

Article

Vegetation Identification in Hyperspectral Images Using Distance/Correlation Metrics

Gabriel E. Chanchí Golondrino *, Manuel A. Ospina Alarcón  and Manuel Saba 

Faculty of Engineering, University of Cartagena, Cartagena de Indias 130015, Colombia; mospinaa@unicartagena.edu.co (M.A.O.A.); msaba@unicartagena.edu.co (M.S.)

* Correspondence: gchanchig@unicartagena.edu.co

Abstract: Distance/correlation metrics have emerged as a robust and simplified tool for assessing the spectral characteristics of hyperspectral image pixels and effectively categorizing vegetation within a specific study area. Correlation methods provide a readily deployable and computationally efficient approach, rendering them particularly advantageous for applications in developing nations or regions with limited resources. This article presents a comparative investigation of correlation/distance metrics for the identification of vegetation pixels in hyperspectral imagery. The study facilitates a comprehensive evaluation of five distance and/or correlation metrics, namely, direct correlation, cosine similarity, normalized Euclidean distance, Bray–Curtis distance, and Pearson correlation. Direct correlation and Pearson correlation emerged as the two metrics that demonstrated the highest accuracy in vegetation pixel identification. Using the selected methodologies, a vegetation detection algorithm was implemented and validated using a hyperspectral image of the Manga neighborhood in Cartagena de Indias, Colombia. The spectral library facilitated image processing, while the mathematical calculation of correlations was performed using the numpy and scipy libraries in the Python programming language. Both the approach adopted in this study and the implemented algorithm aim to serve as a point of reference for conducting detection studies on diverse material types in hyperspectral imagery using open-access programming platforms.



Citation: Chanchí Golondrino, G.E.; Ospina Alarcón, M.A.; Saba, M. Vegetation Identification in Hyperspectral Images Using Distance/Correlation Metrics. *Atmosphere* **2023**, *14*, 1148. <https://doi.org/10.3390/atmos14071148>

Academic Editor: Riccardo Buccolieri

Received: 28 May 2023
Revised: 30 June 2023
Accepted: 12 July 2023
Published: 14 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: distance-based method; urban vegetation; vegetation clustering; vegetation classification

1. Introduction

Remote sensing technology has revolutionized the field of vegetation mapping and monitoring, providing a powerful tool for scientists and researchers to study and understand complex terrestrial ecosystems [1–4]. The ability to obtain high-resolution images of the Earth’s surface from satellites, aircraft, and drones has transformed the way humans perceive and analyze vegetation patterns, providing insights into the distribution, structure, and health of plant communities across different regions and biomes [1,5–10].

Despite the significant advances in remote sensing technology over the past few decades, the identification of vegetation in multi- and hyperspectral images remains a challenging task due to the complex spectral signatures of different plant species, as well as the influence of environmental factors, such as soil, water, and atmospheric conditions [11]. Traditional methods for vegetation mapping using remote sensing involve the interpretation of spectral indices, such as the Normalized Difference Vegetation Index (NDVI), which provide a measure of the vegetation cover based on the contrast between the red and near-infrared bands of the electromagnetic spectrum [9,12–21].

While these methods have proven to be effective in many applications, they suffer from several limitations, including the need for manual interpretation and the limited spatial resolution of satellite images. Moreover, traditional methods are often time-consuming and costly, requiring extensive field surveys and ground truth data to calibrate and validate the results [11,22,23].

To overcome these challenges, machine learning techniques have emerged as a promising approach for vegetation mapping and classification, offering a cost-effective and accurate solution for remote sensing applications, especially in developing countries. Machine learning algorithms can analyze vast amounts of hyperspectral data and identify patterns and features that are difficult or impossible to discern through human interpretation, allowing for the automated and objective detection of vegetation cover. The use of machine learning applications in vegetation mapping has been on the rise in recent years [24–26], driven by the availability of large hyperspectral datasets and the development of advanced algorithms and computational tools. Machine learning techniques, such as artificial neural networks [27–29], decision trees [30,31], support vector machines [32], and random forests [33,34], have shown promising results in identifying and classifying vegetation cover from hyperspectral images, offering a more robust, reliable, and cost-effective solution for remote sensing applications.

In addition to machine learning techniques, correlation methods between curves have emerged as another powerful tool for comparing spectral signatures of hyperspectral image pixels and efficiently classifying vegetation in a study area. Spectral signatures are graphical representations of the reflectance values of an object or material across the electromagnetic spectrum. In the context of vegetation mapping, spectral signatures can be used to identify different plant species or vegetation types based on their unique spectral properties [35]. Correlation methods between curves can be used to compare the spectral signatures of different pixels in a hyperspectral image and identify similarities or differences in their spectral properties. Among the most common correlation methods used in vegetation mapping are Pearson's correlation coefficient and spectral angle mapper, among many other methods available [30,36].

Pearson's correlation coefficient measures the linear relationship between the two sets of data and ranges from -1 to 1 , with values closer to 1 indicating a strong positive correlation between the two sets of data. In vegetation mapping, Pearson's correlation coefficient can be used to compare the spectral signatures of different pixels and identify areas with similar vegetation cover [37,38].

On the other hand, the spectral angle mapper measures the angular similarity between two spectral signatures and ranges from 0 to 1 , with values closer to 1 indicating a high degree of spectral similarity between the two signatures. The spectral angle mapper can be used to compare the spectral signatures of different pixels and identify areas with similar vegetation [30,39,40].

These correlations offer a complementary approach to machine learning techniques. Moreover, these correlation methods can be easily implemented and computationally efficient, making them particularly useful for applications in developing countries or regions with limited resources.

The need to efficiently classify vegetation in a study area using remote sensing technology has led to the development of various methods for comparing spectral signatures of hyperspectral image pixels. The focus of the present work is to determine the best distance/correlation-based method for vegetation detection in hyperspectral images with 380 bands from 400 nm to 2400 nm in an urban area through five different correlation methods, "Correlation methods", "Cosine" "Euclidean", "Bray Curtis", and "Pearson" [41–45].

2. Methodology

For the development of this research, the following 5 methodological phases were considered (Figure 1): identification of reference pixels, obtaining the average or characteristic pixel, determination of distance/correlation metrics, selection of the best metrics, and detection of vegetation in the reference image.

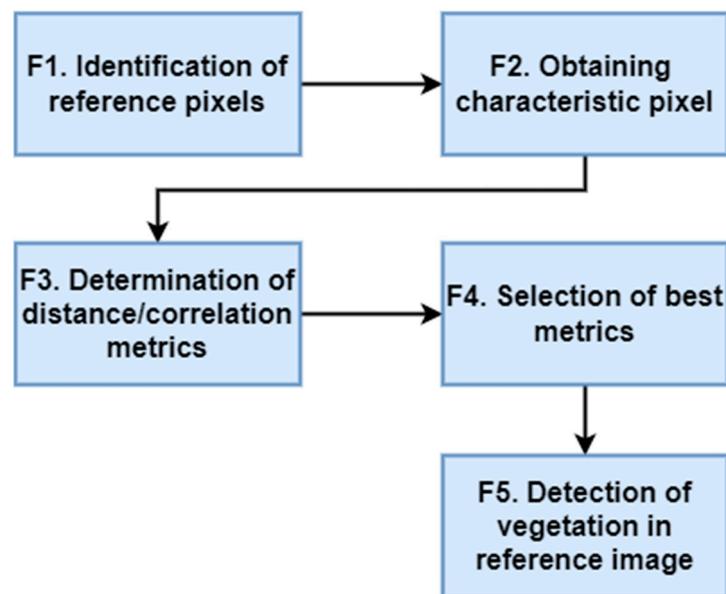


Figure 1. The figure depicts the methodology utilized in the study.

According to Figure 1, in phase 1 of the methodology, a set of 100 coordinates corresponding to vegetation pixels and a set of 100 coordinates corresponding to non-vegetation pixels (roofs, roads, sea, and containers) were obtained manually from the reference image, in order to compare the accuracy of the different distance and correlation metrics in both sets. In phase 2 of the methodology, descriptive statistics were used to determine the characteristic pixel or average pixel. The mean was employed as a measure of central tendency by averaging the 380 positions (frequency bands spanning from 400 nm to 2400 nm) of the 100 vegetation pixels. The aim was for the characteristic pixel to collect, across its 380 positions, the average information from the different frequency bands. Once the characteristic pixel is obtained, in phase 3 of the methodology, the calculation of the distance/correlation metrics between the average pixel and the vegetation and non-vegetation sets is performed, in order to determine the precision of each metric. The 5 distance/correlation metrics used in this phase were direct correlation (correlation distance), similarity or cosine distance, normalized Euclidean distance, Bray–Curtis distance (also known as Normalized Manhattan distance), and Pearson’s correlation coefficient. The mathematical description of each of these metrics is presented below.

Direct correlation: It is a measure of dependence between two paired random variables (vectors) of arbitrary dimension, not necessarily equal. That is, it calculates the correlation distance between two one-dimensional arrays. The correlation distance between two vectors \vec{u} and \vec{v} is defined as [41], Equation (1):

$$d_{(\vec{u}, \vec{v})} = 1 - \frac{(\vec{u} - \bar{\vec{u}}) \cdot (\vec{v} - \bar{\vec{v}})}{\|\vec{u} - \bar{\vec{u}}\|^2 \|\vec{v} - \bar{\vec{v}}\|^2} \quad (1)$$

where $\bar{\vec{u}}$ and $\bar{\vec{v}}$ are the means of the elements of \vec{u} and \vec{v} , $(\vec{u} - \bar{\vec{u}}) \cdot (\vec{v} - \bar{\vec{v}})$ is the scalar product, and $\|\vec{u} - \bar{\vec{u}}\|^2$ is its norm or variance.

Similarity or cosine distance: It is a measure of similarity between two non-zero vectors defined in an inner product space. The cosine distance is the cosine of the angle between the vectors. It is the scalar product of the vectors divided by the product of their lengths. It differs from direct correlation as it circumvents the utilization of the arithmetic mean or

average of the vectors under examination. The cosine distance, ascribed as the disparity between two vectors \vec{u} and \vec{v} , is rigorously delineated as [45], Equation (2):

$$d_{(\vec{u}, \vec{v})} = 1 - \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|^2 \|\vec{v}\|^2} \tag{2}$$

where $\vec{u} \cdot \vec{v}$ is the scalar product and $\|\vec{u}\|^2$ is its norm.

Normalized Euclidean Distance: It gives the squared distance between two vectors, where the lengths have been scaled to have a unit norm. This is useful when the direction of the vector is significant but the magnitude is not [43]. According to [43], the normalized Euclidean distance of two vectors is defined by Equation (3):

$$NED^2_{(\vec{u}, \vec{v})} = \frac{1}{2} \cdot \frac{\|(\vec{u} - \bar{u}) - (\vec{v} - \bar{v})\|^2}{(\|\vec{u} - \bar{u}\|^2 + \|\vec{v} - \bar{v}\|^2)} = 0.5 \frac{var(\vec{u} - \vec{v})}{var(\vec{u}) + var(\vec{v})} \tag{3}$$

where \bar{u} and \bar{v} are the means of the elements of \vec{u} and \vec{v} and $\|\cdot\|^2$ is the variance.

Bray–Curtis distance: The Bray–Curtis distance is a normalization method commonly used in the field of botany, ecology, and environmental sciences [44]. It is a statistic used to quantify the compositional dissimilarity between two different sites, based on counts at each site. It is defined by Equation (4) [44]:

$$d_{(\vec{u}, \vec{v})} = \frac{\sum_{i=1}^n (u_i - v_i)}{\sum_{i=1}^n (u_i + v_i)} \tag{4}$$

where n is the size of the vectors or sample. The Bray–Curtis distance is in the range [0, 1] if all coordinates are positive and is undefined if the entries are of length zero.

Pearson correlation coefficient: It is a measure of linear dependence between two quantitative random variables (vectors). Unlike covariance, Pearson’s correlation is independent of the scale of measurement of the variables [41]. Pearson’s population correlation coefficient (also denoted by $r_{(\vec{u}, \vec{v})}$) is defined as [43], Equation (5):

$$r_{(\vec{u}, \vec{v})} = \frac{\sigma_{(\vec{u}, \vec{v})}}{\sigma_{\vec{u}} \cdot \sigma_{\vec{v}}} = \frac{\sum_{i=1}^n (u_i - \bar{u}) \cdot (v_i - \bar{v})}{\sqrt{\sum_{i=1}^n (u_i - \bar{u})^2} \cdot \sqrt{\sum_{i=1}^n (v_i - \bar{v})^2}} \tag{5}$$

where n is the sample size, $\sigma_{(\vec{u}, \vec{v})}$ is the covariance of (\vec{u}, \vec{v}) , $\sigma_{\vec{u}}$ is the standard deviation of vector \vec{u} , $\sigma_{\vec{v}}$ is the standard deviation of vector \vec{v} , and \bar{u} and \bar{v} are the means of \vec{u} and \vec{v} , respectively.

Based on the previous equations corresponding to correlation metrics and considering that these metrics have values close to 0 when the correlation is high, Equation (6) was proposed to obtain a percentage of correlation or similarity.

$$porc_corr = (1 - abs(corr)) \cdot 100 \tag{6}$$

In phase 4 of the methodology, the results obtained by calculating the correlations and similarities described in phase 3 were consolidated in relation to vegetation and non-vegetation pixels. The aim was to determine the precision thresholds for the correlation percentages of each metric in both cases, in order to identify the best metrics to be used in vegetation detection. These metrics are the ones that best classify a pixel as vegetation. Finally, in phase 5, the vegetation detection process was performed on the entire hyperspectral reference image using the selected similarity and/or correlation metrics.

The study area of the present work is shown in Figure 2. It is the Manga neighborhood in the city of Cartagena de Indias, Colombia. Regarding the hyperspectral image, it measures 1500×1500 pixels with 380 bands per pixel ranging from 400 nm to 2400 nm. A total of 100 arbitrary vegetation pixels were extracted from various locations across the image. The purpose is to obtain a representative or characteristic pixel by averaging the corresponding spectral signatures of the 100 pixels. This representative pixel was used to compare the correlation and similarity metrics described above. Thus, in Figure 2, the vegetation pixels are shown in blue, while the non-vegetation pixels are shown in red.

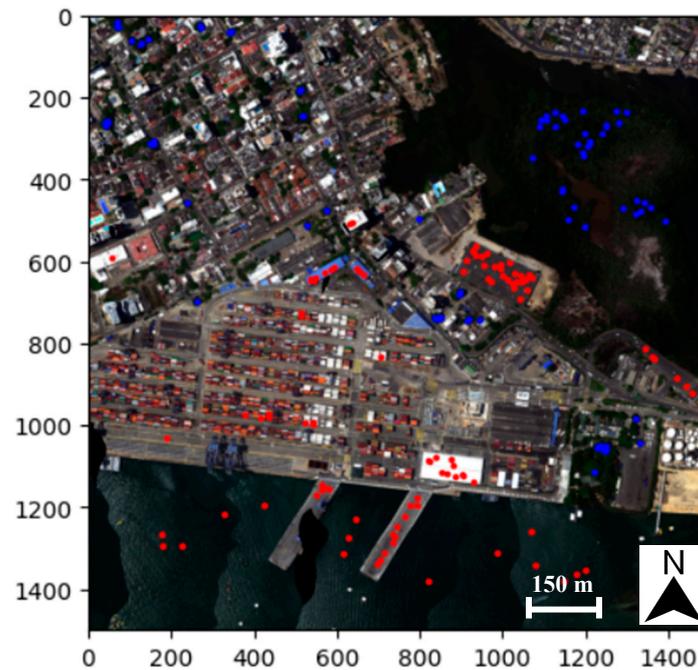


Figure 2. Hyperspectral image used in the present study, in blue vegetation pixels and in red non-vegetation pixels.

3. Results and Discussion

In Figure 3a, the spectral signatures of the selected 100 pixels and their distribution across 380 bands can be observed. It is possible to appreciate that although the reflectance values exhibit variations, the shape of the 100 curves remains the same. They possess a set of peaks that are repeated in a similar manner across different pixels. Similarly, in Figure 3b, the spectral signature corresponding to the zone of no vegetation is presented. Finally, in Figure 3c, the spectral signature corresponding to the characteristic or average pixel can be observed. This pixel is obtained by averaging the curves of the 100 pixels, incorporating the representative maximum and minimum values of the considered sample. The image reading and processing of the different pixels that compose it were performed using the functionalities provided by the Python spectral library. By employing an average pixel normalization approach, it was ascertained that the incorporation of over 100 pixels did not result in significant modifications to the mean pixel value.

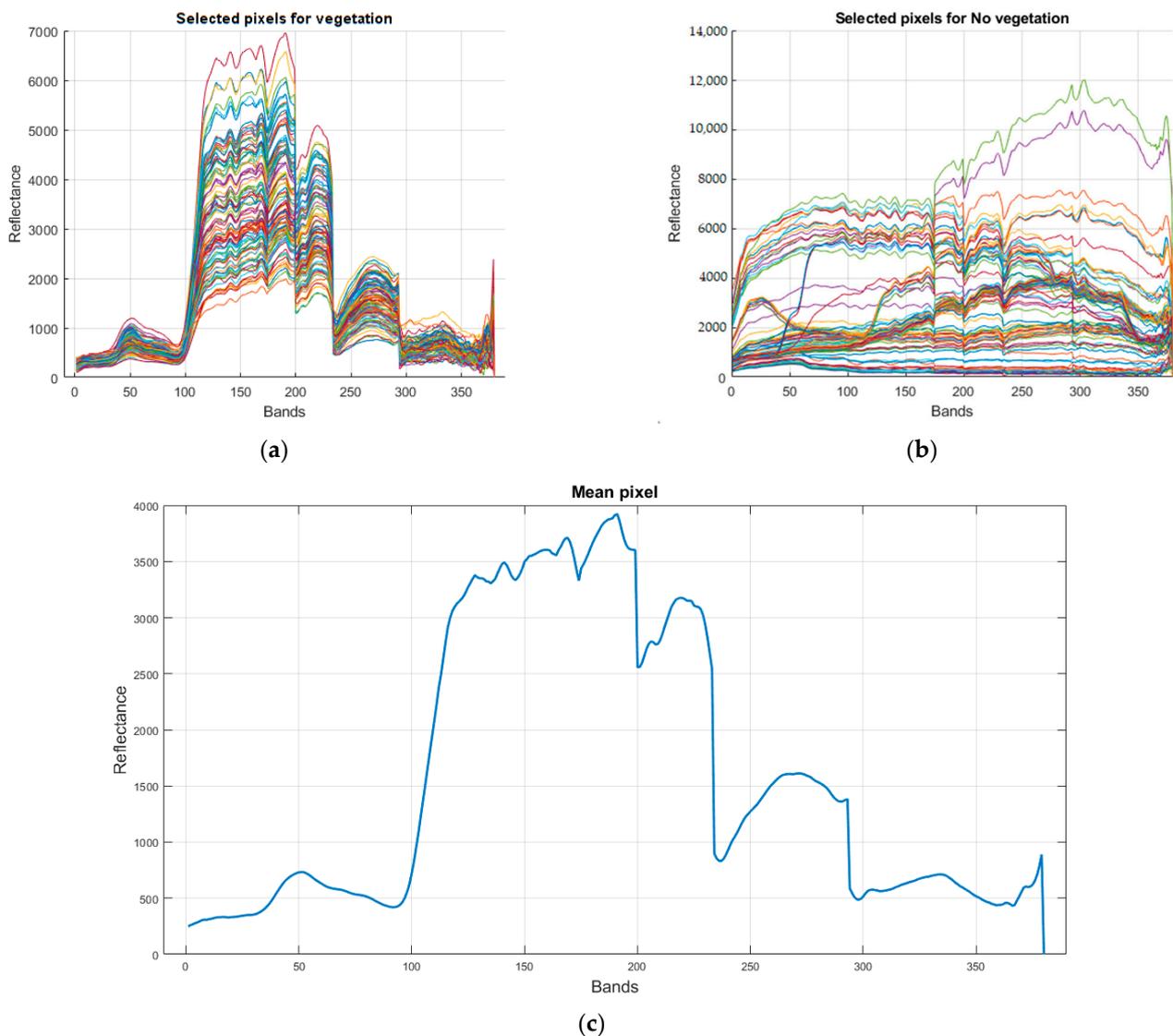


Figure 3. Pixels selected for comparative study: (a) vegetation, (b) non-vegetation, and (c) mean pixel for vegetation.

Based on the above, initially, the representative or average pixel was correlated with each of the 100 selected vegetation pixels using the five metrics considered in Section 2. This was completed to determine both the average and the thresholds obtained for each metric, which helps to identify the most accurate and effective method for identifying the vegetation's spectral signature. For this purpose, the authors started with the idea that each pixel in the image is a vector representation in space. In other words, each pixel is a vector with 380 positions corresponding to the 380 bands, each of which stores a reflectance value. By performing vector operations using the numpy library and the statistical tool from scipy between the average pixel and the 100 vegetation pixels, the bar chart presented in Figure 4 displays the minimum, mean, and maximum values obtained for each of the five metrics considered, as well as the respective standard deviation value for the mean.

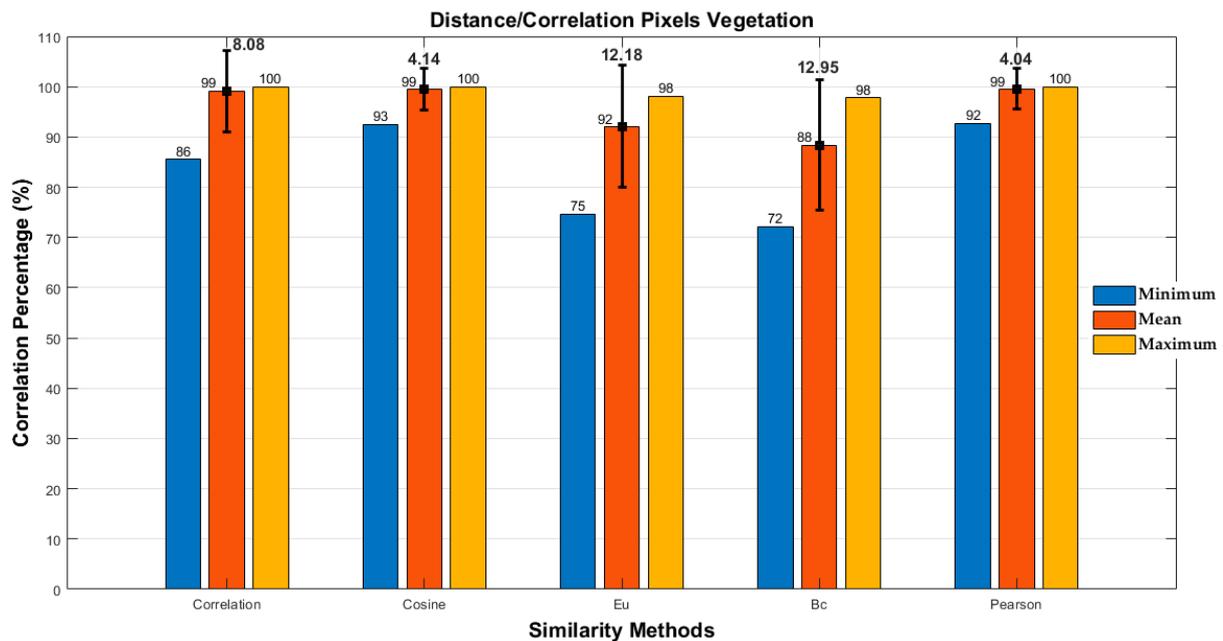


Figure 4. Result of the comparative study of similarity or correlation methods in vegetation pixels.

In Figure 4, it is possible to observe that Pearson correlation, cosine distance, and direct correlation are the metrics that respectively present the highest average values of correlation percentage between the average pixel and the 100 vegetation pixels, with respective values of 99.54%, 99.47%, and 99.09%. Similarly, it can be appreciated that these same three metrics, in the same order, present the highest minimum percentage values of correlation between the average pixel and the 100 vegetation pixels, with respective values of 92.77%, 92.55%, and 85.53%. On the other hand, upon consolidating and comparing the threshold values of these three top metrics, it is possible to observe that the metric with the smallest difference between the maximum and minimum thresholds is Pearson correlation with a difference of 7.2%, followed by cosine distance with a difference of 7.41%.

Once the 100 vegetation pixels were studied, the correlation between the average pixel and the 100 randomly selected non-vegetation pixels was performed using the same five metrics and following a similar procedure as with the vegetation pixels. Hence, by conducting vector operations using the same Python libraries between the average pixel and the 100 non-vegetation pixels, the bar chart in Figure 5 presents the minimum, mean, and maximum values for the five studied metrics, as well as the respective standard deviation value for the mean.

According to Figure 5, it is possible to observe that direct correlation, Bray–Curtis distance (Bc), and Pearson correlation are the metrics that respectively present the lowest average values of correlation percentage between the average pixel and the 100 non-vegetation pixels, with respective values of 31.24%, 56.81%, and 65.62%. Similarly, in Figure 5, it can be appreciated that Bray–Curtis distance (Bc), direct correlation, and Euclidean distance (Eu) are the metrics that respectively present the lowest maximum values of correlation percentage between the average pixel and the 100 non-vegetation pixels, with respective values of 74.24%, 78.52%, and 82.09%. It is also important to mention that cosine distance and Pearson correlation, which showed the two best results for vegetation pixels in this case, have the two highest maximum percentages for non-vegetation pixels with values of 90.12% and 89.26%, respectively. Although these values do not surpass the minimum threshold detected with vegetation pixels, they are close to a difference of 2.43% and 3.51% from the minimum threshold in vegetation pixels.

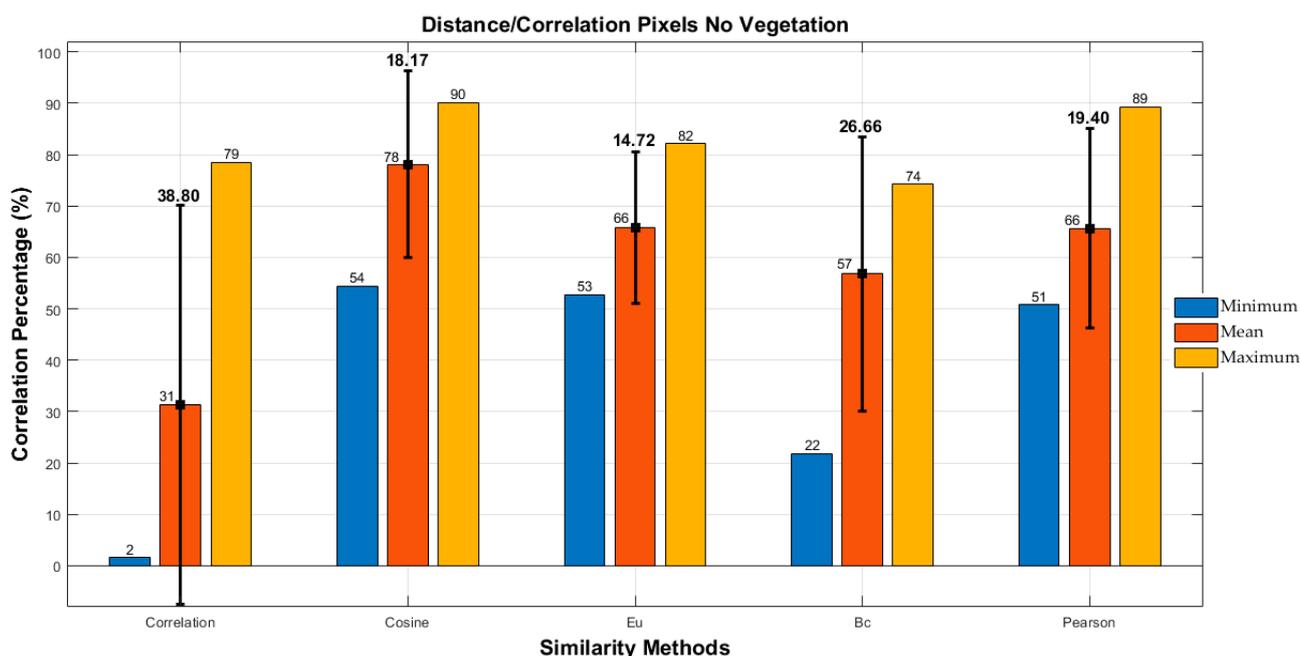


Figure 5. Result of the comparative study of similarity or correlation methods in non-vegetation pixels.

Continuing with the comparison between the results of correlation in vegetation and non-vegetation pixels, Figure 6 presents a comparative graph that shows, for each of the five considered metrics, the minimum correlation percentage value for vegetation and the maximum correlation percentage value for non-vegetation. It can be observed initially that Euclidean distance (Eu) and Bray–Curtis distance (Bc) metrics exhibit an overlap between these thresholds (the maximum value is greater than the minimum value, indicating a negative difference). This implies that applying these methods to the image may lead to confusion between vegetation and non-vegetation pixels. On the other hand, it is evident that the three metrics without overlap between the thresholds are direct correlation, cosine distance, and Pearson correlation. Among them, direct correlation shows the greatest difference between the thresholds (limits), with a percentage difference value of 7.01%.

In Figure 6, it is possible to verify how the overlap of thresholds in the Euclidean distance and Bray–Curtis distance metrics is evidenced by the negative difference between the minimum correlation percentage in vegetation pixels and the maximum correlation percentage in non-vegetation pixels. Similarly, it can be observed that the best results are respectively obtained for the direct correlation and Pearson correlation metrics, with percentage difference values of 7.01% and 3.51%. It is worth mentioning that these metrics are also among the top three metrics with the highest average correlation percentage in vegetation pixels. Therefore, based on the results, they represent the best options for vegetation detection in hyperspectral images.

According to the previous analysis, Figure 7 presents the results of applying the direct correlation and Pearson correlation methods to the hyperspectral image of the Manga neighborhood in Cartagena. This was completed using the spectral library in Python and considering a threshold correlation percentage of 95%. In this regard, the implemented vegetation detection algorithm traverses each pixel of the image matrix, correlating each pixel of the image (a 380 positions vector) with the average vegetation pixel (380 positions) using direct correlation and Pearson correlation to determine the correlation percentage. If the correlation percentage is greater than 98%, the corresponding pixel in the displayed image is colored blue. Visually, it can be observed that in both cases, the detection algorithm adequately identifies the regions where vegetation is present, and the density of the detected vegetation pixels depends on the threshold adjustment for each correlation.

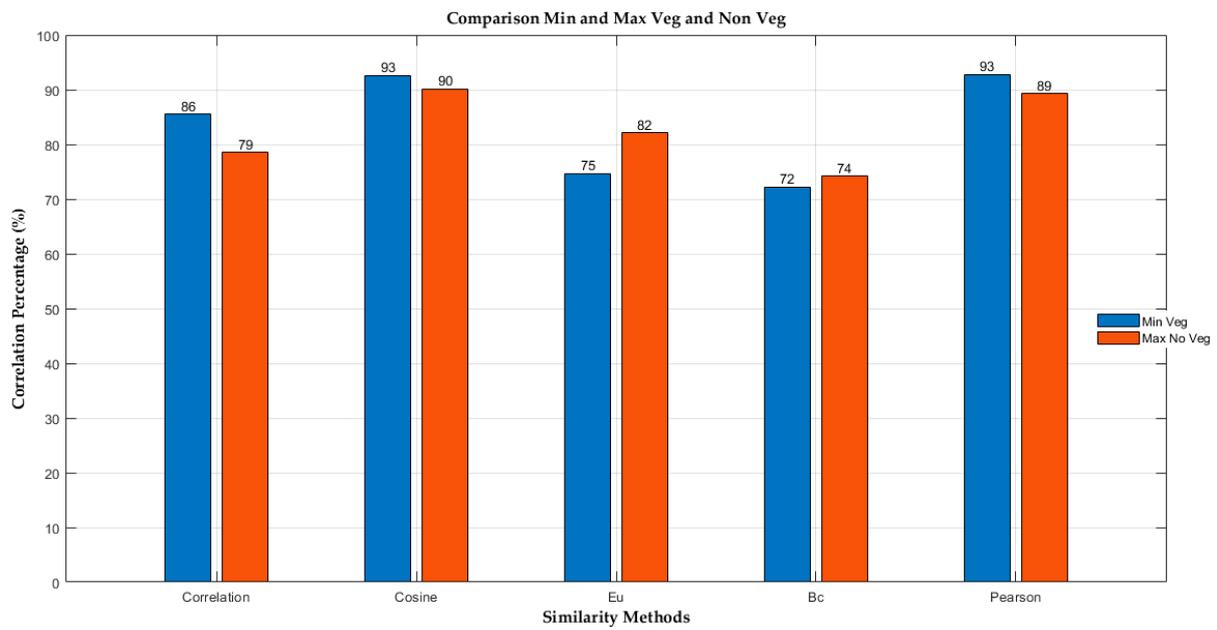
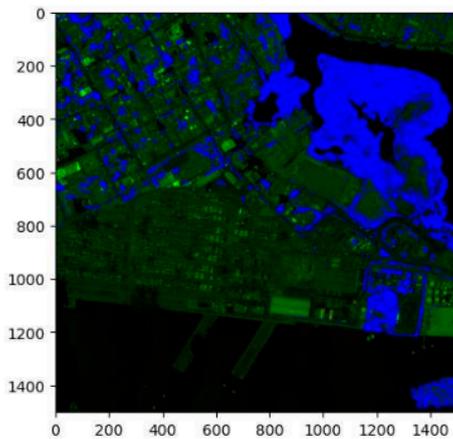


Figure 6. Comparison between minimum and maximum thresholds in vegetation and non-vegetation pixels.

```
In [26]: arr_copia=np.copy(arr)
tam=arr.shape
cont_corr=0
for i in range (0, tam[0]):
    for j in range(0,tam[1]):
        arr_temp=arr[i,j,:]
        nz=arr_temp.size - np.count_nonzero(arr_temp)
        if(nz!=380):
            corr=ssd.correlation(pix_prom,arr_temp)
            porc=(1-corr)*100
            if(porc>95.0):
                arr_copia[i][j]=150
                cont_corr+=1
```

```
In [27]: view=imshow(arr_copia)
view
```

```
Out[27]: ImageView object:
Display bands      : [0, 190.0, 379]
Interpolation     : <default>
RGB data limits   :
R: [0.0, 14968.0]
G: [0.0, 15000.0]
B: [0.0, 150.0]
```



```
In [28]: arr_copia=np.copy(arr)
tam=arr.shape
cont_pear=0
for i in range (0, tam[0]):
    for j in range(0,tam[1]):
        arr_temp=arr[i,j,:]
        nz=arr_temp.size - np.count_nonzero(arr_temp)
        if(nz!=380):
            corr_pr=np.corrcoef(pix_prom, arr_temp)
            mean_pr=np.mean(np.absolute(corr_pr))
            porc=mean_pr*100
            if(porc>95.0):
                arr_copia[i][j]=150
                cont_pear+=1
```

```
In [29]: view=imshow(arr_copia)
view
```

```
Out[29]: ImageView object:
Display bands      : [0, 190.0, 379]
Interpolation     : <default>
RGB data limits   :
R: [0.0, 14968.0]
G: [0.0, 15000.0]
B: [0.0, 150.0]
```

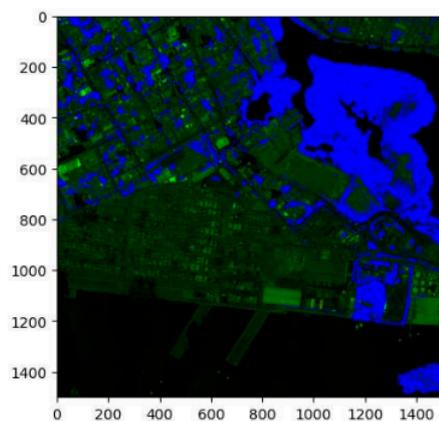


Figure 7. Application of direct correlation and Pearson correlation metrics on the hyperspectral test image.

Finally, by performing a count on the vegetation pixels of the hyperspectral image of the Manga neighborhood that were detected by the algorithm based on correlation metrics, Pearson correlation, and cosine distance, with a 95% threshold, the percentages of detected vegetation were obtained, as presented in Figure 8. It can be observed that, according to the results obtained by the considered metrics, between 15% and 17% of the pixels correspond to vegetation within the hyperspectral image.

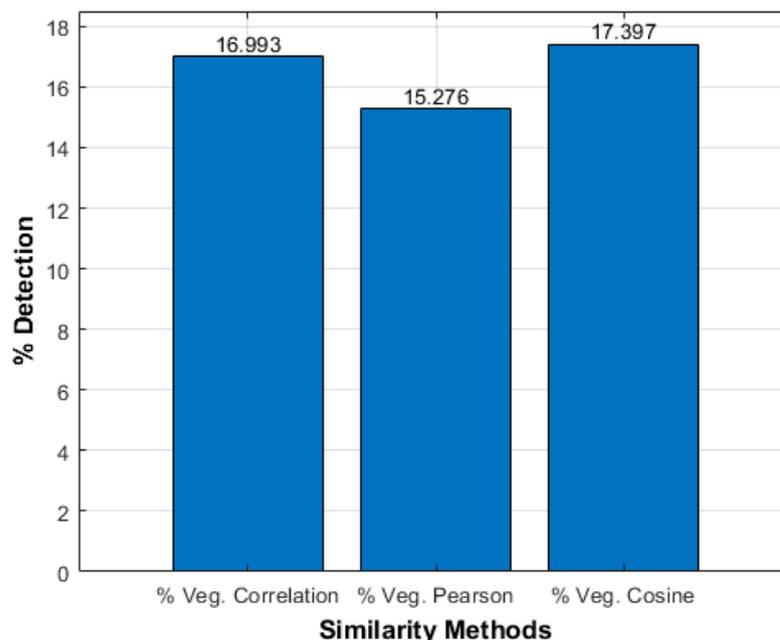


Figure 8. % Vegetation detection.

The present study introduces a novel approach to identify vegetation in hyperspectral images. This research offers valuable insights into the field of remote sensing and image analysis by utilizing distance and correlation metrics. The significance of this study is particularly relevant for developing countries, as it highlights the importance of open-access programming platforms for analyzing hyperspectral images. Accessible tools enable these countries to harness the potential of hyperspectral data, aiding in land monitoring, agricultural management, and environmental conservation efforts.

From a scientific standpoint, in relation to conducting research involving proprietary tools like Environment for Visualising Images Software (ENVI) for analyzing hyperspectral images [46,47], the utilization of open-source technologies like spectral, numpy, and other machine learning libraries such as scikit-learn enables the development of unhindered technological solutions capable of addressing specific remote sensing needs. This approach facilitates the integration of diverse artificial intelligence techniques, exemplifying its potential. These open-source tools allow for the implementation of unrestricted technological solutions that can cater to the unique requirements of remote sensing applications [48], thereby enhancing the accuracy, efficiency, and robustness of the hyperspectral and multispectral image analysis process.

Torres-Gil et al. [11] present distinct spectral signatures (asbestos, oil, soil, and vegetation) suggesting to analyze curve peaks as a viable approach for material detection and similarity determination. According to this, the present study has yielded empirical evidence supporting the effectiveness of the correlation method for comparing curve peaks. This effectiveness stems from the ability to attain a substantial spatial correlation by identifying points of alignment in both the highest and lowest peaks among the curves. However, the findings of the present research indicate that restricting the analysis to specific bands alone is inadequate for comparing the curves comprehensively. It is advisable to perform a correlation analysis between pixels employing the maximum feasible number of

bands, while also considering not only the utmost and lowest points but also the points of inflection.

Ultimately, concerning the techniques employed for vegetation analysis in hyperspectral image processing, specifically the normalized spectral mixture method as outlined in [49], our research introduces an alternative approach that incorporates representative pixel and correlation methods, facilitating the identification and characterization of vegetation in a broader sense. In contrast, the normalized spectral mixture method demonstrates enhanced precision in detecting objects exhibiting greater homogeneity or possessing a higher degree of purity. Consequently, the proposed methodology in our study can find practical application in urban environmental studies, specifically in investigating the correlation between overall green areas and population size. By incorporating the representative pixel and correlation techniques, our approach offers a comprehensive means of discerning vegetative elements within hyperspectral imagery. This methodology surpasses the limitations of solely relying on the normalized spectral mixture method, which primarily excels in detecting homogeneous and pure objects. Therefore, the methodology proposed in our research serves as a valuable tool for analyzing vegetation in a broader context, enabling a more comprehensive understanding of urban environmental dynamics and the relationship between green spaces and population density.

4. Conclusions

This study aimed to propose a comparative approach to assess the accuracy of different distance and correlation metrics in identifying vegetation within hyperspectral images of an urban area in the Manga neighborhood of Cartagena, Colombia. A total of 100 vegetation and non-vegetation pixels were selected from the hyperspectral image to create a reference or average pixel. Subsequently, correlations were computed using five distinct metrics, revealing that the direct correlation and Pearson metrics exhibited superior precision and discrimination of vegetation pixels, outperforming supervised learning-based methods.

Despite the differences between these two metrics, it was observed that the Euclidean and Bray–Curtis distance metrics encountered challenges in accurately classifying pixels as vegetation or non-vegetation. The proposed method based on correlation and/or vector distance metrics proved effective in differentiating spectral signatures within hyperspectral images. Visual inspection confirmed that the detection algorithm employing both metrics correctly identified vegetation in the designated areas, although the number of detected pixels could be adjusted based on specific correlation thresholds. The algorithm implemented using the most effective correlation metrics can be employed in studies related to determining the spatial distribution of vegetation across various latitudes. Furthermore, it can serve as a reference for detecting diverse objects or surfaces, such as water contaminants, asphalt, containers, and asbestos roofs, among others.

An innovative aspect of this study lies in the utilization of open-source software tools for accessing, processing, and analyzing hyperspectral images. The spectral library facilitated image retrieval, reading, and pixel extraction in vector representation, while the Python numpy and scipy libraries were employed for vector operations and correlation calculations. Additionally, the Python matplotlib library facilitated the management and visualization of pixel and hyperspectral image spectral signatures. These tools can serve as a valuable resource for testing and validating detection algorithms within the realm of hyperspectral imaging.

Future endeavors will involve incorporating machine learning models and specialized classification approaches for vegetation pixel detection in hyperspectral images. Furthermore, there are plans to test and calibrate the considered metrics for detecting asbestosement on urban rooftops.

Author Contributions: Conceptualization, G.E.C.G. and M.A.O.A.; methodology, G.E.C.G.; software, G.E.C.G.; validation, G.E.C.G. and M.A.O.A.; formal analysis, G.E.C.G.; investigation, G.E.C.G.; resources, M.S.; data curation, M.A.O.A.; writing—original draft preparation, G.E.C.G.; writing—review and editing, M.A.O.A. and M.S.; visualization, M.A.O.A.; supervision, M.S.; project administration, M.S.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This article is considered a product in the framework of the project “Formulation of an integral strategy to reduce the impact on public and environmental health due to the presence of asbestos in the territory of the Department of Bolivar”, financed by the General System of Royalties of Colombia (SGR) and identified with the code BPIN 2020000100366. This project was executed by the University of Cartagena, Colombia, and the Asbestos-Free Colombia Foundation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank David Enrique Valdelamar Martinez.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Pérez-Cabello, F.; Montorio, R.; Alves, D.B. Remote sensing techniques to assess post-fire vegetation recovery. *Curr. Opin. Environ. Sci. Health* **2021**, *21*, 100251. [CrossRef]
- Andreatta, D.; Gianelle, D.; Scotton, M.; Dalponte, M. Estimating grassland vegetation cover with remote sensing: A comparison between Landsat-8, Sentinel-2 and PlanetScope imagery. *Ecol. Indic.* **2022**, *141*, 109102. [CrossRef]
- Sripada, R.P. *Determining In-Season Nitrogen Requirements for Corn Using Aerial Color-Infrared Photography*; North Carolina State University: Raleigh, NC, USA, 2005.
- Shikwambana, L.; Xongo, K.; Mashalane, M.; Mhangara, P. Climatic and Vegetation Response Patterns over South Africa during the 2010/2011 and 2015/2016 Strong ENSO Phases. *Atmosphere* **2023**, *14*, 416. [CrossRef]
- García-Pardo, K.A.; Moreno-Rangel, D.; Domínguez-Amarillo, S.; García-Chávez, J.R. Remote sensing for the assessment of ecosystem services provided by urban vegetation: A review of the methods applied. *Urban For. Urban Green.* **2022**, *74*, 127636. [CrossRef]
- Neinavaz, E.; Schlerf, M.; Darvishzadeh, R.; Gerhards, M.; Skidmore, A.K. Thermal infrared remote sensing of vegetation: Current status and perspectives. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102415. [CrossRef]
- Meusbürger, K.; Bänninger, D.; Alewell, C. Estimating vegetation parameter for soil erosion assessment in an alpine catchment by means of QuickBird imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 201–207. [CrossRef]
- Henrich, V.; Krauss, G.; Götze, C.; Sandow, C. Index DataBase. A Database for Remote Sensing Indices. 2012. Available online: <https://www.indexdatabase.de/db/s-single.php?id=9> (accessed on 29 September 2022).
- Huang, S.; Tang, L.; Hupy, J.P.; Wang, Y.; Shao, G. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *J. For. Res.* **2021**, *32*, 1–6. [CrossRef]
- Hernández, D.H.B. *Aplicación de Índices de Vegetación para Evaluar Procesos de Restauración Ecológica en el Parque Forestal Embalse del Neusa*; Universidad Militar Nueva Granada: Neusa, Colombia, 2017.
- Gil, L.K.T.; Martínez, D.V.; Saba, M. The Widespread Use of Remote Sensing in Asbestos, Vegetation, Oil and Gas, and Geology Applications. *Atmosphere* **2023**, *14*, 172. [CrossRef]
- Birth, G.S.; McVey, G.R. Measuring the Color of Growing Turf with a Reflectance Spectrophotometer. *Agron. J.* **1968**, *60*, 640–643. [CrossRef]
- Wolf, A.F. Using WorldView-2 Vis-NIR multispectral imagery to support land mapping and feature extraction using normalized difference index ratios. In Proceedings of the Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII, SPIE Defense, Security, and Sensing, Baltimore, MD, USA, 23–27 April 2012; Volume 8390, pp. 188–195. [CrossRef]
- Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. *Monitoring Vegetation Systems in the Great Plains with Erts*; Nasa Special Publication; NASA: Washington, DC, USA, 1974.
- Kauth, R.J.; Thomas, G.S.P. The tasselled cap—A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, IN, USA, 29 June–1 July 1976.
- Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
- Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
- Crippen, R.E. Calculating the vegetation index faster. *Remote Sens. Environ.* **1990**, *34*, 71–73. [CrossRef]

19. Gitelson, A.A.; Merzlyak, M.N. Remote sensing of chlorophyll concentration in higher plant leaves. *Adv. Sp. Res.* **1998**, *22*, 689–692. [[CrossRef](#)]
20. Bannari, A.; Asalhi, H.; Teillet, P.M. Transformed difference vegetation index (TDVI) for vegetation cover mapping. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Toronto, ON, Canada, 24–28 June 2002; Volume 5, pp. 3053–3055. [[CrossRef](#)]
21. MaxMax. Enhanced Normalized Difference Vegetation Index (ENDVI). 2015. Available online: <https://www.maxmax.com/endvi.htm> (accessed on 26 September 2022).
22. Adão, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J.J. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sens.* **2017**, *9*, 1110. [[CrossRef](#)]
23. Wan, L.; Li, H.; Li, C.; Wang, A.; Yang, Y.; Wang, P. Hyperspectral Sensing of Plant Diseases: Principle and Methods. *Agronomy* **2022**, *12*, 1451. [[CrossRef](#)]
24. Wang, S.; Guan, K.; Zhang, C.; Jiang, C.; Zhou, Q.; Li, K.; Qin, Z.; Ainsworth, E.A.; He, J.; Wu, J.; et al. Airborne hyperspectral imaging of cover crops through radiative transfer process-guided machine learning. *Remote Sens. Environ.* **2023**, *285*, 113386. [[CrossRef](#)]
25. Khan, A.; Vibhute, A.D.; Mali, S.; Patil, C.H. A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications. *Ecol. Inform.* **2022**, *69*, 101678. [[CrossRef](#)]
26. Chen, D.; Zhang, F.; Tan, M.L.; Chan, N.W.; Shi, J.; Liu, C.; Wang, W. Improved Na+ estimation from hyperspectral data of saline vegetation by machine learning. *Comput. Electron. Agric.* **2022**, *196*, 106862. [[CrossRef](#)]
27. Gakhar, S.; Tiwari, K.C. Spectral–Spatial urban target detection for hyperspectral remote sensing data using artificial neural network. *Egypt. J. Remote Sens. Sp. Sci.* **2021**, *24*, 173–180. [[CrossRef](#)]
28. Ma, B.; Zeng, W.; Hu, G.; Cao, R.; Cui, D.; Zhang, T. Normalized difference vegetation index prediction based on the delta downscaling method and back-propagation artificial neural network under climate change in the Sanjiangyuan region, China. *Ecol. Inform.* **2022**, *72*, 101883. [[CrossRef](#)]
29. Trombetti, M.; Riaño, D.; Rubio, M.A.; Cheng, Y.B.; Ustin, S.L. Multi-temporal vegetation canopy water content retrieval and interpretation using artificial neural networks for the continental USA. *Remote Sens. Environ.* **2008**, *112*, 203–215. [[CrossRef](#)]
30. Davies, B.F.R.; Gernez, P.; Geraud, A.; Oiry, S.; Rosa, P.; Zoffoli, M.L.; Barillé, L. Multi- and hyperspectral classification of soft-bottom intertidal vegetation using a spectral library for coastal biodiversity remote sensing. *Remote Sens. Environ.* **2023**, *290*, 113554. [[CrossRef](#)]
31. Badola, A.; Panda, S.K.; Roberts, D.A.; Waigl, C.F.; Jandt, R.R.; Bhatt, U.S. A novel method to simulate AVIRIS-NG hyperspectral image from Sentinel-2 image for improved vegetation/wildfire fuel mapping, boreal Alaska. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102891. [[CrossRef](#)]
32. Rumpf, T.; Mahlein, A.-K.; Steiner, U.; Oerke, E.-C.; Dehne, H.-W.; Plümer, L. Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Comput. Electron. Agric.* **2010**, *74*, 91–99. [[CrossRef](#)]
33. Wang, L.; Wang, Q. Fast spatial-spectral random forests for thick cloud removal of hyperspectral images. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102916. [[CrossRef](#)]
34. Ding, X.; Wang, Q.; Tong, X. Integrating 250 m MODIS data in spectral unmixing for 500 m fractional vegetation cover estimation. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *111*, 102860. [[CrossRef](#)]
35. Shore, S.N. Astrochemistry. In *Encyclopedia of Physical Science and Technology*; Academic Press: Cambridge, MA, USA, 2003; pp. 665–678. [[CrossRef](#)]
36. Galle, N.J.; Brinton, W.; Vos, R.; Basu, B.; Duarte, F.; Collier, M.; Ratti, C.; Pilla, F. Correlation of WorldView-3 spectral vegetation indices and soil health indicators of individual urban trees with exceptions to topsoil disturbance. *City Environ. Interact.* **2021**, *11*, 100068. [[CrossRef](#)]
37. Thorp, K.R.; French, A.N.; Rango, A. Effect of image spatial and spectral characteristics on mapping semi-arid rangeland vegetation using multiple endmember spectral mixture analysis (MESMA). *Remote Sens. Environ.* **2013**, *132*, 120–130. [[CrossRef](#)]
38. Zhu, Y.; Zhang, Y.; Zheng, Z.; Liu, Y.; Wang, Z.; Cong, N.; Zu, J.; Tang, Z.; Zhao, G.; Gao, J.; et al. Converted vegetation type regulates the vegetation greening effects on land surface albedo in arid regions of China. *Agric. For. Meteorol.* **2022**, *324*, 109119. [[CrossRef](#)]
39. Smyth, T.A.G.; Wilson, R.; Rooney, P.; Yates, K.L. Extent, accuracy and repeatability of bare sand and vegetation cover in dunes mapped from aerial imagery is highly variable. *Aeolian Res.* **2022**, *56*, 100799. [[CrossRef](#)]
40. Tian, J.; Zhang, Z.; Philpot, W.D.; Tian, Q.; Zhan, W.; Xi, Y.; Wang, X.; Zhu, C. Simultaneous estimation of fractional cover of photosynthetic and non-photosynthetic vegetation using visible-near infrared satellite imagery. *Remote Sens. Environ.* **2023**, *290*, 113549. [[CrossRef](#)]
41. Lyons, R. Distance covariance in metric spaces. *Ann. Probab.* **2013**, *41*, 3284–3305. [[CrossRef](#)]
42. Székely, G.J.; Rizzo, M.L. Brownian distance covariance. *Ann. Appl. Stat.* **2009**, *3*, 1236–1265. [[CrossRef](#)]
43. Connor, R. A tale of four metrics. In *International Conference on Similarity Search and Applications—SISAP 2016*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2016; Volume 9939, pp. 210–217. [[CrossRef](#)]
44. Bray, J.R.; Curtis, J.T. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Source Ecol. Monogr.* **1957**, *27*, 325–349. [[CrossRef](#)]

45. Novotn, V.T. Implementation notes for the soft cosine measure. In Proceedings of the CIKM '18: 27th ACM International Conference on Information and Knowledge Management, Torino, Italy, 22–26 October 2018; pp. 1639–1642. [[CrossRef](#)]
46. Curcio, A.C.; Barbero, L.; Peralta, G. UAV-Hyperspectral Imaging to Estimate Species Distribution in Salt Marshes: A Case Study in the Cadiz Bay (SW Spain). *Remote Sens.* **2023**, *15*, 1419. [[CrossRef](#)]
47. ESRI. ENVI 2023. 2023. Available online: <https://www.esri.com/partners/13harris-technologie-a2T39000001dNCnEAM/envi-a2d5x000005jPrfAAE> (accessed on 28 June 2023).
48. Duarte, L.; Teodoro, A.C.; Monteiro, A.T.; Cunha, M.; Gonçalves, H. QPhenoMetrics: An open source software application to assess vegetation phenology metrics. *Comput. Electron. Agric.* **2018**, *148*, 82–94. [[CrossRef](#)]
49. Zhang, Y.; Wang, Y.; Ding, N. Spatial Effects of Landscape Patterns of Urban Patches with Different Vegetation Fractions on Urban Thermal Environment. *Remote Sens.* **2022**, *14*, 5684. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.