SPATIAL OPTICS OF THE SENTINEL-3B OLCI SATELLITE IN AMAZONIAN RIVERS:

a potential tool to understand environmental impacts with Machine Learning modeling potential

Julia de Almeida[1]; Júlia Mognol Scopel[1]; Isadora Cezar Caino[1]; Guilherme Peterle Schmitz[1] e Alcindo Neckel[1]

[1]Atitus Educação, Passo Fundo/RS, Brasil. julia.allm659@gmail.com; juliascopel10@gmail.com; isadoracaino@gmail.com; guilhermepeterle99@gmail.com; alcindo.neckel@atitus.edu.br

Abstract: *The Tapajós River is a tributary to the Amazon. Sentinel-3B OLCI (Ocean Land Color Instrument) Level-2 satellite imagery with Water Full Resolution (WFR) of 300 m was utilized to detect the highest potential for the absorption coefficient of detritus and gelbstoff in 443 m^{-1} (ADG443_NN), chlorophyll-a (CHL_NN) and total suspended matter concentration (TSM_NN), at 25 points in the Amazon and Tapajós rivers (in 2019 and 2021). The Sentinel-3B OLCI images, based on the Neural Network (NN) were calibrated by the European Space Agency (ESA), with a standard average normalization of 0.83 µg/mg, containing a maximum error of 6.62% applied to the sampled points. This data generated by images from the Sentinel-3B OLCI satellite enhanced better Machine Learning analyses, resulting in future scenarios regarding the creation of new public policies aimed at preserving Amazon rivers.*

Keywords: *Sentinel-3b OLCI. Romete sensing. Geospatial analysis.*

Resumo: *O rio Tapajós é um afluente do rio Amazonas. Imagens de satélite Sentinel-3B OLCI (Ocean Land Color Instrument) de Nível 2, com resolução total da água (WFR) de 300 m, foram utilizadas para detectar o maior potencial do coeficiente de absorção de detritos e gelbstoff em 443 m^{-1} (ADG443_NN), clorofila-a (CHL_NN) e concentração de material particulado total (TSM_NN), em 25 pontos nos rios Amazonas e Tapajós (em 2019 e 2021).As imagens Sentinel-3B OLCI, baseadas em Rede Neural (NN), foram calibradas pela Agência Espacial Europeia (ESA), com uma normalização média padrão de 0,83 µg/mg, apresentando um erro máximo de 6,62% aplicado aos pontos amostrados. Esses dados gerados pelas imagens do satélite Sentinel-3B OLCI aprimoraram as análises de Aprendizado de Máquina (Machine Learning), resultando em cenários futuros voltados à criação de novas políticas públicas destinadas à preservação dos rios amazônicos.*

Palavras chave: *Sentinel-3b OLCI. Sensoriamento remoto. Análise geoespacial.*

Resumen: *El río Tapajós es un afluente del Amazonas. Se utilizaron imágenes satelitales Sentinel-3B OLCI (Ocean Land Color Instrument) de Nivel 2, con resolución total del agua (WFR) de 300 m, para detectar el mayor potencial del coeficiente de absorción de detritos y gelbstoff en 443 m^{-1} (ADG443_NN), clorofila-a (CHL_NN) y la concentración de materia en suspensión total (TSM_NN), en 25 puntos de los ríos Amazonas y Tapajós (en 2019 y 2021). Las imágenes Sentinel-3B OLCI, basadas en redes neuronales (NN), fueron calibradas por la Agencia Espacial Europea (ESA), con una normalización promedio estándar de 0,83 µg/mg y un error máximo de 6,62% aplicado a los puntos muestreados. Estos datos generados por las imágenes del satélite Sentinel-3B OLCI mejoraron los análisis de aprendizaje automático (Machine Learning), resultando en escenarios futuros orientados a la creación de nuevas políticas públicas destinadas a la preservación de los ríos amazónicos.*

Palabras clave:*Sentinel-3b OLCI. Teledetección. Análisis geoespacial.*

1. Introduction

The confluence of the Amazon and Tapajós rivers occurs in the city of Santarém, in the Brazilian state of Pará (Teixeira *et al*., 2021). The practice of clandestine gold mining negatively impacts more than 300,000 inhabitants of the city of Santarém, in addition to strongly compromising riverine and soil health and regional tourism in the Tapajós River watershed (Teixeira *et al*., 2021). It becomes necessary to understand the proportion of pollutants present in the particles in this upstream watershed in addition to demonstrating a need to identify potentially hazardous elements present in the riverbank sediments of the Amazon and Tapajós rivers. This must be done in order to alert the population and public authorities to the possible need for environmental remediation and preservation of these hydrological environments.

The main method of sediment transport in the Amazon Basin is via the action of water (Silva *et al*., 2023). The Amazon River and its tributaries run through large regions. This study uses satellite imagery to examine the absorption coefficient of detritus and gelbstoff in 443 m^{-1} (ADG443_NN), chlorophyll-a (CHL_NN) and total suspended matter concentration (TSM_NN g^{-3}) throughout the rivers' course as the rivers move over the landscape. One method of analysis that has no direct impact on land or the Amazonian ecosystem is the use of geospatial analysis carried out using images from the Sentinel-3B OLCI satellite, one of the most modern analytical techniques available today (Neckel *et al*., 2022). According to Neckel *et al*. (2022), by studying the variables of ADG443_NN, CHL_NN and TSM_NN in these images.

This remote sensing data, through the use of satellites and ground-based sensors, offers the ability to acquire data on a global scale, enabling a comprehensive assessment of changes in hydrographic environmental quality (Silva *et al*., 2023). Machine Learning, in turn, has revolutionized the processing and interpretation capacity of this geospatial data (Silva *et al*., 2023). Machine Learning algorithms, such as convolutional neural networks, have demonstrated effectiveness in identifying and classifying complex patterns present in remote sensing images (Silva *et al*., 2023).

The general objective of this study is to analyze the highest potential for the absorption coefficient of detritus and gelbstoff in 443 m^{-1} (ADG443_NN), chlorophyll-a (CHL_NN) and total suspended matter concentration (TSM_NN), at 25 points in the Amazon and Tapajós rivers (in 2019 and 2021) by Neural Network (NN) configured by ESA, together with the extrapolation of variables that allow the use of future Machine Learning techniques.

2. Material and methods

The headwaters of the Amazon River are located in the Andes Mountains of Peru. The river flows throughout the Amazon Basin, eventually emptying into the Atlantic Ocean in Brazilian territory (Salgado *et al*., 2021). Throughout the course of its journey, the Amazon is fed by over 1,000 tributaries encompassing the largest overall volume of riverine fresh water in the world. The Amazon flows through the territory of eight countries in South America: Brazil, Colombia, Peru, Venezuela, Bolivia, Guyana, French Guiana and Suriname. The basin is characterized by an equatorial humid climate with average temperatures varying from 25.8º C in winter to 27.9º C in summer (Salgado *et al*., 2021). According to Park and Latrubesse (2014), the Amazon River watershed area covers more than 7 million km^2, of which 4 million km^2 (63%) are in Brazilian territory. 23,000 km of navigable rivers and an estimated population exceeding 20.3 million inhabitants make up the Amazon Basin. The Amazon itself is the area's largest river, with a length of over 6,000 km (Park and Latrubesse, 2014). At 840 km in length, the Tapajós River is a major Amazonian tributary that supports considerable marine biodiversity and indigenous populations as it flows through the Amazon rainforest (Park and Latrubesse, 2014; Silva *et al*., 2023).

The accompanying geospatial methodological procedure utilized for this study was the use of cloud-free images from the Sentinel-3B OLCI satellite (Neckel *et al*., 2021) at the identical 25 fixed sampling points along the Amazon and Tapajós rivers. These images were analyzed based on the following Neural Network (NN), water color algorithms: surface dissolved material (ADG443_NN), chlorophyll-a (CHL_NN) and total suspended matter concentration (TSM_NN). In this study, the methodological procedures recommended by Silva *et al*. (2023) were followed.

The Sentinel-3B OLCI satellite images used were calibrated by the European Space Agency (ESA) with a standard average normalization of 0.83 µg/mg and a maximum error of 6.62% in relation to the 25 points sampled along the Amazon and Tapajós rivers. According to the Neckel *et al*. (2023), the images are already normalized, attributed to error according to the ESA system's use of SNAP software (Sentinel Application Platform). Researcher login enabled a geospatial application of high reliability, Level- 2 geospatial data processing, containing the mapping of variables on macroscales with uniform space and time consistency. Although the Sentinel-3B OLCI images used in this study were calibrated by ESA, it is important to highlight that the correction is performed by dark current and resampled in relation to the terrestrial surface of water analyzed.

The images utilized in this study were Sentinel-3B OLCI water images with 300 m full resolution (WFR), which corresponds to the product generated by the Sentinel-3 satellite that contains the OLCI Level-2 variable. The following parameters are associated with these images: considering a flow of quantum energy emitted by the sun in a spectral range of 400 to 700 nm, with aerosol charge, in optical depth with a wavelength of 865 nm and the optical depth of the aerosol containing the spectral dependence of 779 to 865 nm (Neckel *et al*., 2023). The authors generated quantitative data from these analyzed images for each of the examined parameters (ADG443_NN, CHL_NN and TSM_NN) in periods of flooding (April 6, 2019 and March 15, 2021) and in periods of low water levels (September 14, 2019 and September 16, 2021), totaling 4 images from the Sentinel-3 OLCI satellite where the 25 sampled points were collected. March and April are typically the two months of highest rainfall in the Amazon Basin, with rates of up to 300 mm monthly, whereas rainfall typically falls below 100 mm per month in the months of August and September (Silva *et al*., 2023). Thus, data collection was carried out by orthogeolocation in the Sentinel-3B OLCI satellite images, through the collected points (Neckel *et al*., 2023).

These data obtained from the analysis of the Sentinel-3B OLCI satellite images in relation to the SPM analytical results, generated a data sheet in CSV (Character-separated values), manipulated using JASP software (JASP workshops), version 0.14.1.0. The data were modeled in K-means clusters (Machine Learning) based on the position of the centroids in relation to the mean center of the resulting dataset (Naghizadeh and Metaxas, 2020). The algorithmic application considered the sum of squares error (SSE) in the internal area of the clusters allocated to the centroid with greater proximity based on the Silhouette index (Naghizadeh and Metaxas, 2020; Neckel *et al*., 2023). Naghizadeh and Metaxas (2020) demonstrated that the Silhouette index of each categorized cluster where the value obtained is greater than 1 presents high reliability, whereas a value lower than -1 indicate the presence of incorrect coherence.

3. Results and Discussion

Figure 1 consists of representations of the Sentinel-3B OLCI image compositions. Figure 1 (A and B) shows the difference between the dissolved material on the surface (ADG443_NN) of the Tapajós River, while the Amazon River remains unchanged. During the low waters in 2019, the Tapajós River presents a greater relative reduction in the volume of water due to the fact that it is characterized as a tributary and covers a smaller watershed, whereas the Amazon River continues to receive water volume from other tributaries throughout its much more extensive geographical watershed. Observations from the low flow periods of 2021 also yielded lower ADG443_NN readings in the Tapajós River than the Amazon River (Figure 1 (C and D). During high water flow periods, this difference in ADG443_NN between the Tapajós and Amazonas rivers becomes insignificant, regardless of the year sampled (Figure 1 (A and C). However, during the 2021 floods, a higher ADG443_NN is observed compared to the 2019 floods. The year 2021 had a greater volume of rainfall than historical averages, considerably increasing the volume of water in both the Amazon and Tapajós rivers (Silva *et al*., 2023). This provided a large carryover of ADG443_NN during times with higher rainfall. During the hydrological low season in the year 2021, there was a slight reduction in the amount of ADG443_NN compared to the low water flow season of 2019.

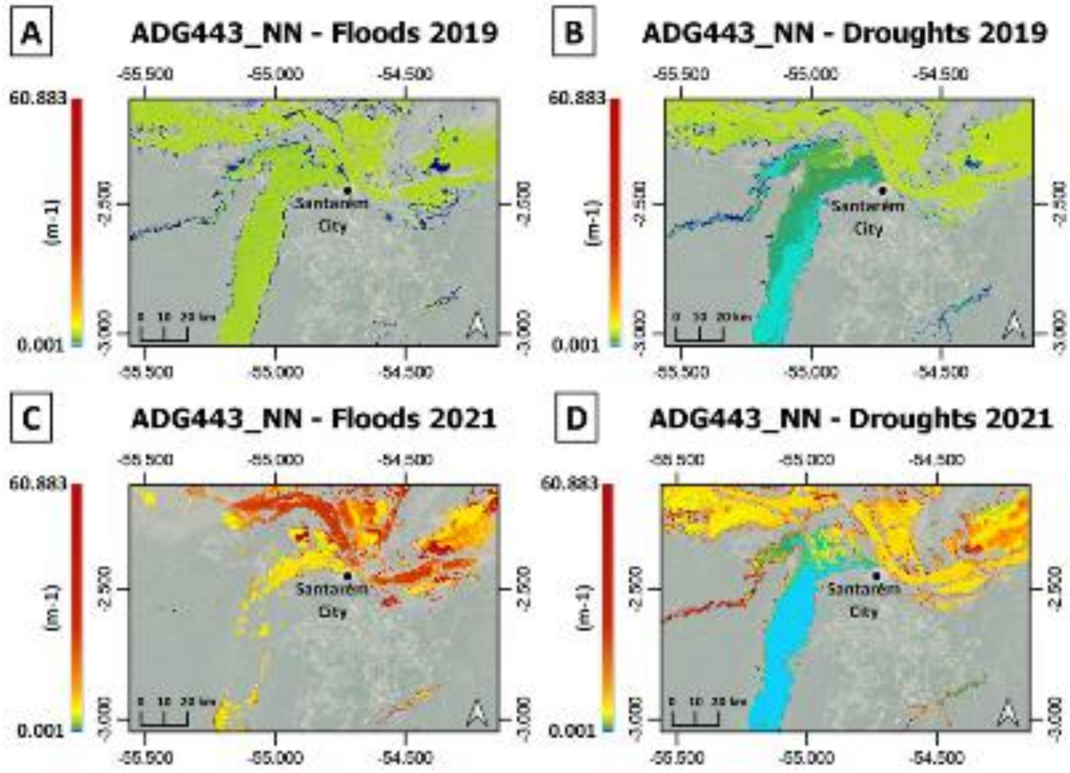

Figure 1. Compositions of Sentinel-3B OLCI satellite imagery, variable: material dissolved in the water surface (ADG443_NN), (A) material dissolved on the water surface during the 2019 floods, (B) material dissolved on the water surface during the 2019 lows, (C) material dissolved on the water surface during the 2021 floods, (D) dissolved material on the water surface during the 2021 lows.

Through Figure 2 (A and B) it is possible to observe that the amount of chlorophyll-a in the water (CHL_NN) decreased during low hydrological flows as compared to the flood season in the Tapajós River, with no perceivable differences in the Amazon River. This is due to the fact that the Amazon always contains high levels of sediments throughout the year (Merschel *et al*., 2017), whereas the Tapajós presents greater turbidity levels primarily during the flood season, attributed to either greater agitation of its precipitated sediments, a further erosion of its banks, or both. Another factor that may influence the greater presence of sediments in the Tapajós River during the flood period is the direct natural injection of sediments from the Amazon into the Tapajós as the two rivers overspill their banks. The influence of turbidity of river mouths and adjacent regions during periods of high flows is evident (Berzas Nevado *et al*., 2010).

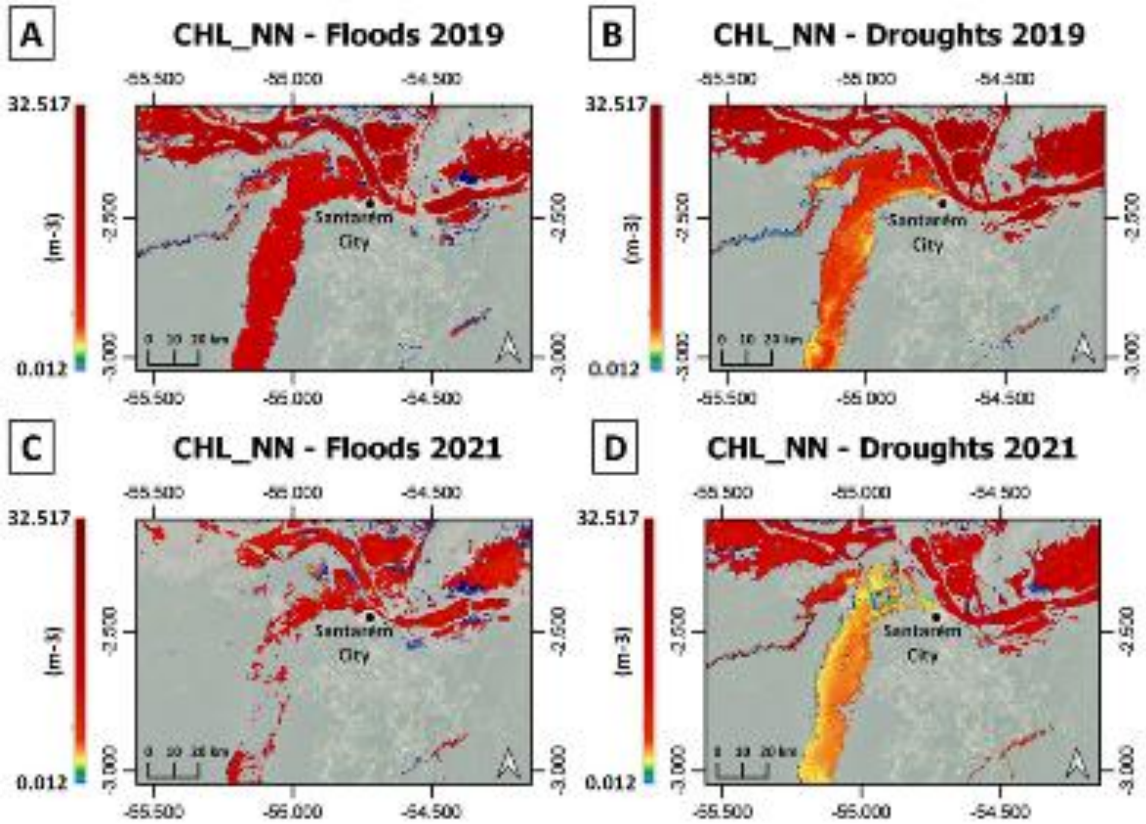

Figure 2. Compositions of Sentinel-3B OLCI satellite imagery, variable: chlorophyll-a concentrations (CHL_NN), (A) chlorophyll-a concentrations in water during the 2019 floods, (B) chlorophyll-a concentrations in water during the 2019 hydrologic lows, (C) chlorophyll-a concentrations in water during the 2021 floods, (D) chlorophyll-a concentrations in water during the hydrological lows of 2021.

Despite this, during sampled periods of high-water flow, the CHL_NN is not altered in relation to the sampled period (Figure 2 (A and C)). However, during the low flow periods of 2021, a reduction in the amount of CHL_NN was observed as compared to the 2019 low flow season (Figure 2 (B and D)). The behavior of CHL_NN was very similar to the behavior of ADG443_NN. The large volume of water that flows into the Amazon River during the flood season carries with it a large amount of material (soil, sediment, wood, mining waste, agricultural waste) (Berzas Nevado *et al*., 2010).

In Figure 3, great differences in total suspended matter concentration were observed during the low flow levels of the Amazon and Tapajós rivers. There was a reduction in TSM_NN in these times compared to the flood season, in both 2019 and 2021. The difference between the Amazon and Tapajós rivers were not at all evident in the 2019 flood season (Figure 3(A)), whereas in the other periods there was an observed departure in the hydrological dynamics of the analyzed rivers (Figures 3 (B, C and D).

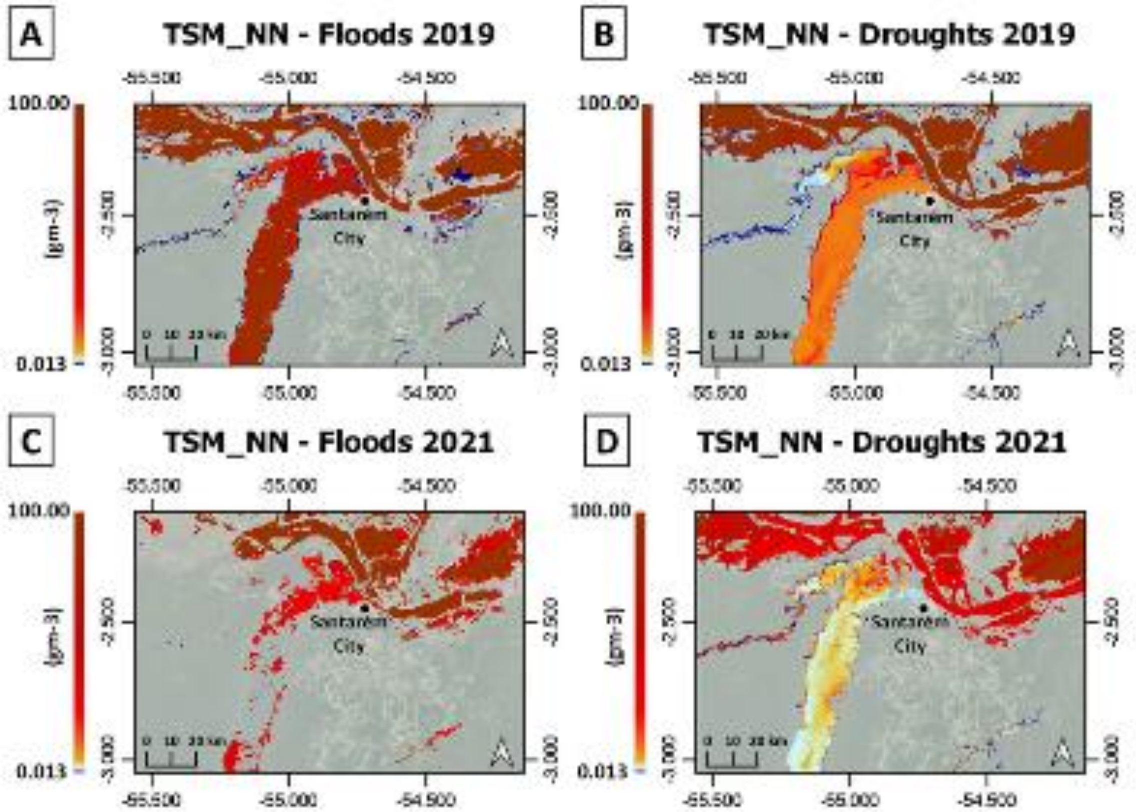

Figure 3. Compositions of Sentinel-3B OLCI satellite imagery, variable: total suspended matter concentration (TSM_NN), (A) water turbidity during the 2019 floods, (B) water turbidity during the 2019 lows, (C) water turbidity during the 2021 floods, (D) water turbidity during the 2021 hydrological lows.

Table 1 demonstrates the linear descriptive statistics of the data from the analysis of Sentinel-3B OLCI images, according to the application method outlined by Neckel *et al*. (2022). ADG443_NN was at its maximum value (55.475 m^{-1}) during the 2021 flood period, much higher than in 2019 (3.494 m^{-1}). The highest average of CHL_NN (16.341 m^{-3}) was also observed in the 2019 flood period However, in the floods of 2021, a corresponding value of 11.155 m^{-3} was recorded, with a small reduction in the amount of CHL_NN. The standard deviation values (CHL_NN) of the flood seasons of the analyzed period of the years 2019 and 2021 were 3.103 and 2.624, respectively. The minimum value in the flood season of the Amazon and Tapajós rivers of 2019 was 0.39 m^{-3}; while the maximum value of CHL_NN was 12.658 m^{-3}. The highest CHL_NN values were measured in 2019, regardless of the volume of water in the Amazon river, with CHL_NN values of 20.281 m^{-3} and 21.806 m^{-3}, while in 2021 the CHL_NN values were 14.113 m^{-3} and 15.734 m^{-3}. The TSM_NN averages were higher in the flood season, with 91.428 gm^{-3} in 2019 and 70.787 gm^{-3} in 2021, while in the low flow season, the averages of TSM_NN were 59.714 gm^{-3} in 2019 and 25.106 gm^{-3} in 2021, respectively. In the low flow season of 2021 there was a drastic reduction in the amount of TSM_NN.

Table 1. Descriptive Statistics from the sampling of the 25 points sampled in the Sentinel-3B OLCI images.

Items	ADG443_Floods2019	CHL_Floods2019	TSM_Floods2019
Valid	25	25	25
Missing	0	0	0

Mean	2.653	16.341	91.428
Std. Deviation	0.251	3.103	15.903
Minimum	2.221	12.658	48.422
Maximum	3.494	20.281	100.00
Items	ADG443_Droughts2019	CHL_Droughts2019	TSM_Droughts2019
Valid	25	25	25
Missing	0	0	0
Mean	2.048	13.048	59.714
Std. Deviation	1.114	7.131	44.981
Minimum	0.498	3.077	7.619
Maximum	3.050	21.806	100.00
Items	ADG443_Floods2021	CHL_Floods2021	TSM_Floods2021
Valid	25	25	25
Missing	0	0	0
Mean	16.303	11.155	70.787
Std. Deviation	12.483	2.624	31.812
Minimum	2.785	0.390	10.559
Maximum	55.475	14.113	100.00
Items	ADG443_Droughts2021	CHL_Droughts2021	TSM_Droughts2021
Valid	25	25	25
Missing	0	0	0
Mean	4.251	8.961	25.106
Std. Deviation	3.374	5.882	22.638
Minimum	0.346	0.624	1.840
Maximum	8.650	15.734	69.357

The K-means (Machine Learning) analysis consists of the adjustment scores of the statistical model, where K = 3 clusters for the set represented by the data of the 25 collection points carried out from the Sentinel-3B OLCI satellite images, generating the R^2 value of 0.856. This value demonstrates that the model presents a good fit for the research question, as it is close to the value of 1 (Naghizadeh and Metaxas, 2020). The silhouette index for k=3 clusters was 0.69, demonstrating that there is a good tendency to cluster the clusters according to the established variables. As for heterogeneity, the total value of 1 is proportionally divided (Naghizadeh and Metaxas, 2020). In this, Cluster 1 is the largest with 13 collection points and a heterogeneity rating of 0.293, demonstrating low overall heterogeneity. Meanwhile, Cluster 3 is the most heterogeneous, with 11 points and a rating of 0.707, demonstrating high overall heterogeneity. Cluster 2 has only one collection point, due to this, it has no heterogeneity and no sum of squares.

The clustering results are linked to the geospatial visualization maps, showing the clustering results according to the 25 collection points. The location of the clusters in relation to the study area is of fundamental importance to understand the water quality variables (ADG443_NN, CHL_NN and TSM_NN) collected using the Sentinel-3B OLCI images (Vanhellemont and Ruddick, 2021; Neckel *et al*. 2022). Cluster 1 is located in the Amazon River and requires a great amount of monitoring and care by public municipalities and authorities due to its enormous pollutant transport potential. High values of dissolved material on the surface (ADG443_NN), concentrations of chlorophyll-a (CHL_NN) and total suspended matter concentration (TSM_NN) predominated in all Cluster 1 analyzed environments. Cluster 3 is located on the Tapajós River, has a lower overall amount of suspended sediment, representing better water quality, regardless of the level of precipitation or year examined. In Cluster 2, the period of high water is the opposite, as it is a region of floodplain in the area of the confluence of the two rivers. Amazonian waters enter the Tapajós channel during high Amazonian flows.

4. Conclusion

The analysis of Sentinel-3B OLCI images showed that in 2021, the Amazon and Tapajós rivers exhibited high values of ADG443_NN (55.475 m^{-1}), CHL_NN (15.734 m^{-3}), and TSM_NN (70.787 g/m^3), indicating significant sediment transport. It is recommended that public authorities use this study as a basis to improve the management and monitoring of these rivers.

References

Berzas Nevado, J.J., Rodríguez Martín-Doimeadios, R.C., Guzmán Bernardo, F.J., Jiménez Moreno, M., Herculano, A.M., do Nascimento, J.L.M. and Crespo-López, M.E. Mercury in the Tapajós River basin, Brazilian Amazon: A review. *Environment International*, 36: 593-608, 2010.

Merschel, G., Bau, M., Schmidt, K., Münker, C. and Dantas, E.L. Hafnium and neodymium isotopes and REY distribution in the truly dissolved, nanoparticulate/colloidal and suspended loads of rivers in the Amazon Basin, Brazil. *Geochimica et Cosmochimica Acta*, 213: 383-399, 2017.

Naghizadeh, A. and Metaxas, D.N. Condensed Silhouette: An Optimized Filtering Process for Cluster Selection in K-Means. *Procedia Computer Science*, 176: 205-214, 2020.

Neckel, A., Oliveira, M.L.S., Castro Bolaño, L.J., Maculan, L.S., Moro, L.D., Bodah, E.T., Moreno-Ríos, A.L., Bodah, B.W. and Silva, L.F.O. Biophysical matter in a marine estuary identified by the Sentinel-3B OLCI satellite and the presence of terrestrial iron (Fe) nanoparticles. *Marine Pollution Bulletin*, 173: 112925, 2021.

Neckel, A., Oliveira, M.L., Dotto, G.L., Stolfo Maculan, L., William Bodah, B., and Silva, L.F. Sentinel-3 OLCI satellite imagery and advanced electron microscopy of nanoparticle analysis in a marine estuary and watershed: Robust multi-analytics and geospatial assessment of key contaminants. *Journal of Hydrology*, 612, 128278, 2022.

Neckel, A., Oliveira, M.L., Maculan, L.S., Adelodun, B., Toscan, P.C., Bodah, B.W., Moro, L.D., and Silva, L.F. Terrestrial nanoparticle contaminants and geospatial optics using the Sentinel-3B OLCI satellite in the Tinto River estuary region of the Iberian Peninsula. *Marine Pollution Bulletin,* 187: 114525, 2023.

Park, E. and Latrubesse, E.M. Modeling suspended sediment distribution patterns of the Amazon River using MODIS data. *Remote Sensing of Environment*, 147: 232-242, 2014.

Salgado, A.A.R., Ribeiro Marent, B. and Paixão, R.W. Large rivers, slow drainage rearrangements: The ongoing fluvial piracy of a major river by its tributary in the Branco River Basin - Northern Amazon. *Journal of South American Earth Sciences*, 112: 103598, 2021.

Silva, L., Bodah, B. W., Lozano, L. P., Oliveira, M. L., Korcelski, C., Maculan, L. S. and Neckel, A. Nanoparticles containing hazardous elements and the spatial optics of the Sentinel-3B OLCI satellite in Amazonian rivers: a potential tool to understand environmental impacts. *Environmental Science and Pollution Research*, 30: 73780-73798, 2023.

Teixeira, R.A., Pereira, W.V.D.S., Souza, E.S.D., Ramos, S.J., Dias, Y.N., Lima, M.W.D., Souza Neto, H.F.de., Oliveira, E.S.D. and Fernandes, A.R. Artisanal gold mining in the eastern Amazon:

Environmental and human health risks of mercury from different mining methods. *Chemosphere*, 284: 131220, 2021.

Vanhellemont, Q. and Ruddick, K. Atmospheric correction of Sentinel-3/OLCI data for mapping of suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. *Remote Sensing of Environment*, 256: 112284, 2