

# NGHIÊN CỨU CƠ SỞ KHOA HỌC TRONG LỰA CHỌN CHỈ SỐ CẢNH QUAN PHÙ HỢP PHỤC VỤ CÔNG TÁC GIÁM SÁT BIẾN ĐỔI CẤU TRÚC RỪNG NGẬP MẶN. THỰC NGHIỆM TẠI RỪNG NGẬP MẶN MŨI CÀ MAU

PHẠM MINH HẢI, ĐỖ THỊ HOÀI

Viện Khoa học Đo đạc và Bản đồ

## Tóm tắt:

Chỉ số (spatial metrics) được định nghĩa là các chỉ số cảnh quan để mô tả hình thái và cấu trúc của cảnh quan (O Neill và ctv, 1988). Các nghiên cứu đánh giá sự thay đổi cấu trúc rừng dựa trên ảnh vệ tinh sử dụng các chỉ số cảnh quan đang trở thành một xu hướng phổ biến trong nghiên cứu mức độ ảnh hưởng của phân mảnh rừng tới các hệ sinh thái. Tuy nhiên cho tới hiện nay, sự tìm hiểu hiểu toàn diện về lý thuyết về sử dụng đúng các chỉ số cảnh quan phục vụ cho các mục đích trong các điều kiện khác nhau chưa được phổ biến. Có rất nhiều chỉ số cảnh quan đã được phát triển phục vụ phân tích cảnh quan như: phân tích biến động lớp phủ bề mặt, biến động đô thị và rừng. Do vậy, công tác nghiên cứu nhằm lựa chọn chỉ số cảnh quan phù hợp cho nghiên cứu biến động về cấu trúc rừng, đặc biệt là rừng ngập mặn, khi kết hợp với công nghệ viễn thám là một vấn đề đặt ra cần được quan tâm giải quyết. Mục tiêu của bài báo này nhằm nghiên cứu cơ sở khoa học trong lựa chọn chỉ số cảnh quan phù hợp cho tiêu chí phân mảnh rừng phục vụ công tác giám sát biến đổi cấu trúc rừng ngập mặn. Nghiên cứu được thực nghiệm tại rừng ngập mặn Mũi Cà Mau.

## 1. Giới thiệu chung

Chỉ số (spatial metrics) được định nghĩa là các chỉ số cảnh quan để mô tả hình thái và cấu trúc của cảnh quan (O Neill và ctv, 1988). Các nghiên cứu đánh giá sự thay đổi cấu trúc rừng dựa trên ảnh vệ tinh sử dụng các chỉ số cảnh quan đang trở thành một xu hướng phổ biến trong nghiên cứu mức độ ảnh hưởng của phân mảnh rừng tới các hệ sinh thái. Mỗi chỉ số cảnh quan được sử dụng thích hợp cho một môi trường khác nhau do vậy khi sử dụng và giải đoán chúng bị nhiều hạn chế hạn nhất định trong các bối cảnh khác nhau. Tuy nhiên cho tới hiện nay, sự tìm hiểu hiểu toàn diện về lý thuyết về sử dụng đúng các chỉ số cảnh quan phục vụ cho các mục đích trong các điều kiện khác nhau chưa được phổ biến. Bên cạnh đó, với các kiểu cảnh quan đa dạng và các chỉ số đã được áp dụng trong khung cảnh thế giới thực làm cho sự

nghiên cứu và nâng cao nhận thức về cách sử dụng các chỉ số cảnh quan trong điều kiện thế giới thực là rất khó.

Có rất nhiều chỉ số cảnh quan đã được phát triển phục vụ phân tích cảnh quan như: phân tích biến động lớp phủ bề mặt, biến động đô thị và rừng. Do vậy, công tác nghiên cứu nhằm lựa chọn chỉ số cảnh quan phù hợp cho nghiên cứu biến động về cấu trúc rừng, đặc biệt là rừng ngập mặn, khi kết hợp với công nghệ viễn thám là một vấn đề đặt ra cần được quan tâm giải quyết.

Mục tiêu của bài báo này nhằm nghiên cứu cơ sở khoa học trong lựa chọn chỉ số cảnh quan phù hợp cho tiêu chí phân mảnh rừng phục vụ công tác giám sát biến đổi cấu trúc rừng ngập mặn. Nghiên cứu được thực nghiệm tại rừng ngập mặn Mũi Cà Mau.

## 2. Mục đích và các chỉ số cảnh quan được

Ngày nhận bài: 09/11/2019, ngày chuyển phản biện: 19/11/2019, ngày chấp nhận phản biện: 25/11/2019, ngày chấp nhận đăng: 01/12/2019

## lựa chọn

Nhóm nghiên cứu tiến hành nghiên cứu cơ sở khoa học trong lựa chọn chỉ số cảnh quan phù hợp cho cho tiêu chí phân mảnh rừng phục vụ công tác giám sát biến đổi cấu trúc rừng ngập mặn, với khu vực thực nghiệm tại Mũi Cà Mau. Để phục vụ cho mục đích phân tích cấu trúc phân mảnh rừng ngập mặn, nhóm chỉ số Đồng nhất (Aggregation) được sử dụng để nhằm lựa chọn ra chỉ số phù hợp nhất (MaGarigal, 2002). Giữa các chỉ số cảnh quan có thể có sự giống nhau về khái niệm nhưng chúng không giống nhau về chức năng và cách thức sử dụng. Trong nghiên cứu này, chúng tôi sử dụng 11 chỉ số cảnh quan như AI-Chỉ số đồng nhất (Aggregation index); Cohension-Chỉ số kết liên (Patch Cohesion Index); Division-Chỉ số phân chia (Landscape Division Index); Lsi-Chỉ số hình dạng(Landscape Shape Index); Mesh-Chỉ số kích thước mắt lưới (Effective Mesh Size); Np-Chỉ số số mảng (Number of patch index); Pd-Chỉ số mật độ mảng (Patch density Index); Pladj-Chỉ số điểm ảnh liền kề (Proportion of Like Adjacency Index) ([https://www.umass.edu/landeco/research/fragstats/documents/fragstats.help\\_4.2.pdf](https://www.umass.edu/landeco/research/fragstats/documents/fragstats.help_4.2.pdf)).

### 3. Hệ số tương quan trong thống kê

#### 3.1. Khái lược về phương pháp tính hệ số tương quan

Có nhiều phương pháp tính hệ số tương quan trong thống kê được sử dụng rộng rãi như Pearson, Kendall, hay Spearman, tuy nhiên phương pháp phổ biến nhất là Pearson correlation (r). Phương pháp này đo lường độ mạnh và hướng của mối quan hệ tuyến tính giữa hai biến, không thể áp dụng cho hai biến không có mối quan hệ tuyến tính và cũng không thể phân biệt được biến độc lập và biến phụ thuộc. Giá trị tuyến tính nằm trong khoảng từ -1 đến 1, ở đó giá trị 1 nói lên mối quan dương hoàn hảo giữa hai biến, có nghĩa là nếu biến thứ nhất tăng thì biến thứ hai cũng tăng theo. Còn giá trị -1

cũng phản ánh mức độ tương quan âm được coi là hoàn hảo tuy nhiên các biến thay đổi ngược nhau, nghĩa là biến đầu tăng thì biến thứ hai sẽ giảm và ngược lại. Còn giá trị 0 được hiểu là không có mối quan hệ nào giữa hai biến đó. Độ mạnh của mối quan hệ giữa các biến thể hiện ở giá trị của hệ số, ví dụ giá trị 0.3 có nghĩa là có sự tương quan dương nhưng sự tương quan đó yếu và hầu như là không có ý nghĩa. Thông thường sẽ không xem xét ý nghĩa của sự tương quan nếu giá trị tương quan nhỏ hơn 0.8. Còn với hệ số tương quan từ 0.9 trở lên được đại diện sự tương quan mạnh.

Hệ số tương quan theo phương pháp Pearson được biểu thị theo công thức dưới đây:

$$\rho_{xy} = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

Trong đó:

$\rho_{xy}$ : hệ số tương quan Pearson

$\text{Cov}(x, y)$ : hiệp phương sai của biến x và y

$\sigma_x$ : độ lệch chuẩn của biến x

$\sigma_y$ : độ lệch chuẩn của biến y

Để tính hệ số tương quan giữa kết quả ảnh phân loại và ảnh chỉ số, ở đây sẽ sử dụng phương pháp Pearson. Do phải tính hệ số tương quan cho từng cặp (ảnh kết quả phân loại và ảnh chỉ số hình thái) với số lượng lớn cặp chỉ số và cần thiết có một kết quả minh họa tổng thể mối liên hệ giữa các chỉ số một cách trực quan, nên tác giả đã xây dựng chương trình tự động tính hệ số tương quan và tạo biểu đồ nhiệt thể hiện mức độ và chiều của sự tương quan giữa các cặp ảnh. Nhóm thực hiện đề tài tiến hành tính toán trên dữ liệu thử nghiệm sử dụng ngôn ngữ Python, có các bộ thư viện hỗ trợ như: Rasterio, Seaborn, Numpy, Pandas, và bộ thư viện học máy Sklearn (hình 1).

#### 3.2. Thực nghiệm tính hệ số tương quan

##### 3.2.1. Khu vực thử nghiệm và dữ liệu đầu vào

Cà Mau là một tỉnh ven biển ở cực nam của Việt Nam, nằm trong khu vực Đồng bằng sông Cửu Long. Phần lãnh thổ đất liền của tỉnh Cà Mau nằm trong tọa độ từ 8°30' - 9°10' vĩ Bắc và 104°80' - 105°5' kinh Đông. Điểm cực Đông tại 105°24' kinh Đông thuộc xã Tân Thuận, huyện Đầm Dơi. Điểm cực Nam tại 8°33' vĩ Bắc thuộc xã Viên An, huyện Ngọc Hiển. Điểm cực Tây tại 104°43' kinh Đông thuộc xã Đất Mũi, huyện Ngọc Hiển. Điểm cực Bắc tại 9°33' vĩ Bắc thuộc xã Biển Bạch, huyện Thới Bình.

```
import numpy as np
import rasterio
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
import os

d_agg={}
d_agg['cls_a'] = rasterio.open('Path_to_Classified_Image').read().flatten()
path = 'Path_to_Index_Image'
band_id = [band for band in os.listdir(path) if band[-4:] == '.tif']
for i in band_id:
    print(i.split('.')[0])
    d_agg[i.split('.')[0]] = rasterio.open(path+i).read().flatten()
df_a = pd.DataFrame(d_agg)

#Set StandardScale for data
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
scaled_data = scaler.fit_transform(df_a)
import pandas as pd
dd={}
for i,j in enumerate(df_a.keys(),start=0):
    dd[j] = scaled_data[:,i]
df_scl = pd.DataFrame(dd)

# Correlation calculation
corr_a = df_scl.corr()
# Checking the value of heatmap correlation
img = corr_a.style.background_gradient(cmap='coolwarm', axis=None)
# Masking to make the diagonal heatmap of correlation
mask = np.zeros_like(corr_a, dtype=np.bool)
mask[np.triu_indices_from(mask)] = True

# Draw correlation heatmap plot
sns.set_style(style = 'whitegrid')
sns.heatmap(corr_a, mask=mask, cmap='coolwarm',
            square=True, annot=True,
            linewidths=.5, cbar_kws={"shrink": .5}, ax=ax)
plt.tight_layout()

# Save to file
plt.savefig('Path_to_Save_Result')
```

Hình 1: Minh họa ngôn ngữ Python và các bộ thư viện hỗ trợ sử dụng trong nghiên cứu

Bài báo thực hiện với dữ liệu ảnh vệ tinh SPOT6 được chụp tại thời điểm năm 2015. Dữ liệu được cung cấp bởi Đài thu Viễn thám Quốc Gia, Cục Viễn thám Quốc Gia, Bộ Tài Nguyên và Môi Trường thuộc đề tài nghiên cứu khoa học Chương trình KHCN vũ trụ giai đoạn 2016-2020: “Nghiên cứu ứng dụng công nghệ viễn thám và phương pháp trắc lượng hình thái trong xây dựng bộ tiêu chí giám sát quá trình suy thoái rừng ngập mặn phục vụ công tác bảo tồn, phục hồi rừng ngập mặn và sử dụng hợp lý đất ngập nước ven biển. Thử nghiệm tại Tỉnh Cà Mau và Quảng Ninh”. Ảnh vệ tinh trong đề tài có độ phủ mây nhỏ hơn 10%, và được hiệu chỉnh bức xạ và khí quyển bằng công cụ ATCOR (Atmospheric

correction) trong phần mềm PCI Geomatic 2018. Quá trình tiền xử lý ảnh gồm 3 bước: Hiệu chỉnh bức xạ trước đầu thu; hiệu chỉnh bức xạ ảnh hưởng mây; hiệu chỉnh ảnh hưởng khí quyển. Nhóm thực hiện nghiên cứu tiến hành công tác nắn chỉnh hình học và trộn ảnh về độ phân giải 2.5m. Ảnh được nắn về hệ tọa độ WGS84 và hệ quy chiếu UTM múi 48. Để đánh giá độ chính xác của kết quả thực hiện, nhóm thực hiện đề tài sử dụng bản đồ kiểm kê rừng tại cùng khu vực nghiên cứu được thành lập năm 2015.



Hình 2: Phạm vi khu vực thực nghiệm (<http://www.camau.gov.vn/wps/portal/bando>)

### 3.2.2. Công tác phân loại ảnh vệ tinh

Nhóm thực hiện đề tiến hành để thu thập mẫu giải đoán ảnh tại khu vực nghiên cứu. Công tác thu thập mẫu được tiến hành cả bằng phương pháp thực địa và trên phần mềm Google Earth. Sau đó, Phương pháp Multiresolution segmentation trên phần mềm xử lý ảnh Ecognition đã được sử dụng để phân đoạn. Sử dụng thuật toán Random Forest và dữ liệu ảnh SPOT6, nhóm nghiên cứu đã tiến hành phân loại ảnh viễn thám và đánh giá độ chính xác của kết quả phân loại ảnh phục vụ xây dựng được hiện trạng rừng ngập mặn khu vực thực nghiệm với các đối tượng cây ngập mặn được chia ra thành các loại như: mắm, đước, hỗn giao.v.v. (Phạm Minh Hải và ctv, 2019) (hình 3). (Xem hình 3)

Sau khi phân loại ảnh, kết quả phân loại ảnh

được sử dụng là dữ liệu đầu vào của quá trình xử lý với các chỉ số cảnh quan đã lựa chọn. Đối với các chỉ số trong nhóm Aggregate, công tác xử lý ảnh thực hiện khu vực có rừng và không có rừng. Do đó, các sản phẩm phân loại ảnh sẽ được xử lý còn hai lớp rừng – không rừng, được lý hiệu là lớp Cls khi phân tích hồi quy tuyến tính khi xử lý với các chỉ số khác (hình 4). (Xem hình 4)

Kết quả tính toán chạy các chỉ số cảnh quan với dữ liệu đầu vào là ảnh phân loại được thể hiện dưới hình 5. Ở đó, hiện tượng phân mảnh rừng được thể hiện qua các chỉ số khác nhau. (Xem hình 5)

Nhóm nghiên cứu tiến hành cả phương pháp so sánh định tính và định lượng nhằm lựa chọn được chỉ số phù hợp nhất trong nhóm Aggregate. Tuy các sản phẩm ảnh chỉ số được tính toán trong cùng nhóm nhưng của kết quả tính toán thể hiện cấu trúc phân mảnh khác nhau với giá trị khác nhau.

Về so sánh định tính, từ hình 5 cho thấy sản phẩm tính toán chỉ số PLADJ thể hiện rõ nét các khu vực bị phân mảnh so với các chỉ số khác.

Về định lượng, nhóm nghiên cứu tiến hành chạy tìm hệ số tương quan giữa cặp dữ liệu với nhau để chọn ra chỉ số phù hợp nhất trong nhóm Aggregate mô tả cấu trúc về độ độ phân mảnh rừng. Với Cls là lớp phân bố của rừng ngập mặn thì chỉ số PLADJ trong nhóm Aggregate tương quan cao trên biểu đồ Heatmap với hệ số tương quan  $R=0.92$ .

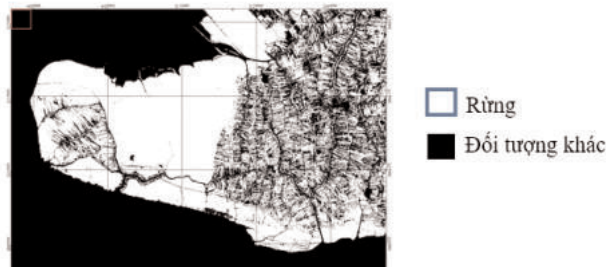
Kết hợp so sánh định tính và định lượng, nhóm nghiên cứu nhận thấy chỉ số PLADJ thể hiện sự tương quan cao nhất đối với CLS, qua đó được đánh giá là phù hợp nhất của chỉ số này khi so sánh với các chỉ số còn lại. (Xem hình 6)

#### 4. Kết luận

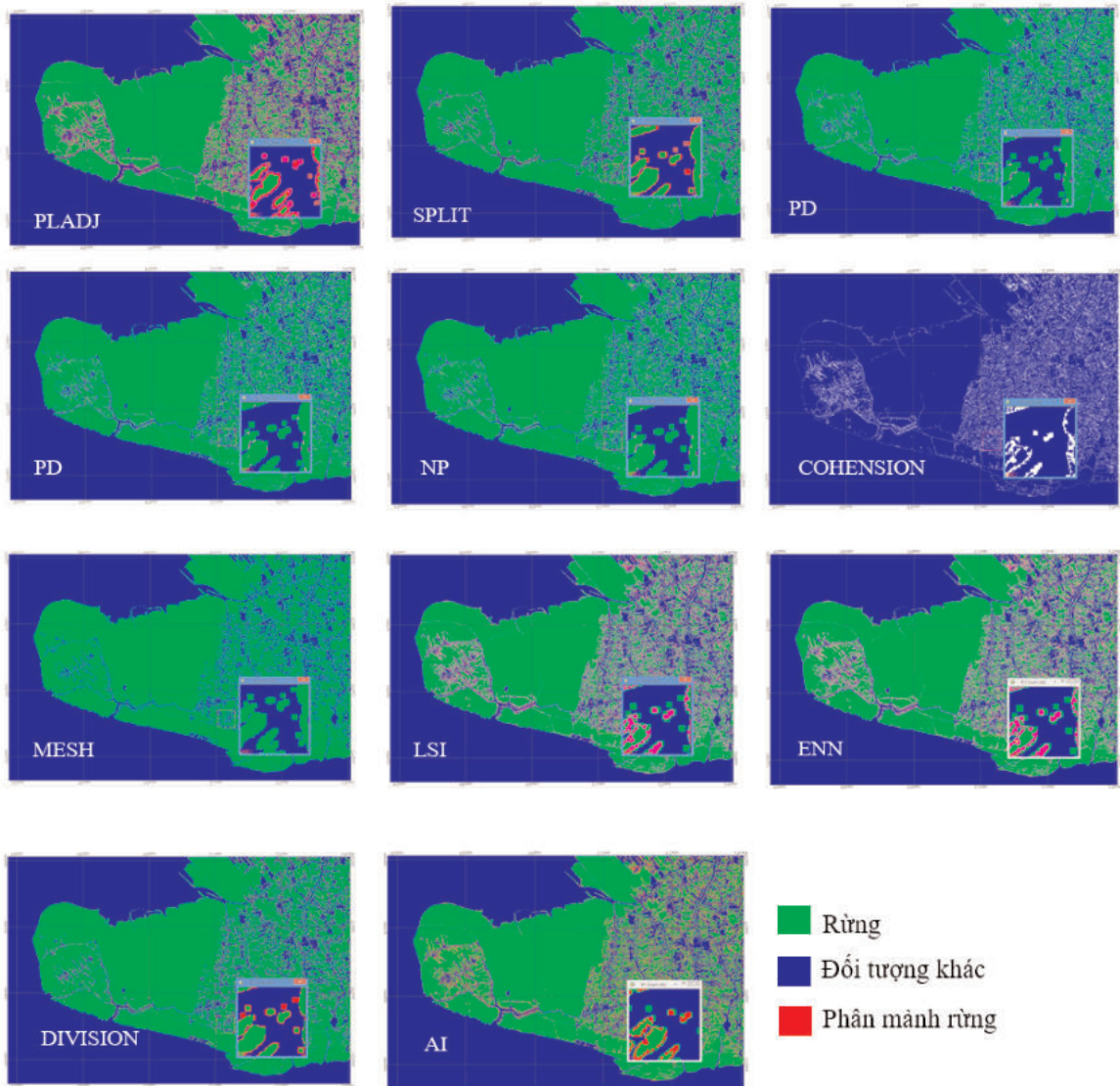
Nhóm thực hiện nghiên cứu đã thực hiện phân tích và dựng tổng hợp các chỉ số cảnh quan



Hình 3: Kết quả phân loại ảnh sử dụng thuật toán Random Forest



Hình 4: Kết quả phân loại sau khi xử lý



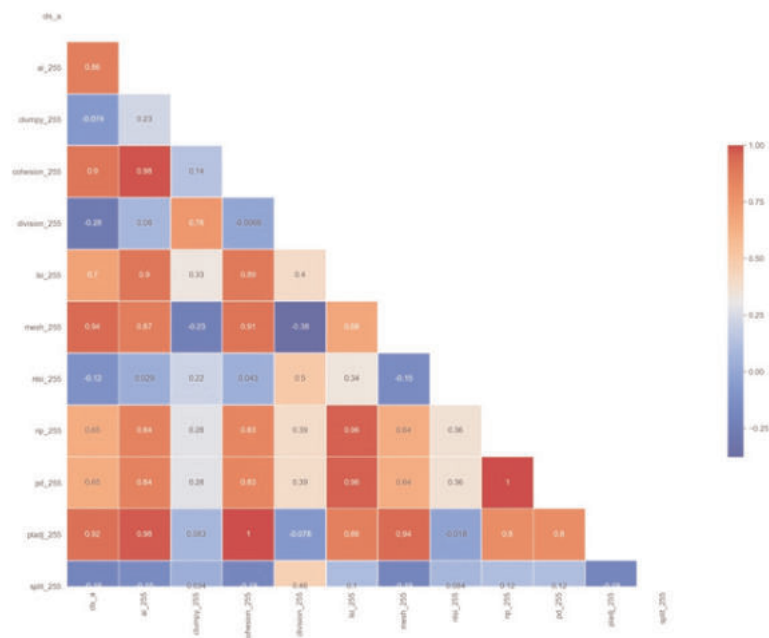
Hình 5: Kết quả tính các chỉ số cảnh quan

trong nhóm Aggregate thực hiện mô tả cấu trúc phân mảnh của rừng ngập mặn. Dựa trên những kết quả thu được, với sự thể hiện rõ ràng sự phân mảnh của rừng ngập mặn trên kết quả tính toán chỉ số cảnh quan và hệ số tương quan cao với nhóm CLS (phân bố rừng ngập mặn) là 0.92, chỉ số cảnh quan PLADJ trong nhóm Aggregate đáp ứng hoàn toàn được cả về mặt định tính (thể hiện sự phân bố rõ ràng) và mặt định lượng (hệ số tương quan cao) khi mô tả tiêu chí phân mảnh của rừng ngập mặn với khu vực thử nghiệm tại

Mũi Cà Mau. Đây là một tiền đề có ý nghĩa vô cùng quan trọng trong nghiên cứu suy thoái rừng ngập mặn trong tương lai ứng dụng kết hợp ảnh vệ tinh và chỉ số cảnh quan. ○

#### Tài liệu tham khảo

- [1]. McGarigal, 2002.  
 (<https://www.umass.edu/landeco/research/fragstats/documents/fragstats.help.4.2.pdf>).
- [2]. O'Neill (1988). Indices of landscape pat-



Hình 6: Các hệ số tương quan giữa các chỉ số cảnh quan - chỉ số cảnh quan và chỉ số cảnh quan - lớp phân bố rừng ngập mặn trên đồ thị Heatmap

tern. Landscape Ecology, 1, 153-162.

Khoa học Đo đạc và Bản đồ. Số 40-6/2019.

[3]. Phạm Minh Hải và ctv, 2019. Nghiên cứu cơ sở khoa học ứng dụng thuật toán Random Forest trong phân loại ảnh vệ tinh SPOT6 với khu vực thử nghiệm tại Tỉnh Cà Mau. Tạp chí

[4]. Pearson's correlation (2012).

<http://www.statstutor.ac.uk/resources/uploaded/pearsons.pdf>

### Summary

**The scientific research in selecting suitable spatial metrics for monitoring changes of mangrove structure. The study area in the mangrove forests of Mui Ca Mau, Ca Mau province of Vietnam**

*Pham Minh Hai, Do Thi Hoai, Vietnam Institute of Geodesy and Cartography*

Spatial metrics are defined as landscape indicators to describe the morphology and structure of a landscape (O'Neill et al., 1988). Recent studies investigating changes in forest structure using satellite images and Spatial metrics are becoming a new trend in studying the extent of forest fragmentation on ecosystems. However, a comprehensive understanding of the theory of the selections of suitable indicators for purposes under different conditions has not been disseminated. There are many spatial metrics developed for landscape analysis, such as: analysis of land use and land cover change, and forest changes. Therefore, studying suitable spatial metrics for analyzing changes in forest structure, especially mangroves, using remote sensing and spatial metrics is getting very important. The objective of this manuscript is to investigate the scientific basis for selecting suitable spatial metrics the forest fragmentation in order to manage changes in mangrove structure. The study area in the mangrove forests of Mui Ca Mau, Ca Mau province of Vietnam. ○