

Development of the Statistical Model for Monitoring Salinization in the Mekong Delta of Vietnam using Remote Sensing data and In-situ measurements

Nguyen Nguyen Vu ^{1*}, Le Van Trung ¹, Tran Thi Van ¹

¹ Ho Chi Minh City University of Technology, Vietnam National University Ho Chi Minh City
E-mail: nnvu2310@gmail.com (Nguyen Nguyen Vu), lvtrungbk@gmail.com (Le Van Trung),
tranthivankt@hcmut.edu.vn (Tran Thi Van)

Received: date; Accepted: date; Published: date

Abstract: This article presents the methodology for developing a statistical model for monitoring salinity intrusion in the Mekong Delta based on the integration of satellite imagery and in-situ measurements. We used Landsat 8 OLI & TIRS satellite data to establish the relationship between the planetary reflectance and the ground measured data in dry season 2014. The three spectral bands (Blue, Green, Red) and the the principal component band were used to obtain the most suitable models. The selected model showed a good correlation with the exponential function of the principal component band and the ground measured data ($R^2 > 0.8$). Simulation of the salinity distribution along the river shows that the intrusion of 4g/l salt boundary from the estuary to the inner field more than 50km. The developed model will be an active contribution, providing managers with adaptation and response solutions suitable for intrusion in the estuary as well as the inner field of the Mekong Delta.

Keywords: Landsat 8; Mekong Delta; Remote Sensing; Salinization; Statistical model

1. Introduction

The Mekong Delta is a major agricultural area, contributing importantly to export as well as the food supply to the people of Vietnam. However, this is a hollow terrain area, located in the downstream Mekong basin, is being heavily impacted by sea level rise, climate change. This has made drought and salinization more severe and is expected to continue with negative developments in the near future [1].

The traditional measurement of water salinity requires in-situ sampling, which is a costly and time-consuming effort. Because of these limitations, it is impractical to cover the whole water body or obtain frequent repeat sampling at a site. This difficulty in achieving successive water salinity sampling becomes a barrier to water salinity monitoring and forecasting [2].

Remote sensing techniques have the potential to overcome these limitations by providing an alternative means of studying and monitoring water salinity over a wide range of both temporal and spatial scales. Several studies have confirmed that remote sensing can meet the demand for the large sample sizes required for water salinity studies conducted on the watershed scale [3]. Hence, a significant amount of research has been conducted to develop remote sensing methods and indices. These methods ranged from semi-empirical techniques to analytical methods for estimating and producing quantitative water salinity maps [4].

This research aimed at assessing the use of remote sensing data for estimating of the Mekong Delta water salinity using statistical techniques. The work was based on developing an algorithm for estimation of water salinity using the Landsat-8 OLI & TIRS bands reflectance.

2. Materials and Methods

2.1. Study area

The study area was located in the Mekong Delta of Vietnam, including main estuaries: Cua Tu, Cua Dai, Ba Lai, Ham Luong, Co Chien, Cung Hau, Dinh An, Tran De. The area lies between 105°46'16" and 106°05'10" E, 9°18'44 "and 10°21'48" N. The total area was approximately 8407 km² (Figure 1).

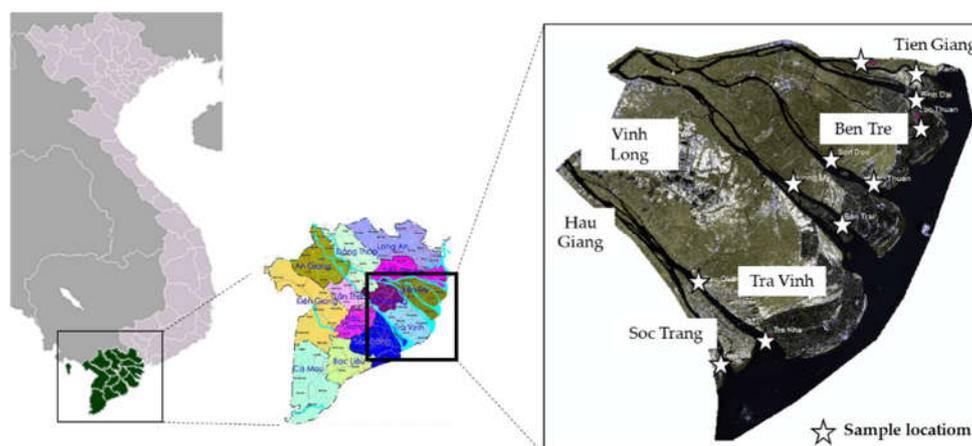


Figure 1. Study area and sample location

2.2. Satellite and ground measured data

We used 5 scenes of Landsat-8 OLI&TIRS multispectral image data for the different dates in 2014 (22/02, 10/3, 26/3, 27/4, 13/5). The size of each scene was 185km x 180 km. The spatial resolution for each of these spectral bands was 30 m. The satellite data was downloaded from the United States Geological Survey in GeoTIFF format with the Level 1T correction. The Level 1T is the standard terrain correction in which systematic radiometric and geometric accuracy is provided using the ground control points and the topographic accuracy is obtained by using the digital elevation model. The scenes of Landsat-8 OLI&TIRS were selected on the basis of the least cloud cover, and at the sampling days. In total, we used ground measured data at eleven sampling locations of the study area in February 2014 to – May 2014 to develop and validate models using the planetary reflectance of the scenes of Landsat-8. The sampling locations and the dates for the ground measured data and Landsat-8 scenes are given in Table 1.

Table 1. Sample location in the study area

No.	Name	River	Province	Longitude	Latitude
1	Hoa Binh	Cua Tieu	Tien Giang	106°36' E	10°18' N
2	Vam Kenh	Cua Tieu	Tien Giang	106°43' E	10°16' N
3	Binh Dai	Cua Dai	Ben Tre	106°43' E	10°12' N
4	Loc Thuan	Cua Dai	Ben Tre	106°43' E	10°10' N
5	Son Doc	Ham Luong	Ben Tre	106°30' E	10°03' N
6	Huong My	Cung Hau	Ben Tre	106°23' E	9°59' N
7	An Thuan	Ham Luong	Ben Tre	106°36' E	9°59' N
8	Ben Trai	Co Chien	Ben Tre	106°31' E	9°53' N
9	Cau Quan	Hau	Tra Vinh	106°09' E	9°43' N
10	Tra Kha	Hau	Tra Vinh	106°20' E	9°35' N
11	Tran De	Cua Tran De	Soc Trang	106°12' E	9°31' N

2.3. Image processing

The satellite scenes were processed to make them workable for the purpose of this research. The operations applied for processing are briefly explained in the following sub-sections:

2.3.1. Conversion of digital numbers into spectral radiance

In the first step, we converted the raw digital numbers of all the Landsat-5 TM images into spectral radiance using Eq. (1) as follows:

$$L_{\lambda} = M_L \times Q_{cal} + A_L \tag{1}$$

Where L_{λ} = spectral radiance at the sensor’s aperture ($Wm^{-2} sr^{-1} \mu m^{-1}$), Q_{cal} = quantized calibrated pixel value, M_L = radiance_mult_band_x, A_L = radiance_add_band_x.

2.3.2. Conversion from spectral radiance to planetary reflectance

In the second step, the spectral radiance was converted into planetary reflectance using Eq. (2) as follows:

$$p_{\lambda} = (M_p + Q_{cal} + A_p) / \cos(\theta_{sz}) \tag{2}$$

Where p_{λ} = unitless planetary reflectance, Q_{cal} = quantized calibrated pixel value, M_p = reflectance_mult_band_x, A_p = reflectance_add_band_x, θ_{sz} = solar zenith angle (degree).

2.3.3. Normalized Difference Vegetation Index (NDVI)

We calculated normalized difference vegetation index (NDVI) – a measure of vegetation greenness - using Eq. (3) [5] as follow:

$$NDVI = (p_{NIR} - p_{Red}) / (p_{NIR} + p_{Red}) \tag{3}$$

Where p_{NIR} = reflectance of NIR band, p_{Red} = reflectance of Red band.

Such NDVI calculations were performed over study area in all scenes of Landsat-8, the negative NDVI values (between 0 and -1) indicated the presence of water in the pixels whereas positive NDVI values showed the possible contamination due to other landuses [6]. In case of a positive NDVI value for any sampling site pixel, we considered the reflectance value of a nearest neighboring water pixel.

2.4. Investigating the correlation between the Landsat-8 satellite image data and the in-situ salinity measurements

We investigated the correlation between the spectral bands of Landsat-8 images (except panchromatic band and cirrus band) and the in-situ salinity measurements by calculating the Pearson correlation coefficient (Table 2).

Table 2. Pearson correlation coefficient (r) between the spectral bands of Landsat-8 images (except panchromatic band and cirrus band) and the in-situ salinity measurements

	Pearson correlation coefficient (r)
Band 1 – Coastal Aerosol	0.67790
Band 2 – Blue	0.785104
Band 3 – Green	0.81233
Band 4 – Red	0.82786
Band 5 – Near Infrared (NIR)	-0.58903
Band 6 – SWIR1	-0.4972
Band 7 – SWIR2	-0.28816
Band 10 – Thermal Infrared 1 (TIR1)	-0.46747
Band 11 – Thermal Infrared 2 (TIR2)	-0.36758

Table 2 shows the Coastal aerosol, Blue, Green and Red band correlate good with the in-situ salinity measurements ($r = 0.68 - 0.83$). Therefore, these spectral bands can be used to develop the model for water salinity.

Principal component analysis technique is an algorithm that creates the principal components (PCs) to remove the overlap data on the image, highlighting spectral characteristics of different objects on the Earth’s surface [7]. We also investigated the correlation between the principal

components analyzed from the original Landsat-8 image bands (Band 1, 2, 3, 4) and the in-situ salinity measurements. The first principal component correlated most strongly with the in-situ salinity measurements ($r = 0.86$) (Table 3).

Table 3. Pearson correlation coefficient (r) between the principal components (PCs) and the in-situ salinity measurements

	Pearson correlation coefficient (r)
PC1	0.860305
PC2	0.562152
PC3	-0.503505
PC4	-0.26186

2.5. Developing models for water salinity from planetary reflectance of Landsat-8 data

We used regression analysis technique to developed 8 individual empirical models in determining salinity as a function of the spectral bands of Blue, Green, Red and PC1 band. We used 40 data records (Landsat-8 as well as the ground data) for the development of models to obtain salinity from the planetary reflectance. The remaining 15 data records were used to validate the selected best models by calculating the root mean square error (RMSE). In all these models, the in-situ salinity measurements were the dependent variable whereas the spectral bands and PC1 band were the independent variables. The strength of correlations between the planetary reflectance and in situ salinity for the development and validation of the models were obtained based on co-efficient of determination (R^2) and RMSE. On the basis of R^2 values, RMSE, we identified the significant empirical models for salinity.

2.6. Spatial analysis for the study area

The NDVI was calculated for the study area to extract the pixels with water and remove the pixels contaminated with other land use types. We selected the most suitable empirical models for salinity on the basis of R^2 values. The selected salinity model was applied on the scenes of Landsat-8 to obtain the spatial distribution along the study area. The salinity values were divided into four classes which are: (1) Low salinity: less than 4 g/l, (2) Moderate salinity: 4–7 g/l; (3) High salinity: 7–10 g/l, (4) Severely salinity: greater than 10 g/l.

3. Results and Discussion

3.1. Empirical models for determining salinity classes

We developed 8 empirical models for determining salinity as given in Table 4. These empirical models could be used to obtain the spatial distribution of salinity using the planetary reflectance of bands. Among these significant models, R^2 was higher (0.73–0.82) for the models with the exponential function (model no. 1, 2, 3, 5 in Table 4) whereas it was lower (0.62–0.74) in the models with the linear function (model no. 4, 6, 7, 8 in Table 4). The best model was the use of the principal component PC1 (model no. 1 in Table 4).

Table 4. Models developed for water salinity for the study area using the spectral bands (Blue, Green, Red) and the principal component PC1 of Landsat-8 image data

Model no.	Model ¹	R^2	RMSE
1	$S = 0.0014.e^{0.0006.(PC1)}$	0.8195	1.592
2	$S = 0.0001.e^{0.0011.(G)}$	0.7859	2.029
3	$S = 0.0131.e^{0.0007.(R)}$	0.7510	2.073
4	$S = 0.0042.(PC1) - 54.015$	0.7401	2.317

5	$S = 0.00001.e^{0.0012.(B)}$	0.7269	2.651
6	$S = 0.0053.(R) - 41.889$	0.6854	2.761
7	$S = 0.0081.(G) - 74.6$	0.6599	3.601
8	$S = 0.0095.(B) - 94.186$	0.6164	4.081

¹ S = Salinity value, G = Green band, B = Blue band, R = Red band, PC1 = principal component band 1

The scatter plot of model no. 1 is shown in Figure 2. This figure shows that the selected model showed the best correlation with the exponential function of the principal component PC1 band and the ground measured data with $R^2 = 0.8195$ and $RMSE = 1.592$. Figure 3 revealed that the the predicted values of salinity matched the observed ones. The value of coefficient of determination computed for the regression between measured and predicted salinity was 0.8948 and $RMSE = 0.802$, they found that both field and calculated salinity data showed a good agreement indicating a good fitting.

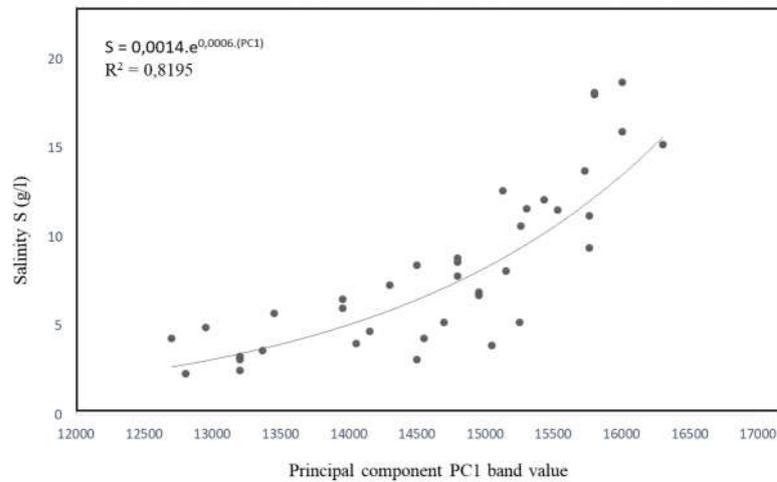


Figure 2. Model for obtaining salinity using the principal component PC1 band of the study area

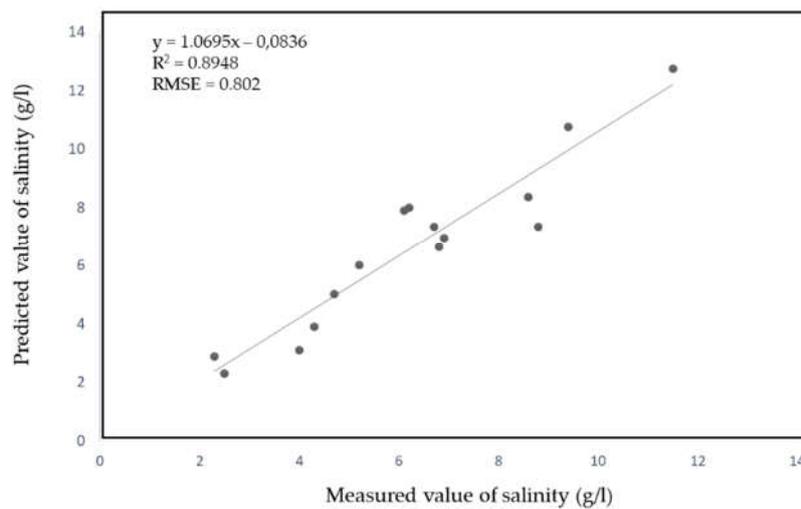


Figure 3. Measured against predicted value of salinity at the study area

3.2. Application of model for spatial analysis

We applied the best model (model no. 1 in Table 4) to quantify the water salinity into the spatial distribution map for the study area. Figure 4 shows four salinity classes which are: (1) Low salinity: less than 4 g/l, (2) Moderate salinity: 4–7 g/l; (3) High salinity: 7–10 g/l, (4) Severely salinity: greater than 10 g/l in February 2014.

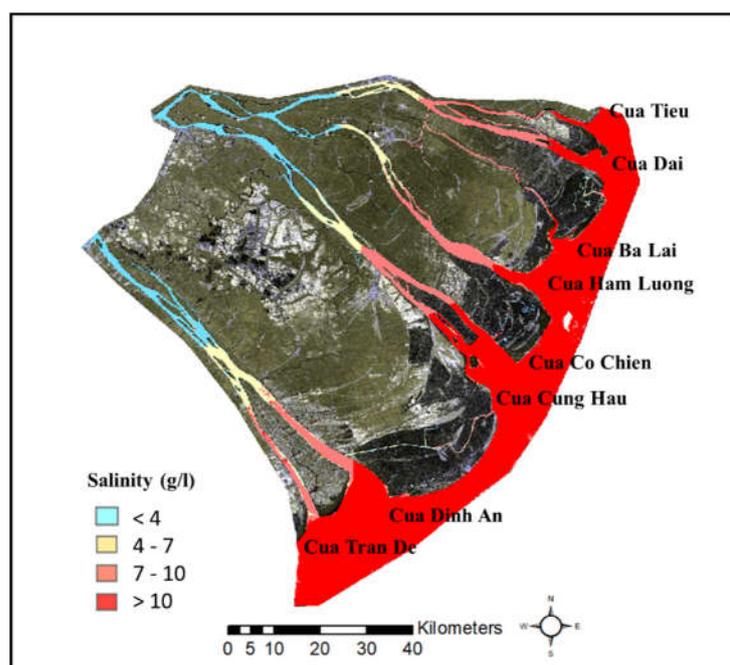


Figure 4. Map of salinity distribution in the study area (February 2014)

In February 2014, the length of saline intrusion of 4g/l salt boundary into the estuaries: Cua Tieu, Cua Dai, Cua Ham Luong more than 50 km, Table 5. shows the salt intrusion length of 4g/l salt boundary into the inner field from eight large estuaries in the study area.

Table 5. The length of saline intrusion of 4 g/l salt boundary into the large estuaries in the study area (In February 2014)

Estuaries	Length of saline intrusion
Cua Tieu	52 km
Cua Dai	56 km
Cua Ba Lai	47 km
Cua Ham Luong	50 km
Cua Co Chien	47 km
Cua Cung Hau	48 km
Cua Dinh An	38 km
Cua Tran De	43 km

4. Conclusions

The purpose of this research was to use in-situ salinity measurements combined with Landsat-8 OLI & TIRS data for deriving water salinity in the Mekong River Delta of Vietnam for sustainable management. Correlation and regression models were developed between each of the salinity measurements and the reflectance of image data. The current research indicated that despite the variation in statistical response, there existed a good correlation between selected salinity and Landsat-8 data in the mekong River Delta of Vietnam. The efficiency of the proposed algorithms was investigated based on the values of determining coefficient and RMSE. As there was a little variation between Landsat-8 predicted values and measured concentrations, this research proved that Landsat-8 data were useful in quantifying water salinity dynamics and preparing digital map. The models and methods used in this research for retrieval of water salinity from Landsat-8 are promising and can be frequently applied to the Mekong Delta for sustainable water quality monitoring.

Funding: This research was funded by Ho Chi Minh City University of Technology – VNU-HCM under grant number TNCS-MTTN-2017-11.

References

1. Binh, N. Vulnerability and Adaptation to Salinity Intrusion in the Coastal Province of Tra Vinh. *Disaster Risk Reduction and Climate Change Adaptation* **2010**.
2. Khalil, T. Integrated Field Study, Remote Sensing and GIS Approach for Assessing and Monitoring Some Chemical Water Quality Parameters in Bardawil Lagoon, Egypt. *International Journal of Innovative Research in Science, Engineering and Technology* **2016**, vol.5, pp.14656-14669.
3. Choubey, V.K. Monitoring surface water conductivity with Indian remote sensing satellite data: a case study from central India. *Hydrological, Chemical and Biological Processes of Transformation and Transport of Contaminants in Aquatic Environments* **1994**, vol.219, pp.317-326.
4. Abdelmalik, K. Role of statistical remote sensing for Inland water quality parameters prediction. *The Egyptian Journal of Remote Sensing and Space Sciences* **2016**, pp. 1-8.
5. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring vegetation. *Remote Sensing Environments* **1979**, vol.8, pp.127-150.
6. Weier, J.; Herring, D. Measuring vegetation (NDVI&EVI). *NASA Earth Observatory* 1999.
7. Howari, F.M. Spectral Properties of Salt Crusts Formed on Saline Soils. *Journal Environments* **2002**, vol.31, pp. 1453-1461.



© 2018 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).