

NGHIÊN CỨU CHIẾT TÁCH THÔNG TIN ĐẤT NHIỄM MẶN TỪ DỮ LIỆU SENTINEL-2, THỬ NGHIỆM TẠI TỈNH ĐỒNG NAI

Chu Xuân Huy¹, Nguyễn Minh Ngọc¹, Đinh Ngọc Đạt¹, Lê Thu Thủy¹
Hoàng Hải¹, Bùi Quang Huy¹, Trần Tân Phong², Tống Thị Huyền Ái¹
Trần Thị Hồng Minh³, Lộc Đức Hiệp⁴

¹ Viện Công nghệ Vũ trụ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

² Sở Khoa học và Công nghệ tỉnh Đồng Nai

³ Trường Đại học Tài nguyên và Môi trường Hà Nội

⁴ Học viện Cảnh sát nhân dân

Tóm tắt

Đất nhiễm mặn là một trong các nguyên nhân có ảnh hưởng rõ rệt đến nông nghiệp tại nhiều quốc gia, vùng lãnh thổ trên thế giới. Nghiên cứu này sẽ xây dựng bản đồ đất nhiễm mặn tại khu vực huyện Long Thành và Nhơn Trạch, thuộc tỉnh Đồng Nai, sử dụng dữ liệu Sentinel-2 được chụp vào mùa khô 2020 - 2021. Các chỉ số được sử dụng là chỉ số độ mặn (Salinity Index, SI) và chỉ số tăng cường thực vật (Enhanced Vegetation Index, EVI). Số liệu đo đạc thực tế thu thập vào thời điểm đó được sử dụng để xây dựng mô hình phù hợp cho khu vực nghiên cứu trên cơ sở tương quan với các chỉ số đã sử dụng. Trên cơ sở các tính toán tương quan bằng mô hình hồi quy tuyến tính đa biến với hệ số tương quan xấp xỉ 0,75. Đồng thời, bộ dữ liệu tháng 3 năm 2021 được sử dụng để thành lập bản đồ đất nhiễm mặn vì đây là thời điểm đất mặn thể hiện rõ ràng nhất.

Từ khóa: Sentinel-2; Đất mặn; Đồng Nai; Hồi quy tuyến tính đa biến.

Abstract

Study on soil salinity by using Sentinel-2 imagery data: a case study in Dong Nai Province, Vietnam

Soil salinity is one of the causes affecting agriculture in many countries and territories around the world. This study aims to develop the saline soil map of Long Thanh and Nhon Trach districts, Dong Nai province, using Sentinel-2 data taken in the dry season 2020 - 2021. Salinity Index (SI) and Enhanced Vegetation Index (EVI) were used in this study. The actual measurement data and satellite data were collected at the same time in order to build a suitable model of index correlation at the study area. On the basis of correlation calculations by a multivariable linear regression model with a correlation coefficient of approximately 0.75. The data set collected in March 2021 was used to build a saline soil map in the study area due to salinity soils were clearly detected at this time.

Keywords: Sentinel-2; Soil salinity; Dong Nai Province; Multiple linear regression.

1. Đặt vấn đề

Suy giảm chất lượng đất đã và đang trở thành vấn đề toàn cầu, đặc biệt là đối với các vùng ven biển, khu vực tập trung đến 50 % dân số thế giới (tính từ bờ biển vào sâu trong lục địa 100 km), đây cũng là nơi chịu ảnh hưởng mạnh mẽ của một trong những hiện tượng suy giảm chất lượng đất đó là đất mặn.

Đất mặn gây ra các vấn đề nghiêm trọng trong công tác quản lý và sử dụng, gây ảnh hưởng không nhỏ đến khả năng canh tác của người dân sống trên đó. Bản chất và đặc tính của các loại đất mặn cũng rất đa dạng, do đó đòi hỏi cách tiếp cận, nghiên cứu cụ thể nếu muốn quản lý, sử dụng để khai thác các nguồn lợi từ đất mặn. Vì vậy, sự phát triển của đất mặn cần được theo dõi thường xuyên để đảm bảo cập nhật về mức độ, mức độ nghiêm trọng, sự phân bố không gian, tính chất và sự phát triển của chúng.

Lập bản đồ đất mặn thường được thực hiện bằng quan sát trực quan và sử dụng các phép đo trong phòng thí nghiệm. Tuy nhiên, quan sát chỉ cung cấp thông tin định tính và phương pháp phòng thí nghiệm thường tốn thời gian, chi phí và nhiều công sức. Sự ra đời của các vệ tinh viễn thám đã mang lại một phương pháp nhanh chóng và hiệu quả về kinh tế để giám sát và lập bản đồ đất mặn; với rất nhiều vệ tinh và cảm biến cung cấp thông tin hữu ích trong việc phát hiện, giám sát đất mặn như dữ liệu đa thời gian của LANDSAT, SPOT, IKONOS, EO-1, IRS, Terra-ASTER,... có độ phân giải từ trung bình đến cao cũng như các cảm biến siêu phổ. Sentinel-2 là một trong số các vệ tinh đa phổ cung cấp dữ liệu miễn phí có độ

phân giải không gian 10 m đang được ứng dụng rộng rãi và phù hợp cho những khu vực có diện tích đối tượng đồng nhất nhỏ. Tuy nhiên vẫn còn ít các nghiên cứu sử dụng dữ liệu ảnh này để chiết tách thông tin đất mặn.

2. Khu vực nghiên cứu

Khu vực được quan tâm nằm dọc theo sông Đồng Nai thuộc hai huyện Long Thành và Nhơn Trạch, tỉnh Đồng Nai. Đây là khu vực tập trung nhiều diện tích trồng lúa trong tỉnh nhưng cũng là khu vực có diện tích đất nhiễm mặn ngày càng tăng do chịu ảnh hưởng của sự xâm nhập mặn từ biển theo Sông Đồng Nai lên đến gần thành phố Biên Hòa. Tại đây hoạt động sản xuất lúa khá phức tạp, tại các xã chịu ảnh hưởng mạnh của hiện tượng xâm nhập mặn như Phước Khánh, Phú Hội, Phú Đông của huyện Nhơn Trạch chỉ có một đến hai vụ lúa, thậm chí tại khu vực xã Phước Khánh, các diện tích lúa được trồng cũng có sự khác nhau về giống lúa chịu mặn là giống 3 tháng và giống 6 tháng; tuy nhiên tại một số xã như Long Phước của huyện Nhơn Trạch, do nước được cung cấp đầy đủ nên một năm có 3 vụ lúa. Điều này gây ra khá nhiều thách thức trong công tác theo dõi diễn biến đất mặn trong khu vực từ dữ liệu ảnh viễn thám, đòi hỏi cần có loại dữ liệu có độ phân giải không gian đủ chi tiết theo dõi các vùng diện tích nhỏ; đồng thời cũng có độ phân giải thời gian cao để có dữ liệu vào nhiều thời điểm khác nhau.



Hình 1: Vị trí khu vực nghiên cứu

3. Dữ liệu sử dụng

Dữ liệu ảnh vệ tinh thu thập được là dữ liệu ảnh Sentinel-2 vào hai thời điểm đầu và cuối mùa khô tại khu vực nghiên cứu và gần với thời điểm tiến hành khảo sát đo đạc là ngày 03/12/2020 và 28/3/2021, bao gồm các cảnh ảnh có phiên hiệu: T48PXS, T48PXT, T48PYS.

Các dữ liệu đo đạc được thu thập qua các chuyến khảo sát đầu và cuối mùa khô, đợt 1 từ 29 - 30/11/2020 với 50 điểm đo đạc (được sử dụng để xây dựng mô hình) và đợt hai là từ 22 - 23/3/2021 với 35 điểm đo đạc (được sử dụng để kiểm chứng mô hình), tại khu vực các xã thuộc hai huyện Long Thành và Nhơn Trạch. Tại mỗi điểm đo, độ mặn của đất được thu thập bằng

thiết bị đo độ mặn dùng trong nông nghiệp FieldScout. Các mẫu này đều được lấy trên cơ sở căn cứ vào bản đồ thổ nhưỡng và hiện trạng sử dụng đất trong khu vực [5, 6].

Bảng 1. Dữ liệu ảnh và số liệu đo

Dữ liệu ảnh	Dữ liệu đo đạc
03/12/2020	29 - 30/11/2020
28/3/2021	22 - 23/3/2021

4. Phương pháp nghiên cứu

Dữ liệu ảnh vệ tinh trước khi tính toán được hiệu chỉnh khí quyển để đưa về giá trị phản xạ phổ của các đối tượng trên mặt đất.

Hiện nay có rất nhiều các chỉ số độ mặn được sử dụng, tuy nhiên nhóm tác giả chỉ chọn một số chỉ số phù hợp với các kênh phổ của Sentinel-2 như sau:

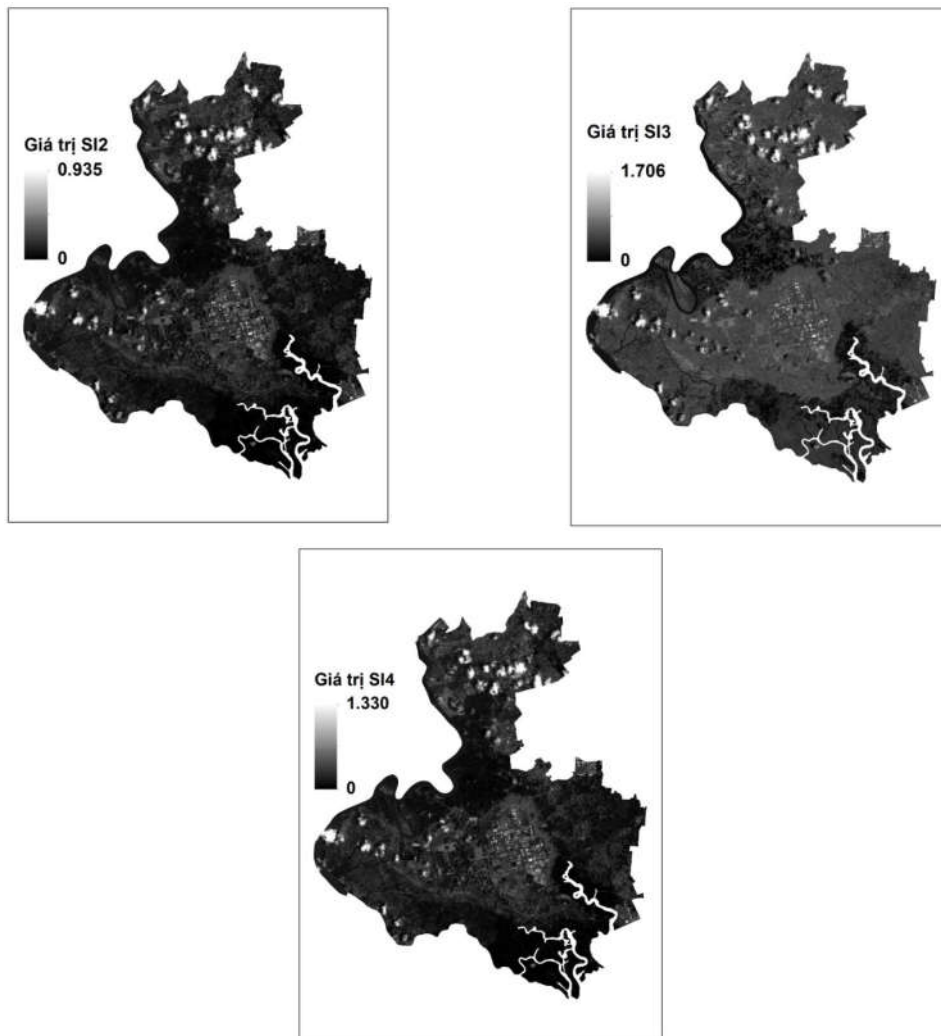
Bảng 2. Các chỉ số dùng trong nghiên cứu

Chỉ số	Công thức	Tham khảo
SI1	$SI = \sqrt{B} \times R$	[8]
SI2	$SI = \sqrt{G} \times R$	[2]
SI3	$SI = \sqrt{G^2 + R^2 + NIR^2}$	[2]
SI4	$SI = \sqrt{G^2 + R^2}$	[2]
SI5	$SI = \frac{B}{R}$	[10]
SI6	$SI = \frac{B - R}{B + R}$	[10]
SI7	$SI = \frac{G \times R}{B}$	[10]

Chỉ số	Công thức	Tham khảo
SI8	$SI = \frac{B \times R}{G}$	[10]
SI9	$SI = \frac{NIR \times R}{G}$	[10]
SI10	$SI = \sqrt{G + R}$	[10]
EVI	$EVI = 2.5 \times \frac{NIR - R}{NIR + c_1 \times R - c_2 \times B + L}$	[4]

(Trong công thức tính EVI, $c_1 = 6$; $c_2 = 7,5$; $L = 1$)

Các chỉ số này được chọn theo các tài liệu liên quan dựa trên hiệu quả của chúng trong việc lập bản đồ độ mặn của đất. Do đó, các phân tích tương quan giữa từng chỉ số với giá trị độ dẫn điện (Electrical Conductivity, EC) được thực hiện bằng phương pháp hồi quy tuyến tính để khảo sát mối tương quan tốt nhất với các giá trị độ mặn của đất được lấy mẫu.



Hình 2: Kết quả tính toán các chỉ số SI2, SI3, SI4, SI9

Tương quan giữa độ dẫn điện thông qua các mẫu thu thập ngoài thực địa và giá trị tính toán từ các kênh phổ sẽ được tính toán để tìm ra mối quan hệ giữa các giá trị tính toán từ ảnh và giá trị đo đạc thực tế. Phương pháp lựa chọn là phương pháp hồi quy tuyến tính được sử dụng khá rộng rãi trong nghiên cứu phân bố không gian. Tuy nhiên, phương pháp hồi quy tuyến tính đơn biến khó có thể đáp ứng được yêu cầu về độ chính xác [1, 7, 11], nên nhóm tác giả lựa chọn phương pháp hồi quy tuyến tính đa biến. Trong đó, sử dụng thuật toán mô hình trung bình Bayesian, BMA (Bayesian Model Average) kết hợp tiêu chí thông tin Bayesian (Bayesian Information Criterion, BIC) và Akaike

(Akaike Information Criterion, AIC) để lựa chọn mô hình ước tính chính xác nhất, sau đó mới xét đến hệ số xác định R² giữa giá trị đo và giá trị ước tính [9]. Công tác thực hiện hồi quy tuyến tính đa biến theo tiêu chí BIC và AIC được thực hiện trên phần mềm R, với các bước thực hiện bao gồm: chuẩn hóa dữ liệu đầu vào, lựa chọn thuật toán, và lựa chọn mô hình tối ưu.

Kết quả sau đây (Bảng 3) thể hiện các mô hình tốt nhất trong các mô hình ước tính chỉ số độ mặn sử dụng tiêu chí BIC. Theo đó, mô hình 1 (Model 1) là mô hình tốt nhất với BIC = -35,759; R² = 0.616 và Post prob = 0.273 [2] với các biến độc lập là SI2, SI3, SI4, SI9. Và mô hình ước tính là:

$$EC_{\text{ước tính}} = 11.668 - 697.087*SI2 - 98.301*SI3 + 511.201*SI4 + 86.717*SI9$$

Bảng 3. Kết quả lựa chọn mô hình tối ưu dựa trên thuật toán BMA với tiêu chí thông tin BIC

	p!=0	EV	SD	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	100.0	1.359e+01	4.426	11.668	15.791	10.697	11.481	13.803
SI1	7.0	-1.727e+00	30.245	.	.	.	.	.
SI2	59.1	-3.949e+02	397.555	-697.087	.	-786.402	-714.433	-711.581
SI3	100.0	-8.558e+01	19.244	-98.301	-70.142	-87.427	-100.839	-97.157
SI4	58.4	2.886e+02	294.047	511.201	.	575.738	528.126	519.378
SI5	6.7	3.388e-02	2.609	.	.	.	.	-2.580
SI6	6.5	-1.996e-03	3.781	.	.	.	.	.
SI7	7.3	-1.551e-01	9.703	.	.	.	.	.
SI8	6.3	-2.768e-01	7.798	.	.	.	-8.740	.
SI9	100.0	7.470e+01	19.779	86.717	57.969	81.739	90.049	84.582
SI10	6.4	-6.325e-02	1.762	.	.	.	.	.
EVI	8.2	-2.483e-01	2.765	.	.	-6.504	.	.
<i>nvar</i>				4	2	5	5	5
<i>R2</i>				0.616	0.548	0.621	0.617	0.617
<i>BIC</i>				-35.759	-34.926	-32.515	-31.856	-31.852
<i>Post prob</i>				0.273	0.180	0.054	0.039	0.039

Nghiên cứu

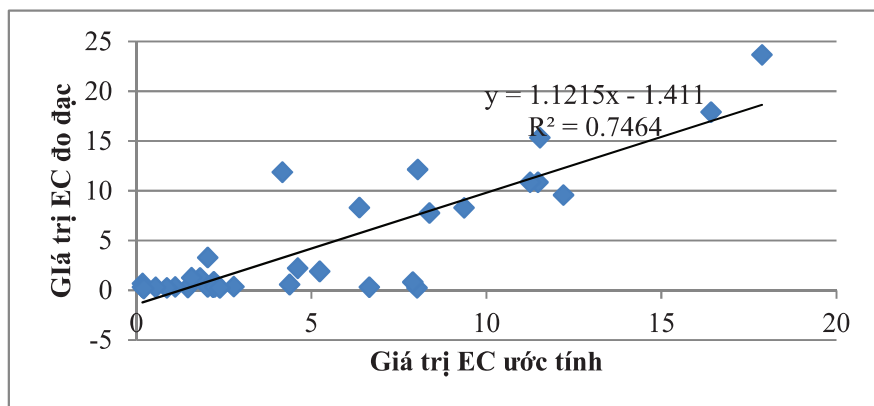
Tương tự như vậy đối với tiêu chí thông tin AIC, nhóm tác giả cũng đã tìm ra mô hình tối ưu để ước tính chỉ số độ mặn có tương quan cao nhất với giá trị đo đạc. Các tính toán được thể hiện trong Bảng 4 sau:

Bảng 4. Kết quả lựa chọn mô hình tối ưu dựa trên thuật toán BMA với tiêu chí thông tin AIC

Residuals					
Min	1Q	Median	3Q	Max	
-14.3467	-1.5275	-0.1436	1.5506	11.8853	
Coefficients					
	<i>Estimate</i>	<i>Std. Error</i>	<i>t value</i>	<i>Pr (> t)</i>	**
<i>(Intercept)</i>	11.668	3.428	3.404	0.00133	**
<i>SI2</i>	-697.087	240.789	-2.895	0.00565	***
<i>SI3</i>	-98.301	13.553	-7.253	2.69e-09	**
<i>SI4</i>	511.201	182.095	2.807	0.00715	***
<i>SI9</i>	86.717	13.860	6.257	9.40e-08	
Signif. codes	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
<i>Residual standard error:</i> 4.795 on 49 degrees of freedom					
<i>Multiple R-squared:</i> 0,6162, <i>Adjusted R-squared:</i> 0.5849					
<i>F- statistic:</i> 19.67 on 4 and 49 DF, <i>p-value:</i> 1.039e-09					

Kết quả thu được cũng cho thấy với các biến SI2, SI3, SI4, SI9 giá trị tương quan R^2 cũng gần tương đồng với tiêu chí thông tin BIC.

Sau khi lựa chọn được mô hình ước tính chỉ số độ mặn từ dữ liệu ảnh Sentinel-2, nhóm tác giả tiến hành kiểm chứng độ chính xác của mô hình. Kết quả cho thấy, tương quan giữa giá trị ước tính và giá trị đo là khá tốt, với $R^2 \approx 0,75$ (Hình 3).



Hình 3: Tương quan giữa kết quả đo đạc và tính toán từ ảnh

5. Kết quả và thảo luận

Sau khi mô hình được kiểm chứng và đạt kết quả tốt, nhóm tác giả đã thực hiện thành lập bản đồ đất nhiễm mặn theo phương pháp nền chất lượng với cấp độ mặn được chia theo tiêu chuẩn của Tổ chức Nông lương Liên hiệp quốc FAO [3], cụ thể như sau:

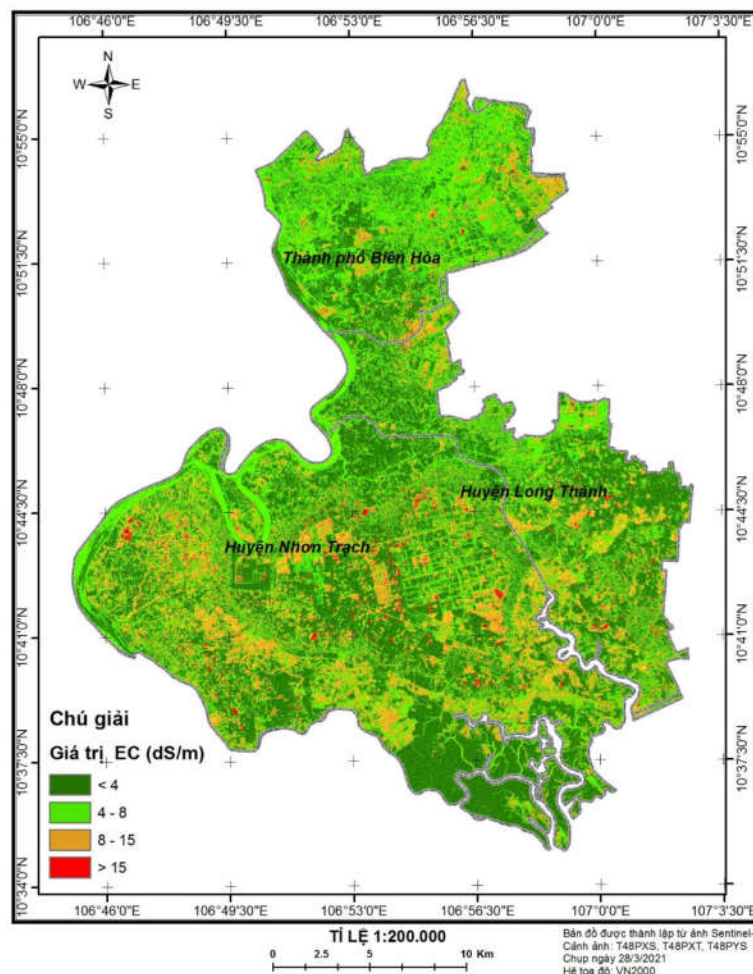
Bảng 5. Cấp độ mặn theo FAO

STT	Mức độ mặn	Giá trị EC (dS/m)
1	Không mặn	< 4
2	Mặn nhẹ	4 - 8
3	Mặn vừa	8 - 15
4	Mặn nặng	> 15

Trên bản đồ tại Hình 4 có thể thấy, các khu vực chịu ảnh hưởng mạnh nhất

là phía Nam của khu vực nghiên cứu, thuộc các xã như Phú Hữu, Phú Đông của huyện Nhơn Trạch, đất trồng lúa bị mặn nhiều (nằm trong khoảng 8 - 15 dS/m) nên sản lượng thấp, một năm chỉ có 1 vụ nhưng có một số khu vực bị mặn nặng, với giá trị thường thấy 17 - 18 dS/m; bên cạnh đó xã Phước Khánh là khu vực chịu ảnh hưởng khá mạnh, đất bị mặn (giá trị thường gặp là khoảng 16 - 16,5 dS/m) không thể trồng được lúa, phải bỏ hoang, chỉ có cỏ lác. Tuy vậy, xã Vĩnh Thanh lại là nơi có diện tích trồng lúa tương đối lớn, do địa hình cao, ít chịu ảnh hưởng như xã Phước Khánh. Ngược lên phía Bắc dọc theo Sông Đồng Nai, độ mặn của đất cũng giảm dần cho ít chịu ảnh hưởng của hiện

tượng xâm nhập mặn. Xã Đại Phước nằm phía trên xã Phú Hữu nhưng đất lại ít mặn hơn hẳn, diện tích đất trồng lúa còn khá nhiều. Lên đến các xã Long Tân, Phú Hội của huyện Nhơn Trạch và xã Tam An của huyện Long Thành, độ mặn của đất xuống khá thấp, chỉ khoảng từ 1 - 2 dS/m, do đó diện tích lúa nơi đây rất lớn và mùa vụ cũng tăng hơn từ 2 - 3 vụ /năm. Sâu vào phía Đông của khu vực nghiên cứu, đất gần như không bị mặn, lượng nước được cung cấp đầy đủ nên những khu vực trồng lúa tại các xã như Long An của huyện Long Thành, Hiệp Phước, Phước Thiện của huyện Nhơn Trạch có thể canh tác 3 vụ/năm.



Hình 4: Bản đồ phân bố đất nhiễm mặn tại một số khu vực thuộc tỉnh Đồng Nai

Khu vực nghiên cứu là nơi có nhiều loại hình sử dụng đất, tuy nhiên nhóm tác giả mới chỉ tập trung vào đất trồng lúa - là đối tượng chịu ảnh hưởng nhiều nhất của đất mặn. Hơn thế nữa, một số khu vực đất đã bị mặn đến mức không thể trồng lúa hoặc cho sản lượng thấp, người dân đã để hoang hoặc bắt đầu chuyển đổi mục đích sử dụng để khai thác đất hiệu quả hơn.

5. Kết luận

Độ mặn của đất là thông số thay đổi theo thời gian, chịu ảnh hưởng mạnh mẽ của quá trình xâm nhập mặn của nước biển, lượng mưa, độ ẩm đất và quá trình bốc hơi; hậu quả của đất mặn là suy thoái đất, giảm sản lượng cây trồng,... Tại khu vực nghiên cứu, nguyên nhân chính gây ra đất mặn là quá trình xâm nhập mặn đặc biệt là vào mùa khô khi lượng mưa ít. Bỏ qua phương pháp hồi quy tuyến tính đơn biến do tương quan giữa kết quả đo đạc và tính toán từ ảnh thấp, nhóm tác giả đã đưa ra mô hình ước tính độ mặn của đất trên cơ sở áp dụng phương pháp hồi quy tuyến tính đa biến, sử dụng thuật toán BMA kết hợp chỉ tiêu thống kê BIC, AIC. Kết quả kiểm chứng cho thấy, giá trị tính toán từ ảnh có tương quan tốt với giá trị đo đạc thực tế (74,64 %). Khu vực nghiên cứu nằm trong vùng khí hậu nhiệt đới gió mùa cận xích đạo, chỉ có hai mùa mưa và khô; thời điểm đất nhiễm mặn cao nhất là cuối mùa khô (khoảng tháng 3, tháng 4 hàng năm) do vậy, dữ liệu ảnh chụp vào ngày 28/3/2021 được sử dụng để thành lập bản đồ có thể phản ánh được thực trạng đất bị nhiễm mặn trong khu vực.

Trên cơ sở phân chia độ mặn theo thang của Tổ chức Nông lương Thế giới (FAO) và kết quả thu được cho thấy khu

vực đất bị mặn cao nhất trong khu vực nghiên cứu đang diễn ra tại khu vực đất thấp, ven biển hay sông. Các phần khác của khu vực chỉ bị ảnh hưởng ở mức nhẹ.

Lời cảm ơn: Bài báo này được hoàn thành dưới sự hỗ trợ của đề tài “Ứng dụng công nghệ viễn thám và GIS thành lập bản đồ đất nhiễm mặn tại một số khu vực thuộc tỉnh Đồng Nai”, mã số UDNGDP.04/20-21.

TÀI LIỆU THAM KHẢO

- [1]. A. Azabdaftari, F.Sunar (2016). *Soil salinity mapping using multitemporal landsat data*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B7, XXIII ISPRS Congress, 12 - 19 July 2016, Prague, Czech Republic.
- [2]. A.E.K. Douaoui, H. Nicolas, C. Walter (2006). *Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data*. Geoderma 134(1): 217 - 230.
- [3]. FAO (2020). *Mapping of salt-affected soils - Technical specifications and country guidelines*. Rome.
- [4]. H.Q. Liu; A.R. Huete (1995). *A feedback based modification of the NDVI to minimize canopy background and atmospheric noise*. IEEE Trans. Geosci. Remote Sensing, 33: 457 - 465.
- [5]. J.D. Rhoades, F. Chanduvi, S. Lesch (1999). *Soil salinity assessment: Methods and interpretation of electrical conductivity measurements*. FAO Irrigation and drainage paper 57.
- [6]. K. George, L. Paraskevi, S. Anastasia (2020). *Comparison of soil EC values from methods based on 1:1 and 1:5 soil to water ratios and ECe from saturated paste extract based method*. Water, 1010; doi:10.3390/w12041010.
- [7]. N. Kim-Anh, L. Yucel-An, T. Ha-Phuong, H. Phi-Phung, N. Thanh-Hung (2020). *Soil salinity assessment by using*

nearinfrared channel and vegetation soil salinity index derived from Landsat 8 OLI data: a case study in the Tra Vinh province, Mekong delta, Vietnam. Progress in Earth and Planetary Science, 7:1, <https://doi.org/10.1186/s40645-019-0311-0>.

[8]. N. M. Khan, V.V. Rastoskuev, E.V. Shalina, Y. Sato (2001). *Mapping salt-affected soils using remote sensing indicators-a simple approach with the use of GIS IDRISI*. 22nd Asian Conference on Remote Sensing, 5(9).

[9]. R-Development Core Team (2011). *A language and environment for statistical computing*.

[10]. S. Khan, A. Abbas (2007). *Remote sensing based modelling applications in*

land and water management: using remote sensing for appraisal of irrigated soil salinity. MODSIM 2007. International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2007.

[11]. Tống Thị Huyền Ái, Nguyễn Vũ Giang, Nguyễn Cao Hạnh, Giang Thị Phương Thảo, Danh Mồn, Phạm Việt Hòa (2018). *Nghiên cứu ước tính độ mặn của đất sử dụng ảnh vệ tinh Sentinel-2 và thiết bị đo dẫn điện mặt đất EM31-EK2*. Tạp chí Khoa học Trường Đại học Sư phạm TP. Hồ Chí Minh: Khoa học tự nhiên và Công nghệ, Tập 15, số 11b (2018):71 - 78.

BBT nhận bài: 16/12/2021; Phản biện xong: 27/12/2021; Chấp nhận đăng: 28/3/2022