6.447	-1.975	-5.348	-5.992	-5.446
B9	100	-62.1756	55.627	-60.896	-67.561	-60.207	-62.428
B10	100	8.9982	26.619	5.258	15.511	23.891	27.744
B11	100	-25.2093	34.387	-20.36	-33.978	-44.43	-47.975
NDVI	100	-7.1104	2.46	-7.855	-6.404	-2.17	-3.16
Alberdo	27.8	25.5221	56.087	.	85.902	140.969	51.452
CRSI	8.9	-0.2438	1.092	.	.	-3.401	-1.782
BI	3.7	1.4721	8.538	.	.	.	40.181
No.ofvariable				8	9	10	11
r2				0.562	0.57	0.578	0.592
BIC				-60	-57.321	-54.758	-54.039
post prob				0.722	0.189	0.052	0.037

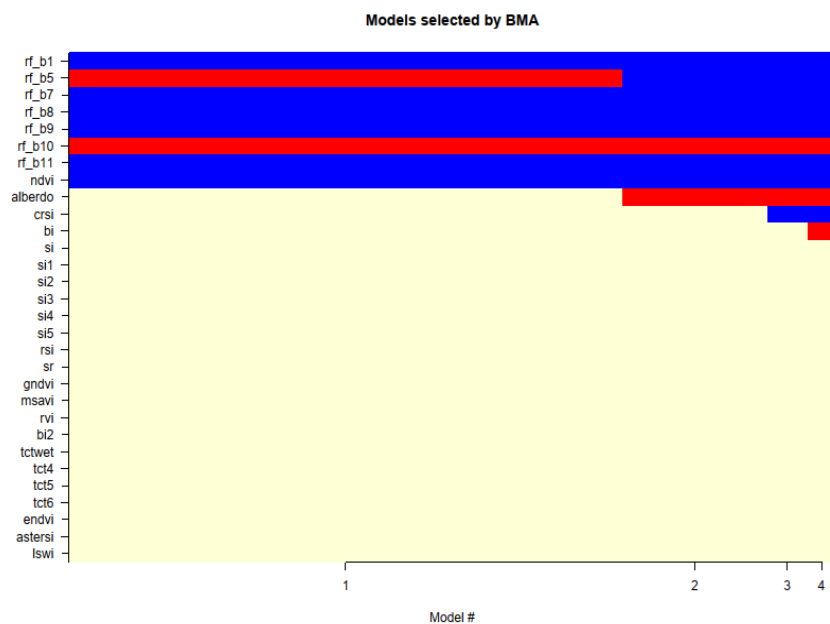


Fig. 6. The importance of potential explanatory variables.

3.2.2. Model development

The development data set with 64 sampling values (dependent variable) and 64 corresponding values of 8 optimally selected variables from model 1 (independent variables) was used to build a multiple linear regression

model. The results of the model is shown in Table 6. In this model, two independent variables (B5 and B8) are statistically significant. The adjusted R-squared is 54% and the model is also statistically significant (p-value <0.001).

Table 6. Multiple linear regression analysis results.

Coefficients	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	5.0306	8.2186	0.612	0.542995	
B1	-5.6832	10.5768	-0.537	0.593211	
B5	-10.6497	2.7283	-3.903	0.000261	***

Coefficients	Estimate	Std. Error	t value	Pr(> t)
B7	-7.0237	5.8269	-1.205	0.233216
B8	16.1316	8.2121	1.964	0.054549
B9	57.2279	107.564	0.532	0.596844
B10	-17.1288	37.8612	-0.452	0.652753
B11	17.5898	46.0384	0.382	0.703884
NDVI	-0.2114	0.8762	-0.241	0.81028

Residual standard error: 1.081 on 55 degrees of freedom

Multiple R-squared: 0.6001

Adjusted R-squared: 0.5419

F-statistic: 10.32 on 8 and 55 DF

p-value: 1.123e-08

Residuals:

Min	1Q	Median	3Q	Max
-1.99568	-0.67507	-0.07298	0.71821	2.25727

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The values of the two independent variables (B5 and B8) and the corresponding sampling values from the development data set were abstracted again to determine a next potential model. The results of this model was displayed in Table 7.

Table 7. Multiple linear regression analysis results for two independent variables B5 and B8.

Coefficients	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.046	0.356	11.366	< 2e-16
B8	8.39	2.082	4.03	0.000157
B5	-12.971	1.569	-8.269	1.52E-11

Residual standard error: 1.107 on 61 degrees of freedom

Multiple R-squared: 0.5348

Adjusted R-squared: 0.5196

F-statistic: 35.07 on 2 and 61 DF

p-value: 7.278e-11

Residuals:

Min	1Q	Median	3Q	Max
-2.0475	-0.8104	0.1333	0.7748	2.1529

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

In this model, the intercept and the two independent variables (B5 and B8) are statistically significant (p-value < 0.001). The adjusted R-squared is 52% and the model is also statistically significant (p-value < 0.001). Also, the median of residuals is approximately zero that means the assumption of normal distribution of residuals is satisfied. Thus, the model may be considered valid. The regression for ground-measured data and satellite-based data is presented in Equation 1 as follows:

$$\text{Soil salinity} = 4.046 - 12.971 \cdot B5 + 8.390 \cdot B8 \quad (1)$$

The total proportion of variance explained by (1) with two predictors (B5 and B8) is 44.9%. The quantified contribution to the model of B5 is higher than the one of B8 which is approximately 4 times.

3.2.3. Model validation

Both quantile-quantile plot (Q-Q plot) in Figure 7 show a good agreement between prediction values from Eq.1 and observed values from the validation dataset. Furthermore, the results of the accuracy of Equ.1 indicate that the Correlation coefficient (r), Root mean squared error (RMSE), and Index of Agreement (IOA) values are 0.67, 1.31, and 0.63, respectively.

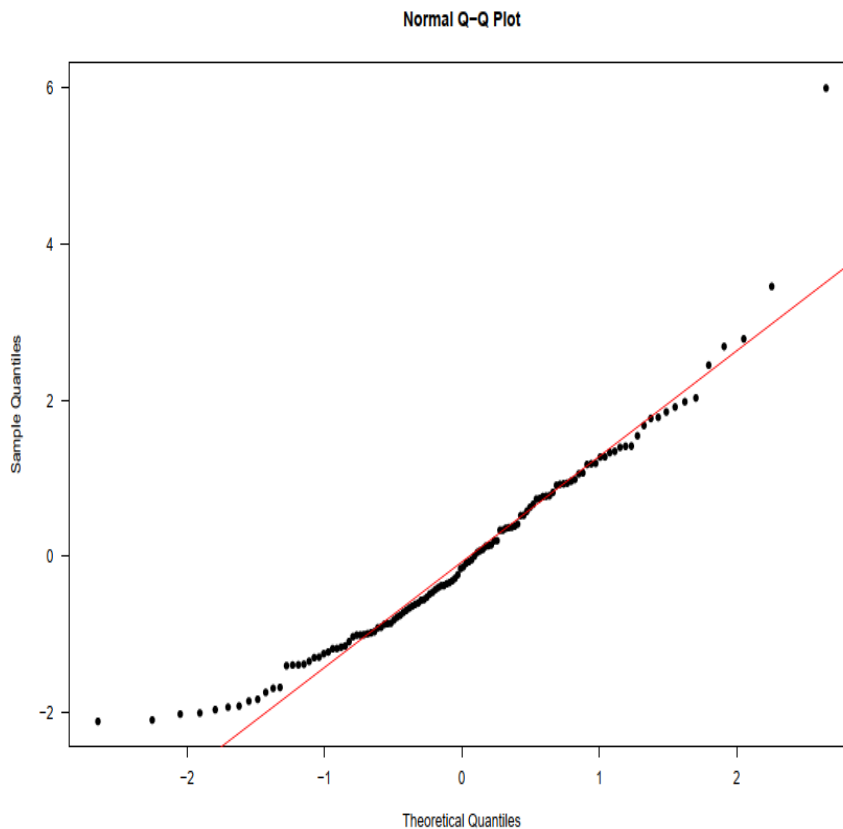


Fig. 7. The quantile-quantile plot (Q-Q plot) between prediction and observed values.

With the Correlation coefficient (r) of 67%, there is a linear relationship between predicted and observed values. The RMSE value of about 1.31 part per thousand provides that there is a good overall measure of the agreement between modelled soil salinity values and predicted soil salinity values. The IOA value of 0.63, which is close to 1, means the model performance is also quite good (Willmott et al., 2012). In overall, the results of model evaluation show that Equ.1 is considered valid and can be used to map a salinity distribution for the study area in the next step.

3.2.4. Salinity zoning map

Based on the developed model (1), salinity zoning maps are built and presented in Figure 8 and 9. From the maps, it can be seen that the main direction of saline intrusion is from the sea to the mainland. Among the three study areas, Thanh Phu and Binh Dai are the two districts having the

lengthiest saline intrusion, with the average saline area accounting for about two-thirds of the districts' area. Most of the saline area of Ba Tri district is along the coastal line. This area also has the lowest rate of saline intrusion among the three study areas, which can be explained that Ba Tri is an urbanized area with a high population density. Having barriers built along rivers also helps to minimize salinity intrusion in Ba Tri when compared to the remaining areas which are mostly agricultural areas. At the end of March, the total area of land affected by salinity increased by 4.63% (equivalent to 4,590 ha) compared to January. Comparing the rate of salinization between January and March, the results show a slight increase in salinity in the central area of Ba Tri in March and a few small inland areas of Thanh Phu and Binh Dai districts. This can be explained by reversed osmosis from underground to surface soil.

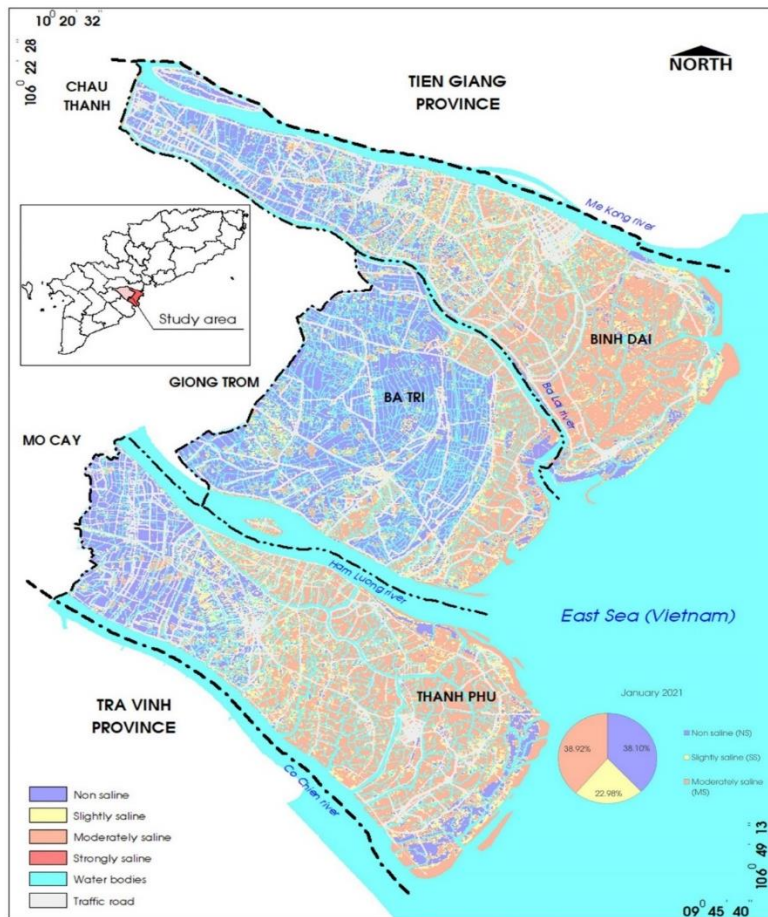


Figure 8. Salinity zoning map in January 2021.

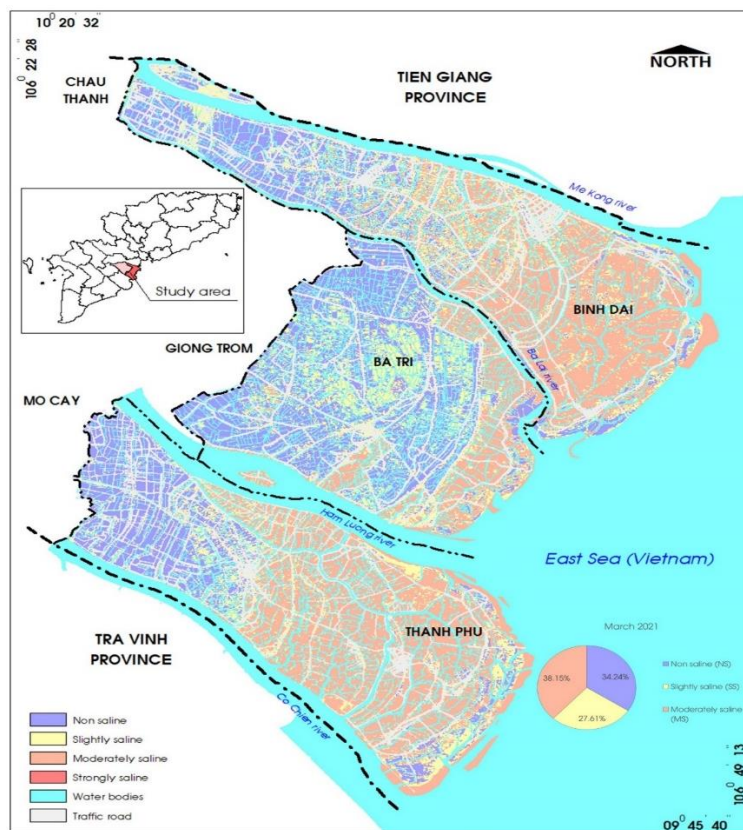


Figure 9. Salinity zoning map in March 2021.

The salinity change from January to March 2021 is shown in Figure 10. The results show that more than 56% of the study area (71,755 hectares) does not change the salinity. Areas with a large increase in salinity (coloured red in Figure 10, changing from NS to MS) are less than 450 ha, accounting for only about 0.35% of the total study area. These areas are scatteredly distributed in Thanh Phu district. Areas with slight increase in salinity (coloured orange in Figure 10, changing from NS to SS or from SS to MS) cover an area of 14,460 ha, accounting for 11.40%

of the study area. This area is concentratedly distributed into 3 areas: (i) along the coastal line with a distance of 2-3km from the coast, (ii) along both sides of the river and canal system, and (iii) distributed in large clusters in the center of Ba Tri district. The areas with a large decrease in salinity (coloured blue in Figure 10, changing from MS to NS) are of about 1,677 ha, accounting for 1.32% of the total study area. This area is mostly located in the mainland, and scatteredly distributed in all 3 districts.

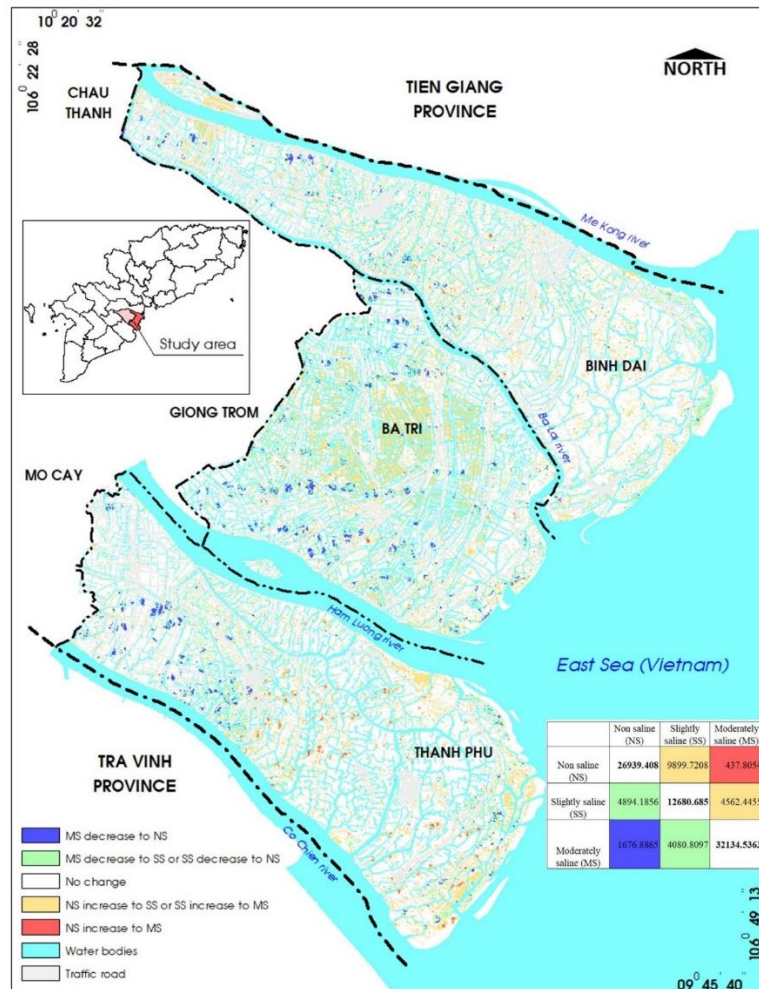


Figure 10. Salinity changes from January to March 2021.

4. Conclusions

Salinity assessment model based on multivariable linear regression model according to Bayesian methodology has a prominent advantage compared to traditional methodology which is not necessary to evaluate the distribution of variable groups before running the model. This advantage can help to save a lot of time and effort, as well as take advantage of the most of the collected data. The model of assessing top-soil salinity based on the Bayesian approach presented in this study has shown its significance, feasibility and applicability to building saline maps of the topsoil. The research results show that the infrared (B5) and panchromatic (B8) channels extracted from Landsat 8 satellite images can represent the

salinity of the topsoil, helping to confirm the previous studies' result that infrared channel can reflect the soil salinity (El hafyani et al., 2019; Bannari and Al-Ali, 2020; Fourati et al., 2017; Allbed et al., 2014; Qian et al., 2019). In the study, the built salinity prediction model can therefore be applied as another fast and cost-effective method in assessing the salinity intrusion from the aquatic into the soil environment, and thereby can predict the extent of damage to agriculture and people's livelihood.

5. Recommendations

From the topsoil salinity results of the three coastal districts of Ben Tre province, and the FAO's soil salinity classification table (Table 3), it can be assessed that the

salinity is mostly in the low to moderate salinity categorizes. However, FAO classification of soil salinity may not be suitable for the assessment of topsoil salinity due to reverse osmosis from below underground to the surface soil layers. The impact of the topsoil salinity can therefore be inaccurate, as the root layers of vegetation are often located deeper below the surface.

In a recent study in a neighboring province, Tra Vinh, the near infrared band and SI index are suggested to predict the top-soil salinity (Nguyen et al., 2020). This study, however, none of former salinity indexes (Table 2) is applicable in the studied region. This is consistent with previous study about other influencing factors such as topography, vegetation cover, or low salinity of the area (Scudiero et al., 2016). One other issue is that as most of the collected data are in the classification from none to slightly saline; the model's efficiency is not high. Therefore, we believe that there is a need for further studies on the relationship between salinity between the underlying and the top-soil layer, especially for groups with different vegetation layers and soil profiles to obtain more accurate assessments of the impacts of saline intrusion to the agricultural activities. Besides, the number of sampling points needs to be more expanded and distributed to increase model performance.

Author Contributions: Conceptualization, Thu Thi Minh Nguyen, Ngan Truong Nguyen, Du Xuan Nguyen, Cong Thanh Tran, Doan Quang Tri; methodology, Thu Thi Minh Nguyen, Ngan Truong Nguyen, Du Xuan Nguyen, Cong Thanh Tran; software, Ngan Truong Nguyen, Cong Thanh Tran, Doan Quang Tri; validation, Thu Thi Minh Nguyen; formal analysis, Thu Thi Minh Nguyen; investigation, Thu Thi Minh Nguyen, Ngan Truong Nguyen; resources, Du Xuan Nguyen; data curation, Du Xuan Nguyen; writing—original draft preparation, Thu Thi Minh Nguyen, Ngan Truong, and Cong Thanh Tran; writing—review and editing, Thu Thi Minh Nguyen, Ngan Truong, Du Xuan Nguyen and Cong Thanh Tran; visualization, Ngan Truong, Cong Thanh Tran, Doan Quang Tri; supervision, Thu Thi Minh Nguyen, Ngan Truong, Doan Quang Tri; All authors have read and agreed to the published version of the manuscript.

Acknowledgement: This study is funded by the Sai Gon University of the project titled “Mapping of top-soil salinity zoning in the coastal area of Ben Tre province, Vietnam” grant number: TD2020-58.

Data Availability Statement: The data presented in this study are available in article.

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

References

- [1] Akramkhanov, A., Vlek, P.L.G., 2012. The assessment of spatial distribution of soil salinity risk using neural network. *Environ. Monit. Assess.* 184, 2475–2485. <https://doi-org/10.1007/s10661-011-2132-5>
- [2] Allbed, A., Kumar, L., Sinha, P., 2014. Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques. *Remote Sens.* 6, 1137–1157. <https://doi-org/10.3390/rs6021137>.
- [3] Adeeb, H.Q., Al-Timiti, Y.K., 2021. Change on detection of vegetation cover and soil salinity using GIS technique in Diyala Governorate, Iraq. *Sci. Rev. Eng. Environ. Sci.* 30 (1), 148–158. **Error! Hyperlink reference not valid.**
- [4] Bannari, A., Al-Ali, Z.M., 2020. Assessing Climate Change Impact on Soil Salinity Dynamics between 1987–2017 in Arid Landscape Using Landsat TM, ETM+ and OLI Data. *Remote Sens.* 12, 2794. <https://doi-org/10.3390/rs12172794>.
- [5] Bouaziz, M., Matschullat, J.; Gloaguen, R., 2011. Improved remote sensing detection of soil salinity from a semi-arid climate in Northeast Brazil. *C.R. Geosci.* 343, 795–803. <https://doi.org/10.1016/j.crte.2011.09.003>.
- [6] Bhat, M.A., Sheoran, H.S., Dar, E.A., Dahiya, H.S., Wani, S.A., Singh, I., Sing, S., 2015. Geoinformatics as a Tool for Appraisal of Salt-Affected Soils-A Review. *Int. J. Innovative Sci. Eng. Technol.* 2 (10), 480–490.
- [7] CGIAR Research Program on Climate Change, Agriculture and Food Security- Southeast Asia (CCAFS-SEA). 2016. Assessment Report: The drought and salinity intrusion in the Mekong River Delta of Vietnam. Hanoi, Vietnam.
- [8] Diana, P.S., Heruna, T., 2018. Linear regression model using bayesian approach for energy performance of residential building. *Procedia Comput. Sci.* 135, 671–677.
- [9] Ding, J., Yang, S., Shi, Q., Wei, Y., Wang, F., 2020. Using Apparent Electrical Conductivity as Indicator for Investigating Potential Spatial Variation of Soil Salinity across Seven Oases along Tarim River in Southern Xinjiang, China. *Remote Sens.* 12, 2601. **Error! Hyperlink reference not valid.**
- [10] El Hafyani, M., Essahlaoui, A., El Baghdadi, M., Teodoro, A.C., Mohajane, M., El hmadi, A., El ouali, A., 2019. Modeling and mapping of soil salinity in Tafilalet plain (Morocco). *Arabian J. Geosci.* 12, 35. <https://doi.org/10.1007/s12517-018-4202-2>.
- [11] Fei, Y.H., She, D.L., Yao, Z.D., Li, L., Ding, J.H.; Hu, W., 2017. Hierarchical Bayesian models for predicting soil salinity and sodicity characteristics in

- a coastal reclamation region. *Ecol. Eng.* 104, 45–56. <https://doi.org/10.1016/j.ecoleng.2017.04.006>.
- [12] Fornacon-Wood, I., Mistry, H., Johnson-Hart, C., Faivre-Finn, C., O'Connor, J.P. B., Price, G.J., 2022. Understanding the Differences Between Bayesian and Frequentist Statistics. *Int. J. Radiat. Oncol. Biol. Phys.* 112, 1076–1082. <https://doi.org/10.1016/j.ijrobp.2021.12.011>.
- [13] Fragoso, T.M., Bertoli, W.; Louzada, F., 2018. Bayesian Model Averaging: A Systematic Review and Conceptual Classification. *Int. Stat. Rev.* 86, 1–28. <https://doi.org/10.1111/insr.12243>.
- [14] Fourati, H.T., Bouaziz, M., Benzina, M., Bouaziz, S., 2017. Detection of terrain indices related to soil salinity and mapping salt-affected soils using remote sensing and geostatistical techniques. *Environ. Monit. Assess.* 189, 177. <https://doi.org/10.1007/s10661-017-5877-7>.
- [15] Fan, X., Liu, Y., Tao, J., Weng, Y., 2015. Soil Salinity Retrieval from Advanced Multi-Spectral Sensor with Partial Least Square Regression. *Remote Sens.* 7, 488–511. <https://doi.org/10.3390/rs70100488>.
- [16] Judkins, G., Myint, S., 2012. Spatial Variation of Soil Salinity in the Mexicali Valley, Mexico: Application of a Practical Method for Agricultural Monitoring. *Environ. Manage.* 50, 478–489. <https://doi.org/10.1007/s00267-012-9889-3>.
- [17] Gao, B.C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58 (3), 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3).
- [18] Gorji, T., Sertel, E., Tanik, A., 2017. Recent satellite technologies for soil salinity assessment with special focus on Mediterranean countries. *Fresenius Environ. Bull.* 26 (1), 196–203.
- [19] Gorji, T., Sertel, E., Tanik, A., 2017. Monitoring soil salinity via remote sensing technology under data scarce conditions: A case study from Turkey. *Ecol. Indic.* 74, 384–391. <https://doi.org/10.1016/j.ecolind.2016.11.043>.
- [20] Ghazavi, R., Pakparvar, M., 2018. Mapping spatial variability of soil salinity in a coastal area located in an arid environment using geostatistical and correlation methods based on the satellite data. *Desert* 23-2, 233-242. Available Online: <http://desert.ut.ac.ir>.
- [21] Guo, B., Han, B., Yang, F., Fan, Y., Jiang, L., Chen, S., Yang, W., Gong, R., Liang, T., 2019. Salinization information extraction model based on VI–SI feature space combinations in the Yellow River Delta based on Landsat 8 OLI image. *Geomatics Nat. Hazards Risk* 10 (1), 1863–1878. <https://doi.org/10.1080/19475705.2019.1650125>.
- [22] Hasab, H.A., Dibs, H., Dawood, A.S., Hadi, W.H., Hussain, H.M., Al-Ansari, N., 2020. Monitoring and Assessment of Salinity and Chemicals in Agricultural Lands by a Remote Sensing Technique and Soil Moisture with Chemical Index Models. *Geosciences* 10, 207. **Error! Hyperlink reference not valid.**
- [23] Hu, J., Peng, J., Zhou, Y., Xu, D., Zhao, R., Jiang, Q., Fu, T., Wang, F., Zhou Shi, Z., 2019. Quantitative Estimation of Soil Salinity Using UAV-Borne Hyperspectral and Satellite Multispectral Images. *Remote Sens.* 11, 736. <https://doi.org/10.3390/rs11070736>.
- [24] Hoa, P.V., Giang, N.V., Binh, N.A., Hai, L.V.H., Pham, T.D., Hasanlou, M., Bui, D.T., 2019. 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.* 11, 128. <https://doi.org/10.3390/rs11020128>.
- [25] Ijaz, M., Ahmad, H.R., Bibi, S., Ayub, M.A., Khalid, S., 2020. Soil salinity detection and monitoring using Landsat data: a case study from Kot Addu, Pakistan. *Arabian J. Geosci.* 13, 510. <https://doi.org/10.1007/s12517-020-05572-8>.
- [26] Judkins, G., Myint, S., 2012. Spatial Variation of Soil Salinity in the Mexicali Valley, Mexico: Application of a Practical Method for Agricultural Monitoring. *Environ. Manage.* 50, 478–489. <https://doi.org/10.1007/s00267-012-9889-3>.
- [27] Kurbatova, A.I., Bouchhima, R.A., Grigorets, E.A. et al., 2021. Methodology for mapping soil salinity and halophyte cover using remote sensing data in Kerkennah, Tunisia. *Euro-Mediterr J. Environ. Integr.* 6, 51. **Error! Hyperlink reference not valid.**
- [28] Kauth, R.J., Thomas, G.S., 1976. The Tasseled Cap - A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by LANDSAT. *LARS Symp.* 159. http://docs.lib.purdue.edu/lars_symp/159.
- [29] Lopez Puga, J., Krzywinski, M., Altman, N., 2015. Points of significance: Bayesian statistics. *Nat. Methods* 12, 377–378. <https://doi.org/10.1038/nmeth.3368>.
- [30] Lhissou, R., El Harti, A., Chokmani, K., 2014. Mapping Soil Salinity in Irrigated Land Using Optical Remote Sensing Data. *Eurasian J. of Soil Sci.* 3, 82–88. <https://doi.org/10.18393/ejss.84540>.
- [31] Matinfar, H.R., Panah, S.K.A., Zand, F., Kamal Khodaei, K., 2013. Detection of soil salinity changes and mapping land cover types based upon remotely sensed data. *Arab. J. Geosci.* 6, 913–919. <https://doi.org/10.1007/s12517-011-0384-6>.

- [32] Meyer, L.H., Heurich, M., Beudert, B., Premier, J., Pflugmacher, D., 2019. Supplementary Materials: Comparison Between Landsat-8 and Sentinel-2 Data for Estimation of Leaf Area Index in Temperate Forests. *Remote Sens.* 11,1160. <https://doi.org/10.3390/rs11101160>.
- [33] Ngo, D., Ha, N.T.T., Dang, Q.T., Koike, K., Trong, N.M., 2019. Effective Band Ratio of Landsat 8 Images Based on VNIR-SWIR Reflectance Spectra of Topsoils for Soil Moisture Mapping in a Tropical Region. *Remote Sens.* 11, 716. **Error! Hyperlink reference not valid.**
- [34] Nouri, H., Borujeni, S.C., Alaghmand, S., Anderson, S.J., Sutton, P.C., Parvazian, S., Beecham, S., 2018. Soil Salinity Mapping of Urban Greenery Using Remote Sensing and Proximal Sensing Techniques; The Case of Veale Gardens within the Adelaide Parklands. *Sustainability* 10, 2826. <https://doi.org/10.3390/su10082826>.
- [35] Nguyen, K.A., Liou, Y.A., Tran, H.P., Hoang, P.P., Nguyen, T.H., 2020. Soil salinity assessment by using near-infrared channel and Vegetation Soil Salinity Index derived from Landsat 8 OLI data: a case study in the Tra Vinh Province, Mekong Delta, Vietnam. *Prog. Earth Planet. Sci.* 7, 1. <https://doi.org/10.1186/s40645-019-0311-0>.
- [36] Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., 2014. Modeling and Mapping of Soil Salinity with Reflectance Spectroscopy and Landsat Data Using Two Quantitative Methods (PLSR and MARS). *Remote Sens.* 6, 10813–10834. <https://doi.org/10.3390/rs61110813>.
- [37] Qian, T., Tsunekawa, A., Peng, F. et al., 2019. Derivation of salt content in salinized soil from hyperspectral reflectance data: A case study at Minqin Oasis, Northwest China. *J. Arid Land* 11, 111–122. <https://doi.org/10.1007/s40333-019-0091-9>.
- [38] Raftery, A.E., 1995. Bayesian model selection in social research. *Sociological Method.* 25 (25), 111–163. <https://doi.org/10.2307/271063>
- [39] Richards, L.A. (ed.) 1954. Diagnosis and improvements of saline and alkali soils. USDA. *Agriculture Handbook* 60. pp. 160.
- [40] Rubio, F.J. Genton, M.G., 2016. Bayesian linear regression with skew-symmetric error distributions with applications to survival analysis. *Stat. Med.* 35, 2441–2454. <https://doi.org/10.1002/sim.6897>.
- [41] Solangi, K.A., Siyal, A.A., Wu, Y., Abbasi, B., Solangi, F., Lakhari, I.A., Zhou, G., 2019. An Assessment of the Spatial and Temporal Distribution of Soil Salinity in Combination with Field and Satellite Data: A Case Study in Sujawal District. *Agronomy* 9, 869. <https://doi.org/10.3390/agronomy9120869>.
- [42] Scudiero, E., Corwin, D.L., Anderson, R.G., Skaggs, T.H., 2016. Moving Forward on Remote Sensing of Soil Salinity at Regional Scale. *Front. Environ. Sci.* 4, 1–5. <https://doi.org/10.3389/fenvs.2016.00065>.
- [43] Sahbeni, G., 2021. Soil salinity mapping using Landsat 8 OLI data and regression modeling in the Great Hungarian Plain. *SN Appl. Sci.* 3, 587. <https://doi.org/10.1007/s42452-021-04587-4>.
- [44] Sahana, M., Rehman, S., Patel, P.P., Dou, J., Hong, H., Sajjad, H., 2020. Assessing the degree of soil salinity in the Indian Sundarban Biosphere Reserve using measured soil electrical conductivity and remote sensing data-derived salinity indices. *Arabian J. Geosci.* 13, 1289. <https://doi.org/10.1007/s12517-020-06310-w>.
- [45] Samra, R.M.A., Ali, R.R., 2018. The development of an overlay model to predict soil salinity risks by using remote sensing and GIS techniques: a case study in soils around Idku Lake, Egypt. *Environ. Monit. Assess.* 190, 706. <https://doi.org/10.1007/s10661-018-7079-3>.
- [46] Shahid, S.A., Abdelfattah, M.A., Omar, S.A.S., Harahsheh, H., Othman, Y., Mahmoudi, H., 2010. Mapping and Monitoring of Soil Salinization Remote Sensing, GIS, Modeling, Electromagnetic Induction and Conventional Methods – Case Studies. *Proceedings of the International Conference on Soils and Groundwater Salinization in Arid Countries*, 59-97. Sultan Qaboos University.
- [47] Shahrayini, E., Noroozi, A.A., 2022. Modeling and Mapping of Soil Salinity and Alkalinity Using Remote Sensing Data and Topographic Factors: a Case Study in Iran. *Environ. Model. Assess.* <https://doi.org/10.1007/s10666-022-09823-8>.
- [48] Sultanov, M., Ibrakhimov, M., Akramkhanov, A. et al., 2018. Modelling End-of-Season Soil Salinity in Irrigated Agriculture Through Multi-temporal Optical Remote Sensing, Environmental Parameters, and In Situ Information. *PFG- J. Photogramm. Remote Sens. Geoinf. Sci.* 86, 221–233. <https://doi.org/10.1007/s41064-019-00062-3>.
- [49] Taghadosi, M.M., Hasanlou, M., Eftekhari, K., 2019. Retrieval of soil salinity from Sentinel-2 multispectral imagery. *Eur. J. Remote Sens.* 52 (1), 138–154. <https://doi.org/10.1080/22797254.2019.1571870>.
- [50] Taghadosi, M.M., Hasanlou, M., 2021. Developing geographic weighted regression (GWR) technique for monitoring soil salinity using sentinel-2 multispectral imagery. *Environ. Earth Sci.* 80, 75. <https://doi.org/10.1007/s12665-020-09345-0>.
- [51] Thiam, S., Villamor, G.B., Faye, L.C., Sène, J.H.B., Diwediga, B., Kyei-Baffour, N., 2021. Monitoring land use and soil salinity changes in coastal landscape: a case study from Senegal. *Environ.*

- Monit. Assess. 193, 259. <https://doi.org/10.1007/s10661-021-08958-7>.
- [52] Vietnam Ministry of Science, Technology and Environment. 2000. Vietnam Standard TCVN 6650:2000 Soil quality - Determination of water-soluble and acid-soluble sulfate.
- [53] Vietnam National Administration of Science and Technology Information. 2016. Salinity intrusion in Cuu Long Delta: Causes, Impacts and Responses.
- [54] Wang, F., Chen, X., Luo, G. et al., 2013. Detecting soil salinity with arid fraction integrated index and salinity index in feature space using Landsat TM imagery. *J. Arid Land* 5, 340–353. <https://doi-org/10.1007/s40333-013-0183-x>.
- [55] Willmott, C.J., Robeson, S.M., Kenji, M., 2012. A refined index of model performance. *Int. J. Climatol.* 32, 2088–2094.
- [56] Wang, N., Xue, J., Peng, J., Biswas, A., He, Y., Zhou Shi, Z., 2020. Integrating Remote Sensing and Landscape Characteristics to Estimate Soil Salinity Using Machine Learning Methods: A Case Study from Southern Xinjiang, China. *Remote Sens.* 12, 4118. <https://doi-org/10.3390/rs12244118>.
- [57] Wang, J., Ding, J., Yu, D., Teng, D., He, B., Chen, X., Ge, X., Zhan, Z., Wang, Y., Yang, X., Shi, T., Su, F., 2019. 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.* 707, 136092. <https://doi.org/10.1016/j.scitotenv.2019.136092>.
- [58] Yahiaoui, I., Douaoui, A., Zhang, Q. et al., 2015. Soil salinity prediction in the Lower Cheliff plain (Algeria) based on remote sensing and topographic feature analysis. *J. Arid Land* 7, 794–805. <https://doi.org/10.1007/s40333-015-0053-9>.
- [59] Yang, L., Huang, C., Liu, G. et al., 2015. Mapping soil salinity using a similarity-based prediction approach: A case study in Huanghe River Delta, China. *Chin. Geogr. Sci.* 25, 283–294. <https://doi.org/10.1007/s11769-015-0740-7>.