

# DELIMITATION OF FLOODED AREAS BASED ON SENTINEL-1 SAR DATA PROCESSED THROUGH MACHINE LEARNING: A STUDY CASE FROM CENTRAL AMAZON, BRAZIL

IVO AUGUSTO LOPES MAGALHÃES<sup>1</sup> 

OSMAR ABILIO DE CARVALHO JUNIOR<sup>1</sup> 

EDSON EYJI SANO<sup>2</sup> 

**ABSTRACT** – Delimitation of areas subject to flooding is crucial to understand water dynamics and fluvial changes. This study analyzed the potential of C-band *Synthetic Aperture Radar* (SAR) images acquired by the Sentinel-1 satellite in 2017, 2018, and 2019 to delineate flooded areas in the Central Amazon. The images were processed by the *Artificial Neural Network Multi-Layer Perceptron* (ANN-MLP) and two *K-Nearest Neighbor* (KNN-7 and KNN-11) machine learning (ML) classifiers. Pre-processing of *Single Look Complex* (SLC) SAR images involved the following methodological steps: *orbit-file* application; radiometric calibration ( $\sigma^0$ ); *Range-Doppler* terrain correction; speckle noise filtering; and conversion of linear data to backscattering coefficients (units in dB). We applied the *Lee* filter, with a window size of 3x3, for speckle filtering. A set of 6000 randomly distributed samples for training (70%), validation (20%), and test (10%) was obtained based on visual interpretation of Sentinel-2 optical satellite image acquired in the same years of SAR images. We found the largest flooded areas in 2019 in the study area (municipality of Parintins and Uruará, Amazonas River, Brazil): 6244km<sup>2</sup> by the ANN-MLP classifier; 6268km<sup>2</sup> by KNN-7; and 6290km<sup>2</sup> by KNN-11, while the smallest flooded areas were found in 2018: 5364km<sup>2</sup> by ANN-MLP; 5412km<sup>2</sup> by KNN-7; and 5535km<sup>2</sup> by KNN-11. The three classifiers presented *Kappa* coefficients between 0.77 and 0.91. ANN-MLP showed the best accuracy. The presence of shadow effects in the SAR images increased the commission errors.

**Keywords:** Remote sensing; water resources; image classifiers; inundation.

---

Recebido: 15/04/2023. Aceite: 14/07/2023. Publicado: 16/08/2023.

<sup>1</sup> Campus Universitário Darcy Ribeiro, ICC Norte, Mezanino B1 573, 70.910-900, Brasília, DF, Brasil. E-mail: [ivosrmagalhaes@gmail.com](mailto:ivosrmagalhaes@gmail.com); [osmarjr@unb.br](mailto:osmarjr@unb.br)

<sup>2</sup> Empresa Brasileira de Pesquisa Agropecuária, EMBRAPA Cerrados, Brasília/Fortaleza Planaltina, DF, Brasil. E-mail: [edson.sano@gmail.com](mailto:edson.sano@gmail.com)

**RESUMO** – DELIMITAÇÃO DE ÁREAS INUNDADAS COM BASE EM DADOS SAR SENTINEL-1 PROCESSADOS ATRAVÉS DE APRENDIZAGEM DE MÁQUINA: UM ESTUDO DE CASO NA AMAZÔNIA CENTRAL, BRASIL. A delimitação de áreas sujeitas a inundações é crucial para entender a dinâmica hídrica e as mudanças fluviais. Este estudo analisou o potencial de imagens de radar de abertura sintética (SAR) adquiridas na banda-C pelo satélite Sentinel-1 em 2017, 2018 e 2019 para delinear áreas inundadas na Amazônia Central. As imagens foram processadas pela *Rede Neural Artificial Multi-Layer Perceptron* (RNA-MLP) e dois classificadores de aprendizagem de máquina (ML) *K-Nearest Neighbor* (KNN-7 e KNN-11). O pré-processamento de imagens SAR *Single Look Complex* (SLC) envolveu as seguintes etapas metodológicas: aplicação do *orbit-file*; calibração radiométrica ( $\sigma^0$ ); correção de terreno *Range-Doppler*; filtragem de ruído *speckle*; e conversão de dados lineares para coeficientes de retroespalhamento (unidades em dB). O filtro de Lee com tamanho de janela de 3x3 foi aplicado para filtragem do ruído *speckle*. Um conjunto de 6000 amostras distribuídas aleatoriamente para treino (70%), validação (20%) e teste (10%) foi obtido com base na interpretação visual da imagem do satélite óptico Sentinel-2 adquiridas no mesmo ano das imagens de radar. As maiores áreas alagadas foram encontradas em 2019 na área de estudo (municípios de Parintins e Uruará, Rio Amazonas, Brasil): 6244km<sup>2</sup> pelo classificador RNA-MLP; 6268km<sup>2</sup> pelo KNN-7; e 6290km<sup>2</sup> pelo KNN-11, enquanto as menores áreas alagadas foram encontradas em 2018: 5364km<sup>2</sup> pelo classificador RNA-MLP; 5412km<sup>2</sup> pelo KNN-7; e 5535km<sup>2</sup> pelo KNN-11. Os três classificadores apresentaram coeficientes *Kappa* entre 0,77 e 0,91. A RNA-MLP apresentou a melhor precisão. A presença de efeitos de sombra nas imagens SAR aumentou os erros de comissão.

**Palavras-chave:** Detecção remota; recursos hídricos; classificadores de imagens; inundação.

**RÉSUMÉ** – DÉLIMITATION DES ZONES INONDÉES SUR LA BASE DES DONNÉES SAR SENTINEL-1 TRAITÉES PAR L'APPRENTISSAGE AUTOMATIQUE: UNE ÉTUDE DE CAS EN AMAZONIE CENTRALE, AU BRÉSIL. La délimitation de zones sujettes aux inondations est cruciale pour comprendre la dynamique hydrique et les changements fluviaux. Cette étude a analysé le potentiel des images radar à ouverture de synthétique (SAR) acquises en bande-C par le satellite Sentinel-1 en 2017, 2018 et 2019 pour délimiter les zones inondées en Amazonie Centrale. Les images ont été traitées par le Réseau de Neurones Artificiels Multicouches Perceptron (ARN-MLP) et deux classificateurs d'apprentissage automatique (ML) *K-Nearest Neighbor* (KNN-7 et KNN-11). Le prétraitement des images SAR images complexes à visée simple (SLC) a impliqué les étapes méthodologiques suivantes : application du fichier d'orbite; calibration radiométrique ( $\sigma^0$ ); correction de terrain *Range-Doppler*; filtrage du bruit de chatoiement; et conversion des données linéaires en coefficients de rétrodiffusion (unités en dB). Un filtre de Lee avec une taille de fenêtre de 3x3 a été appliqué pour filtrer le bruit de chatoiement. Un ensemble de 6000 échantillons assignés au hasard pour la formation (70%), la validation (20%) et les tests (10%) a été obtenu à partir de l'interprétation visuelle de l'image satellite optique Sentinel-2 acquise pour la même année que les images radar. Les plus grandes zones inondées ont été trouvées en 2019 dans la zone d'étude (municipalité de Parintins et Uruará, la Rivière Amazonas, Brésil) : 6244km<sup>2</sup> selon le classificateur RNA-MLP; 6268km<sup>2</sup> pour KNN-7; et 6290km<sup>2</sup> pour KNN-11, alors que les plus petites zones inondées ont été trouvées en 2018: 5364km<sup>2</sup> pour le

classificateur RNA-MLP; 5412km<sup>2</sup> pour KNN-7; et 5535km<sup>2</sup> pour KNN-11. Les trois classificateurs présentaient des coefficients *Kappa* compris entre 0,77 et 0,91. L'ARN-MLP a montré une meilleure précision. La présence d'effets d'ombre dans les images SAR a augmenté les erreurs de commission.

**Mots clés:** Télédétection; ressources en eau; classificateurs d'images; inondation.

**RESUMEN – DELIMITACIÓN DE ÁREAS INUNDADAS BASADA EN DATOS SAR SENTINEL-1 PROCESADOS MEDIANTE APRENDIZAJE AUTOMÁTICO: UN ESTUDIO DE CASO DE AMAZONÍA CENTRAL, BRASIL.** La delimitación de áreas sujetas a inundaciones es crucial para comprender la dinámica del agua y los cambios fluviales. Este estudio analizó el potencial de las imágenes de radar de apertura sintética (SAR) adquiridas en la banda-C por el satélite Sentinel-1 en 2017, 2018 y 2019 para delinear áreas inundadas en la Amazonía Central. Las imágenes fueron procesadas por la *Red Neuronal Artificial de Perceptrón Multicapa* (RNA-MLP) y dos clasificadores de aprendizaje automático (ML) *K-Nearest Neighbor* (KNN-7 y KNN-11). El preprocesamiento de imágenes SAR complejas de una sola mirada (SLC) involucró los siguientes pasos metodológicos: aplicación del archivo de órbita; calibración radiométrica ( $\sigma^0$ ); corrección del terreno *Range-Doppler*; filtrado de ruido moteado; y conversión de datos lineales en coeficientes de retrodispersión (unidades en dB). Se aplicó El filtro *Lee* con un tamaño de ventana de 3x3 para filtrar el ruido moteado. Se obtuvo un conjunto de 6000 muestras asignadas aleatoriamente para entrenamiento (70%), validación (20%) y prueba (10%) en base a la interpretación visual de la imagen del satélite óptico Sentinel-2 adquirida el mismo año que las imágenes de radar. Los humedales más grandes se encontraron en 2019 en el área de estudio (municipio de Parintins y Urucará, Río Amazonas, Brasil): 6244km<sup>2</sup> por el clasificador RNA-MLP; 6268km<sup>2</sup> por KNN-7; y, 6290km<sup>2</sup> por KNN-11, mientras que, los humedales más pequeños se encontraron en 2018: 5364km<sup>2</sup> por el clasificador RNA-MLP; 5412km<sup>2</sup> por KNN-7; y 5535km<sup>2</sup> por KNN-11. Los tres clasificadores presentaron coeficientes *Kappa* entre 0,77 y 0,91. RNA-MLP mostró la mejor precisión. La presencia de efectos de sombra en las imágenes SAR aumentó los errores de comisión.

**Palabras clave:** Sensores remotos; recursos hídricos; clasificadores de imágenes; inundación.

## I. INTRODUCTION

In the Brazilian Amazon region, rivers and their tributaries contain an extensive floodplain that corresponds to approximately 12% of the humid area of the Amazon basin. These floodplains present enormous terrestrial and aquatic biodiversity (Melack & Hess, 2010). Accurate flood monitoring not only in the Brazilian Amazon but also other regions of the world is important for increasing the security of local inhabitants and for reducing infrastructure damages and income losses. Besides, the frequency and magnitude of flood events are expected to increase due to climate change.

Flood monitoring can be conducted based on satellite observations, because of their ability to cover large areas, at high repetition and low costs. Inundation detection has

been addressed based on several optical satellites (e.g., Landsat, Sentinel-2, and Moderate Resolution Imaging Spectroradiometer onboard Terra and Aqua platforms) operating at different spatial, spectral, and temporal resolutions. They exploit the high level of absorption of radiation incident into the water bodies in the near-infrared and shortwave infrared spectra relative to the visible spectrum. However, the Amazon tropical region faces persistent cloud cover conditions most of the year, making the use of optical remote sensing data limited.

Synthetic Aperture Radar (SAR) remote sensing can be an important source of information for mapping flooded areas in the Brazilian Amazon because of its ability to acquire images under cloud-covered conditions. SAR sensors can identify inundation because of the typically lower backscattering returns from water bodies relative to other features. Basically, flooded areas in single SAR images are discriminated from non-flooded areas by thresholding backscatter values at different polarizations (Matgen *et al.*, 2011), subtracting backscattering coefficients between two images (Schlaffer *et al.*, 2015), or calculating variance in time series (DeVries *et al.*, 2020).

More recently, machine learning (ML) and deep learning (DL) classifiers are becoming quite popular in the field of remote sensing image classification. Although there is an overall agreement that DL is more powerful than ML, it requires bigger computational capabilities and so it may not be operational for studies involving large areas such as the Brazilian Amazon. ML-based image classification can be divided into supervised, unsupervised, and reinforcement learning categories. The two most used supervised ML classifiers are the *Random Forest* (RF) and *Support Vector Machine* (SVM) because they usually provide high accuracies in different types of land use and land cover classifications, including flooded and non-flooded classes (Banks *et al.*, 2019; Millard & Richardson, 2013; Mohammadimanesh *et al.*, 2018). The other popular supervised algorithms include naive Bayes and neural networks (Acharya *et al.*, 2019; Boateng *et al.*, 2020; Nemni *et al.*, 2020). Several authors have reported that divergences in the classification results can be substantial due to the differences in sensor systems, timing, and data processing algorithms (e.g., Aires *et al.*, 2013; Pham-Duc *et al.*, 2017; Rosenqvist *et al.*, 2020).

This study aims to evaluate the potential of the *Artificial Neural Network Multi-Layer Perceptron* (ANN-MLP) and two *k-Nearest Neighbor* (KNNs) algorithms to delineate flooded areas in a stretch of the Amazonas River in Central Amazon using Sentinel-1 SAR time series from 2017, 2018, and 2019. To our best knowledge, there is no study evaluating these ML classifiers to identify flooded areas in the Brazilian Amazon, especially using Sentinel-1 SAR data sets. Among the 29 studies listed recently by Fleischmann *et al.* (2022) involving inundation mapping by remote sensing over the Brazilian Amazon, nine relied on SAR data, all acquired by the ALOS/PALSAR mission. Currently, the only SAR data freely available on the internet are the ones acquired by the European Space Agency (ESA) Sentinel-1 satellite (Torres *et al.*, 2012). We addressed the following research question in this study: what are the performances of the ANN-MLP and KNN-based ML algorithms to map flooded areas in tropical rainforests based on Sentinel-1 satellite data?

## II. MATERIALS AND METHODS

### 1. Study area

The study area is located between the municipalities of Urucará and Parintins in the Amazonas State, Brazil, comprising part of the Amazonas River. It is located between the following coordinates:  $22^{\circ}30'48.84''$  and  $22^{\circ}37'59.16''$  of south latitude; and  $44^{\circ}31'35.68''$  and  $44^{\circ}43'25.94''$  of west longitude (fig. 1). The typical climate is classified as *Af*, that is, tropical rainforest climate without dry season, in the *Köppen* classification system, with average annual precipitation ranging from 1355mm to 2839mm (Alvares *et al.*, 2014). The average annual temperature varies from  $25.6^{\circ}\text{C}$  to  $27.6^{\circ}\text{C}$ . The flooding period occurs mostly between May and July.

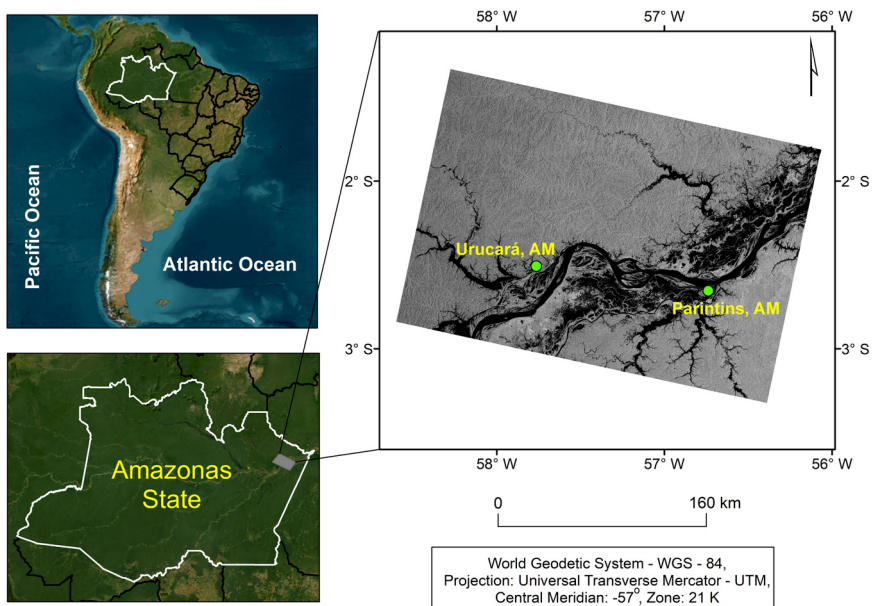


Fig. 1 – Location of the study area in the Central Amazon. The image at right acquired in 2019 by the Sentinel-1 SAR, at the VV polarization. Colour figure available online.

*Fig. 1 – Localização da área de estudo na Amazônia Central. A imagem à direita foi captada em 2019 pelo satélite Sentinel-1 SAR, na polarização VV. Figura a cores disponível online.*

### 2. Remote sensing data sets

This research used three Sentinel-1 SAR images acquired in the VV and VH polarizations during the following flooding periods: 23 June 2017; 18 June 2018; and 7 July 2019. The images were obtained in descending, *Interferometric Wide (IW)* mode, and processed at Level-1, which includes pre-processing and data calibration (table I).

Table I – Sentinel-1 SAR image acquisition modes.

*Quadro I – Modos de aquisição das imagens Sentinel-1 SAR.*

Mode*	Incident Angle (°)	Spatial Resolution	Swath Width (km)	Polarization
SM	20° – 45°	5m×5m	80	HH or VV or (HH and HV) or (VV and VH)
IW	29° – 46°	5m×20m	250	HH or VV or (HH and HV) or (VV and VH)
EW	19° – 47°	20m×40m	400	HH or VV or (HH and HV) or (VV and VH)
WV	22° – 35° 35° – 38°	5m×5m	20	HH or VV

\* SM = StripMap; IW = Interferometric Wide; EW = Extra Wide; and WV = Wave.

Source: ESA (2017)

We also selected Sentinel-2 *MultiSpectral Instrument* (MSI) images acquired near the Sentinel1 SAR overpasses (2 July 2017; 22 June 2018; and 29 July 2019). In this study, the Sentinel2 was used to collect sampling data for training, validation, and testing. The Sentinel-2 MSI images were radiometrically corrected. They have the potential for mapping flooded areas at regional scales, as they are acquired under the spatial resolutions of 10m to 60m and temporal resolution of 10-days (Du *et al.*, 2018) (table II).

Table II – Sentinel-2 MSI image acquisition modes.

*Quadro II – Modos de aquisição dos dados da imagem Sentinel-2 MSI.*

Spatial Resolution (m)	Band	Spectral Bands	Wavelength (nm)
10	B2	Blue	490
	B3	Green	560
	B4	Red	665
	B8	Near Infrared	842

Source: ESA (2017)

### 3. Methodological approach

Figure 2 shows the main steps of the methodological approach used in this study. We conducted the following pre-processing steps: correction by the orbit file; terrain correction; radiometric calibration; conversion of the data to decibels; and spatial filtering. The images were pre-processed by the image orbit file, containing accurate information on the satellite's position, trajectory and speed during the image capture process (ESA, 2017). The terrain correction was based on the digital elevation model (DEM) acquired by the Shuttle Radar Topographic Mission (SRTM) at  $\sim 3$  arc sec<sup>-1</sup>. The radiometric calibration was performed using the Sigma Look-Up Table (LUT) file to generate images converted into backscattering coefficients ( $\sigma^0$ ).

SAR images present speckles that originate from destructive or additive interference from the radar return signal for each pixel (Lee & Pottier, 2009). We used *Lee* filter with a 3×3 window size for processing VV- and VH-polarized images. The *Lee* filter trans-



forms the multiplicative model into an additive model by expanding the first-order *Taylor* series around the average. This technique uses local statistics to minimize the mean square error (MSE) through the *Wiener* filter. In this way, the *Lee* filter is an adaptive filter that has the characteristics of preserving edges (Sant’Anna, 1995). *Lee*’s filter assumes that the mean and variance of the pixel of interest are equal to the mean and variance of all local pixels, which refer to the inside of the adopted window.

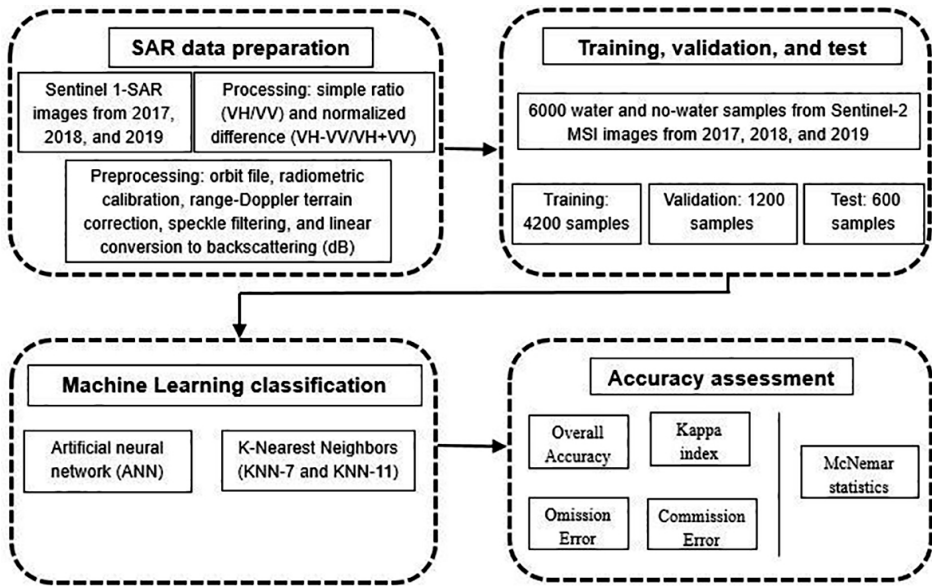


Fig. 2 – Methodological flowchart with the main steps for classifying flooding areas in the Central Amazon in 2017, 2018, and 2019.

Fig. 2 – Fluxograma metodológico com as principais etapas para classificação de áreas inundáveis na Amazônia Central em 2017, 2018 e 2019.

We used the following image processing software: *S1-Toolbox* available in the *Sentinel Application Platform* (SNAP) version 7.0.0; *ArcGIS* version 10.5; and *Abilius*, which uses the *OpenCV* library of artificial intelligence algorithms in the C++ programming language.

Sentinel-1 images converted into backscattering coefficients were normalized based on their averages and standard deviations (eq. 1 and 2).

$$\text{Normalized VH} = \frac{VV - VV_{\text{mean}}}{VV_{\text{std}}} \tag{Eq. 1}$$

$$\text{Normalized VV} = \frac{VH - VH_{\text{mean}}}{VH_{\text{std}}} \tag{Eq. 2}$$

The normalized images were also processed to generate the simple ratio (SR) index and normalized difference (ND) index involving VH and VV polarizations (Hird *et al.*, 2017; Tsyganskaya *et al.*, 2018) (eq. 3 and 4).

$$SR = \frac{VH}{VV} \quad (\text{Eq. 3})$$

$$ND = \frac{VH - VV}{VH + VV} \quad (\text{Eq. 4})$$

As pointed out by Boateng *et al.* (2020), the most widely used nonparametric ML techniques include ensembles of classification trees such as *Random Forest* (RF) (Breiman, 2001), *Artificial Neural Networks* (ANNs) (Brown *et al.*, 2000), *K-Nearest-Neighbors* (KNN) (Breiman & Ihaka, 1984), *Support Vector Machines* (SVMs) (Cortes & Vapnik, 1995). In this study, we selected two classifiers, the ANN and two KNN (KNN-7 and KNN-11) algorithms to verify the performance of these techniques to classify flooded areas in the region of interest. The widely used RF classifier was not selected because we are interested in only two classes (water and non-water), making it impossible to develop a random forest, ensemble architecture that is the basis of these algorithms. In this ensemble architecture, several classification trees are trained based on subsets of the training data (Abdi, 2020). We did not evaluate SVM either because, together with RF, it has been intensively assessed in literature over several different environmental and terrain conditions.

The ANN adopted in this study was the *Multilayer Perceptron* (MLP) type with the backpropagation learning algorithm with insertion of the momentum term, which optimizes the network processing with a learning rate of 0.05 and a momentum factor of 0.5. According to Atkinson and Tatnall (1997), there are several advantages of neural networks, such as the efficient manipulation of large data sets and their use in the classification of remote sensing data without assuming a normal distribution. In the classification by neural networks, we used the *Abilius* program developed by the University of Brasília, Brazil, which is based on the OpenCV library. We used the logistical activation function in which the output result of the neuron, given a set of input signals, assumes real values between zero and one to facilitate the network training process and to simplify its structure (eq. 5):

$$g(\mu) = \frac{1}{1 \pm e^{-\beta\mu}} \quad (\text{Eq. 5})$$

where  $\beta$  = real constant associated with a slope level of the logistic function related to its inflection point; and  $\mu$  = activation potential produced by the difference in value produced between the linear combination and activation threshold.

The KNN classifier is a non-parametric method based on k-training samples closest to the behavior of the analyzed data (Cover & Hart, 1967). The calculation of the nearest



neighbor was performed using the *Euclidean* distance method. In this classification technique, the  $k$ -value refers to the number of neighbors to be used in determining the class assigned by the values of most of the nearest pixels that must be assigned. There are several studies in the literature with different values of  $k$  (Alves *et al.*, 2013). The unknown sample is assigned to the most common class of the  $k$ -training samples that are nearest in the feature space to the unknown sample (Maxwell *et al.*, 2018). In this study, we assigned  $k$ -values of seven and eleven. These numbers are a compromise between too-low and too-high  $k$ -values. Low  $k$ -values will produce complex decision boundaries while high  $k$ -values will result in greater generalization (Maxwell *et al.*, 2018).

The training, validation, and testing data set was produced by manual sample collection of 6000 pixels (denoized time signatures) containing two classes (water and non-water), with equal distribution (3000 samples per class), showing well-distributed sampling design (fig. 3). We considered a total of 4200-pixel samples for training (70%), 1200-pixel samples for validation (20%) and 600-pixel samples for testing (10%), according to the methodology defined by Kuhn and Johnson (2013) and Larose and Larose (2014). The training of ANNs considered different architectures for VV, VH, SR, and ND images. The number of neurons in the hidden layer was determined by the trial and error methods (Hirose *et al.*, 1991). The selected stopping criterion was the number of learning cycles, defined as 10 000. At the end of the training process, 180 sets of independent samples were collected to validate the classification results. The selection of the best classifier was based on the lowest values of mean squared error (MSE).

The accuracy of the classification was analyzed using the confusion matrix, omission and commission errors, overall accuracy, and *Kappa* index (Congalton & Green, 1993). Overall accuracy (OA) and the *Kappa* index were calculated using equations 6 and 7:

$$OA = \sum_{i=1}^m \frac{n_{ii}}{n} \quad (\text{Eq. 6})$$

where  $n_{ii}$  = diagonal elements of the confusion matrix;  $n$  = total number of observations; and  $m$  = number of themes mapped.

$$Kappa = \frac{\sum_{i=1}^M x_i - \sum_{i=1}^M (x_i + x+i)}{n^2 - \sum_{i=1}^M (x_i + x+i)} \quad (\text{Eq. 7})$$

where  $n$  = total number of observations; and  $x_i$  and  $x+i$  are the sums in row and column.

*Kappa* is a coefficient that varies from zero to one, representing a general agreement index. *Kappa* values are associated with the quality of the classification. Cell values were considered for measuring omission and commission errors. The marginal cells in the lines indicate the number of pixels that were not included in a particular category, that is, express the error known by default. Cells on the diagonals represent the pixels that were

not included in no category, expressing the error of commission (Congalton & Green, 1993). The omission error (OE<sub>i</sub>) and the commission error (CE<sub>j</sub>) were calculated for the thematic classes of the classification (eq. 8 and 9):

$$OE_i = \frac{\sum X_{ij} - X_{ii}}{\sum X_{ij}} \tag{Eq. 8}$$

$$CE_j = \frac{\sum X_{ij} - X_{jj}}{\sum X_{ij}} \tag{Eq. 9}$$

where  $\sum X_{ij} - X_{ii}$  = sum of waste per line;  $\sum X_{ij} - X_{jj}$  = sum of waste per column; and  $\sum X_{jj}$  = row or column marginal.

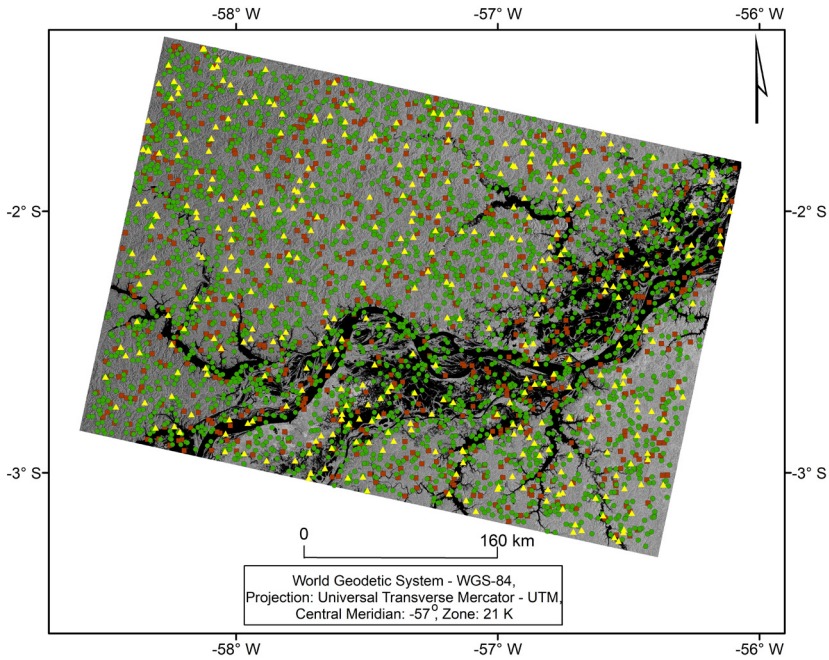


Fig. 3 – Location of the samples for training (green circle), validation (brown square), and test (yellow triangle). Colour figure available online.

*Fig. 3 – Localização das amostras para treino (verde), validação (marron) e teste (amarelo). Figura a cores disponível online.*

We also conducted another validation strategy based on the Sentinel-2 MSI scenes from June and July of 2019 and the *McNemar* chi-squared test ( $c^2$ ). The accuracy analysis used 88 systematic samplings with a 10km diameter in regular grids of 17×17km<sup>2</sup> (fig. 4). We disregarded the cloud-covered samples in the Sentinel-2 images.

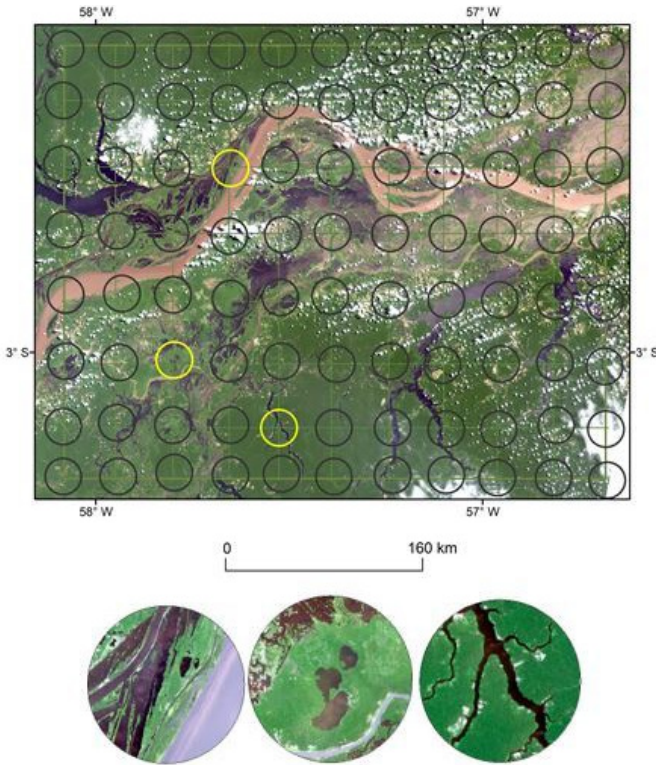


Fig. 4 – Location of systematic samples for validation of flooding maps produced by the machine learning classifiers. Colour figure available online.

*Fig. 4 – Localização das amostras sistemáticas para validação dos mapas de inundação produzidos pelos classificadores de aprendizagem de máquina. Figura a cores disponível online.*

McNemar's  $\chi^2$  test was used considering a statistical level of significance of 0.05 and one degree of freedom to analyze the differences in measured areas between visual interpretation and classified images. According to McNemar (1947) and Leeuw *et al.* (2006), McNemar's analysis is a non-parametric statistical test to analyze pairs and has been widely used in remote detection because it can use the same validation set (Eq. 10):

$$\chi^2 = \frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}} \quad (\text{Eq. 10})$$

where  $f_{12}$  = number of wrong classifications by Method1, but correctly classified by Method2; and  $f_{21}$  = number of correct classifications correct by Method1, but incorrectly classified by Method2.

This precision comparison based on related samples is quite popular in the literature (Abdi *et al.*, 2020; Manandhar *et al.*, 2009; Mayer *et al.*, 2021; Wang *et al.*, 2018). The

*McNemar* statistical test was performed at the level of significance of 0.05 and one degree of freedom between the image classifiers to analyze whether the classified images differ statistically.

### III. RESULTS

#### 1. Backscattering values of flooded areas

Table III presents the statistical results for the backscatter coefficients of the three-year images in dual polarization from the Sentinel-1 SAR satellite before and after the data normalization. The results of the normalization showed a slight increase in the mean values compared to the non-normalized images. The value of variance and standard deviation remained identical with the image before adjustment. Decreasing mean values in the fitted images represents a greater concentration of the distribution data, compressing the backscatter values for the image. In other words, the decrease in the mean and standard deviation values indicates a low dispersion of the backscattered data. The characteristics and multi-temporal patterns of the backscatter values were similar to the VH and VV polarizations, as well as to the SR and ND images for the three years.

Table III – Statistical results before and after normalization of Sentinel-1 SAR images from 2017 to 2019.  
*Quadro III – Resultados estatísticos antes e após a normalização das imagens Sentinel-1 SAR de 2017 a 2019.*

Overpass	Normalization	Mean VH	Mean VV	Standard Deviation VH	Standard Deviation VV	Variance VH	Variance VV
23 June 2017	Non-normalized	-10.68	-7.08	8.77	7.14	76.91	50.97
	Normalized	-9.46	-6.09	8.77	7.14	76.91	50.97
12 July 2018	Non-normalized	-9.96	-6.29	7.69	5.76	59.13	33.17
	Normalized	-8.67	-5.20	7.69	5.76	59.13	33.17
17 June 2019	Non-normalized	-10.50	-6.63	7.86	5.93	61.77	35.16
	Normalized	-9.16	-5.52	7.86	5.93	61.77	35.16

Figure 5 shows the box plot in a range of grouped multi-temporal backscatter values, measured in all scenes (VH, VV, SR, and ND). In comparison to the backscatter values of water bodies in the time series, there was a considerable increase in the average backscatter values in the following order: VH, VV, SR, and ND. The backscatter of water bodies in the VH image was the lowest among the three scenes, with an average value of -25.8dB in the upper limit and -20.0dB in the lower limit and with backscattering in the value of -24.2dB in the first quartile. In contrast, the VV polarization presented the largest data series between the lower and upper limits, with the upper limit lower value at -23.3dB and the upper limit at -13.3dB, with 75% of its backscatter values being represented by -15.8dB, as shown in quartile-3.

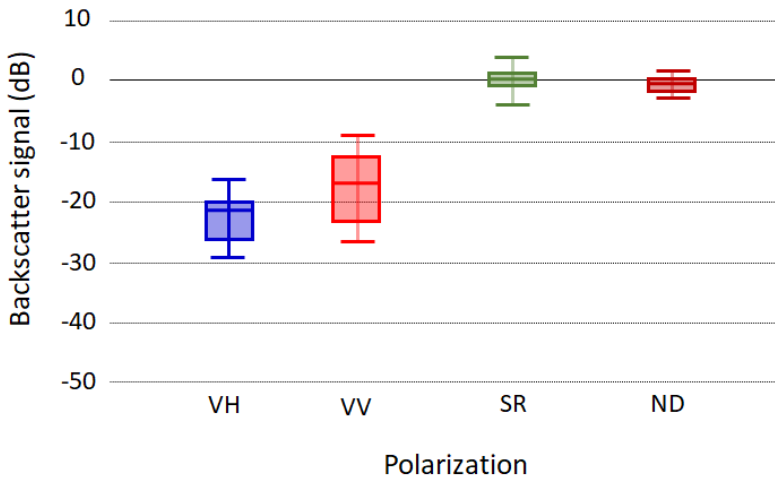


Fig. 5 – VH, VV, SR and ND backscatter values over water bodies obtained by averaging the Sentinel-1 scenes from 2017, 2018, and 2019. Colour figure available online.

*Fig. 5 – Valores médios de coeficientes de retroespalhamento para corpos de água nas polarizações VH, VV, SR e ND relativos aos compósitos adquiridos em 2017, 2018 e 2019. Figura a cores disponível online.*

The VH polarization showed the lowest backscatter values for the water body due to the record of the backscattered value of the return signal on the antenna occurring in the horizontal direction. The backscatter values of the indices became more aggregated compared to the scalar values of the polarizations VH and VV, with median, quartiles, and closest limits for the index images.

## 2. Best ANN model for flooded area detection

The ANN model with the combination of VV and VH polarizations as well as SR and ND presented accurate results in the network learning and training tests, with: 91.3% and 91.9% of overall accuracies for the image acquired in 2017; 90.7% and 90.9% accuracies for 2018; and 90.2% and 90.7% accuracies for 2019. The ANN architecture with the best results was obtained using four neurons in the input layer, two hidden inner layers with eight neurons, and two neurons in the output layer (4-8-8-2 model) (fig. 6).

In the learning phase of the three ANNs, the training errors were slightly higher than the test errors. However, with the increasing number of trainings in some moments, these errors were equalized. The errors stabilized with values of root mean square error (RMSE) in 0.2. The maximum cycles did not exceed 10 000 iterations. These values demonstrate that the maximum learning limit of ANN with 10 000 iterations is sufficient for training, which results in high precision, accuracy, and lower computational processing cost (fig. 7).

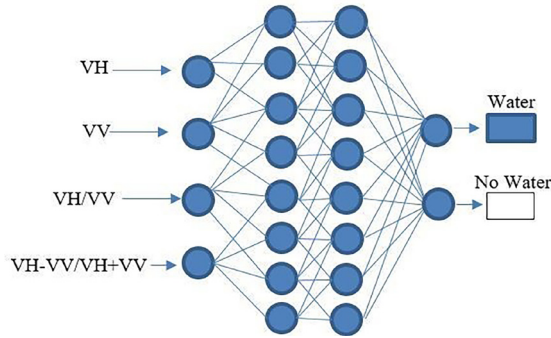


Fig. 6 – The most appropriate artificial neural network model for classifying water bodies in the study area (4-8-8-2). Colour figure available online.

Fig. 6 – Arquitetura da rede neural artificial mais apropriada para classificar corpos d'água na área de estudo (4-8-8-2). Figura a cores disponível online.

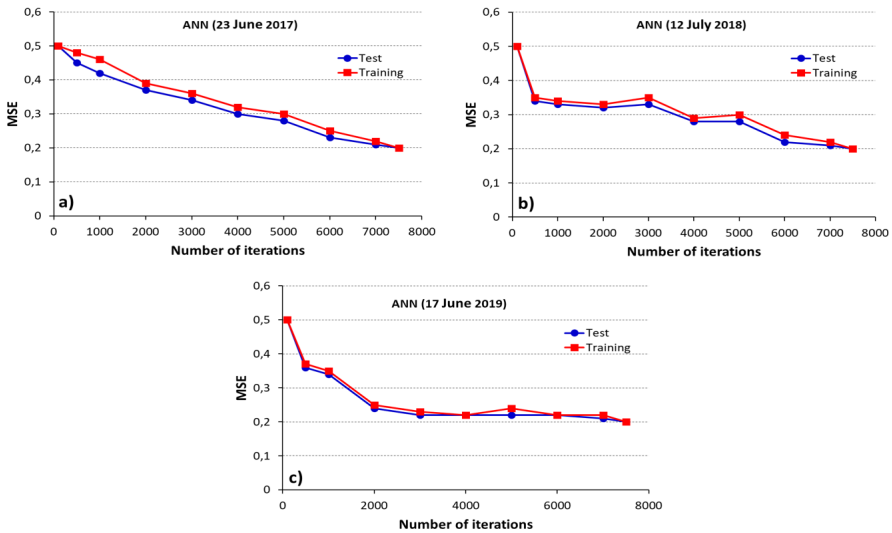


Fig. 7 – Mean square error (MSE) values in relation to on the number of iterations in the training of Artificial Neural Networks (ANN) for scenes acquired in 2017 (a), 2018 (b), and 2019 (c). Colour figure available online.

Fig. 7 – Valores do erro quadrático médio (MSE) em relação ao número de iterações no treino de Redes Neurais Artificiais (ANN) para os compósitos adquiridos em 2017 (a), 2018 (b) e 2019 (c). Figura a cores disponível online.

### 3. Classification results

Figure 8 shows the classification of the flooded areas in the Central Amazon region using the ANN, KNN-7, and KNN-11 classifiers in the period of largest flood pulse during the years 2017, 2018, and 2019. The blue-colored areas correspond to areas classi-



fied as water bodies. The classifiers delimited precisely the main channel of the Amazon River, as well as the flooded areas adjacent to the river channel. In general, the classification through ANN generated products classified with slightly higher clarity when compared to the KNN classifiers.

The noises classified as upper lands increased from KNN-7 to KNN-11, indicating that the increase in the *Euclidean* distance from seven to eleven contributes to the increase of confusion between the water body and the upper land. In addition, there was a misclassification of water bodies for all images. These areas are shown as random blue dots mainly after the boundary of the tributaries of the Amazon River, as well as the presence of random pixels scattered in various regions spread in the entire study area.

The largest presence of water bodies was found in the image acquired in 2019, with a total area of 6244km<sup>2</sup> (ANN), 6268km<sup>2</sup> (KNN-7), and 6290km<sup>2</sup> (KNN-11). In other words, the KNN-7 and KNN11 algorithms presented the largest occurrences of water bodies, as compared with the ANN classification.

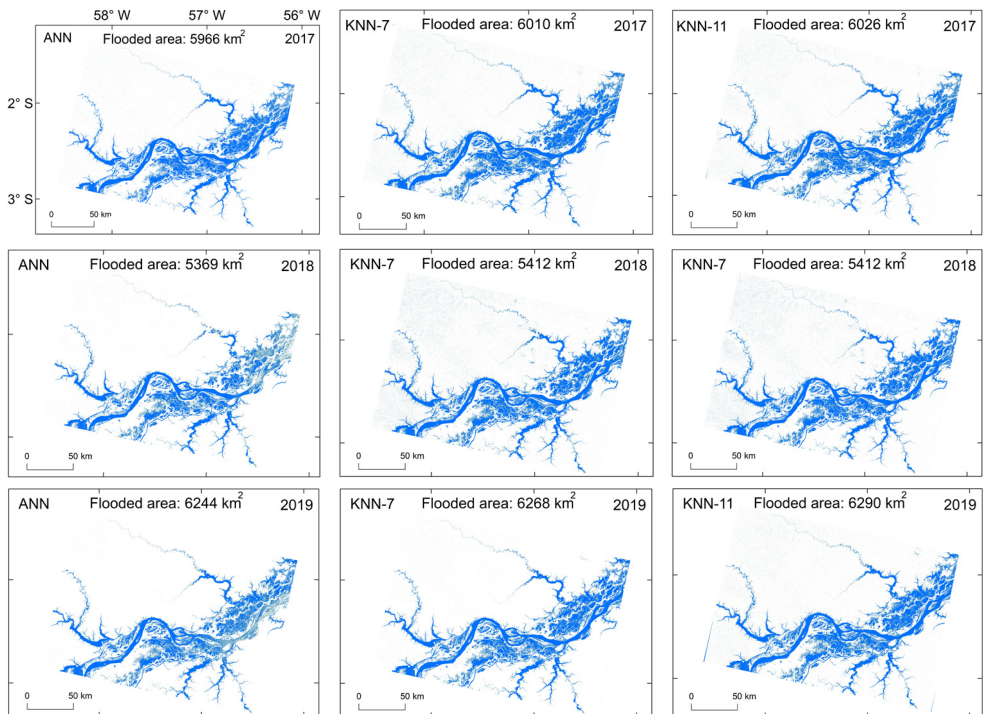


Fig. 8 – Delineation of flooded areas in the Central Amazon region through the ANN, KNN-7, and KNN-11 classifiers applied in the *Sentinel-1* images acquired in 2017, 2018, and 2019.

Colour figure available online.

Fig. 8 – Delimitação de áreas inundadas na região da Amazônia Central através da aplicação dos algoritmos ANN, KNN-7 e KNN-11 nos compostos *Sentinel-1* adquiridos em 2017, 2018 e 2019.

Figura a cores disponível online.

#### 4. Accuracy analysis

The image classification by ML showed *Kappa* coefficient values from 0.77 to 0.91 (table IV). ANN showed higher accuracy for all Sentinel-1 scenes. Second in the classification order, the KNN7 presented results close to the ANN. The KNN-11, on the other hand, presented the highest differences in the accuracies among the three classifiers, obtaining the lowest *Kappa* index (0.77) in the image acquired in 2018.

There was no discrepancy between *Kappa* indexes and the overall accuracies. The classifications by ML showed high overall accuracy values, with ANN presenting the highest values, with 97% in the image from 2019. The lowest value was presented by KNN-11, with 92%, in the image acquired in 2018.

Table IV – *Kappa* coefficient, overall accuracy, and omission and commission errors in the SAR images processed by the ANN, KNN-7, and KNN-11 classifiers.

*Quadro IV – Coeficiente Kappa, precisão geral, erros de omissão e comissão nas imagens SAR obtidas pelos classificadores ANN, KNN-7 e KNN-11.*

Satellite Overpass	<i>Kappa</i>	Overall Accuracy (%)	Commission Error (%)	Omission Error (%)
ANN				
23 June 2017	0.87	96	8.99	2.99
12 July 2018	0.85	93	10.90	3.98
17 June 2019	<b>0.91</b>	<b>97</b>	<b>6.99</b>	<b>1.99</b>
KNN-7				
23 June 2017	0.85	95	10.9	4.9
12 July 2018	0.82	95	12.9	5.9
17 June 2019	0.88	96	8.99	3.9
KNN-11				
23 June 2017	0.83	94	17.9	8.9
12 July 2018	<b>0.77</b>	<b>92</b>	<b>18.9</b>	<b>12.9</b>
17 June 2019	0.85	95	11.5	5.7

The ANN classification technique obtained the lowest commission error, with 7.0% and omission error of 1.9%, in the image from 2019. The largest commission and omission errors were measured in the image from 2018, classified by KNN-11, with a commission error of 18.9% and an omission of 12.9%. There was more commission error when compared to the omission error in all products generated by the classifiers, proving that the biggest classifier errors occurred in the definition of the drylands as the water body.

The image from 2018 was measured with the lowest presence of water bodies, which was also the image that presented the largest errors and the worst statistical indices analyzed. Thus, it is inferred that due to the smaller grouping and the greater distance between the pixels corresponding to the backscatter values of the water bodies was the determining factor for obtaining the worst results obtained by the KNN classifier with *Euclidean* distance of eleven.

The classifiers produced an overall good performance in delineating water bodies, especially the ANN and the KNN-7. They showed the lowest levels of noise in the images (fig. 9). The results of the indexes showed that the ANN obtains the best performance in comparison with the other classifiers. ANN and KNN-7 achieved similar levels of precision in two of the three images, these images being classified with greater quantities of water bodies. On the other hand, it was observed that the worst indices occurred in the KNN-11 classification and were obtained in the image with the least amount of water bodies.

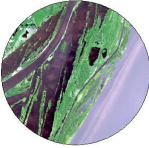
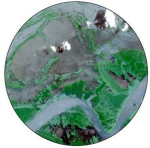
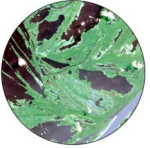












Sentinel-2 R (4), G (3) and B (2) 	Sentinel-2 R (4), G (3) and B (2) 	Sentinel-2 R (4), G (3) and B (2) 
Visual analysis: flooded area = 59.6% 	Visual analysis: flooded area = 55.5% 	Visual analysis: flooded area = 36.1% 
ANN-MLP: flooded area = 57.7% 	ANN-MLP: flooded area = 54.9% 	ANN-MLP: flooded area = 33.9% 
KNN-7: flooded area = 56.1% 	KNN-7: flooded area = 51.2% 	KNN-7: flooded area = 32.7% 
KNN-11: flooded area = 53.8% 	KNN-11: flooded area = 48.8% 	KNN-11: flooded area = 30.2% 

Fig. 9 – Results of visual interpretation (Sentinel-2) and Machine Learning (ML) classification (Sentinel-1) of three enlarged images acquired in 2019. Colour figure available online.

*Fig. 9 – Resultados da interpretação visual (Sentinel-2) e classificação pelo método de Machine Learning (ML) (Sentinel-1) de três imagens ampliadas que foram adquiridas em 2019. Figura a cores disponível online.*

Table V shows the results of the *McNemar* test between the pairs of the three classifications, considering the different combinations of parameters. The ANN showed the best results in the time series, despite obtaining the best values between the methods. The ANN did not differ statistically from the KNN-7 classifier in the time-period. The KNN-11 classifier presented the highest value of  $\chi^2$  with 4.37, in the image from 2018. This result was statistically significant, rejecting the hypothesis of statistical equality between the pairs of classifiers RNA  $\times$  KNN-11 at the p-value of 0.05, since the calculated  $\chi^2$  is higher than the tabulated  $\chi^2$  (3.84).

In general, the biggest confusion for image classifiers was to distinguish the water class from other non-water targets with similar backscatter values. In this way, the presence of the shadow effects was observed in all images, indicating that this effect was the biggest cause of misclassification, with less interference only in ANN classification (fig. 10).

Table V – *McNemar* test between visual interpretation and ANN, KNN-7, and KNN-11 classifiers shown in terms of  $\chi^2$  values.

*Quadro V – Teste de McNemar entre interpretação visual e classificadores ANN, KNN-7 e KNN-11, apresentados em termos de valores de  $\chi^2$ .*

Sentinel-1 Overpass	Visual $\times$ ANN	Visual $\times$ KNN-7	Visual $\times$ KNN-11	ANN $\times$ KNN-7	ANN $\times$ KNN-11	KNN-7 $\times$ KNN-11
23 June 2017	2.10	2.36	3.63	2.64	2.94	2.86
12 July 2018	2.30	3.61	<b>4.28<sup>(*)</sup></b>	2.98	<b>4.37<sup>(*)</sup></b>	2.77
17 June 2019	2.68	3.55	<b>4.46<sup>(*)</sup></b>	2.12	2.53	2.36

(\*) represents statistical significance, with the calculated value of  $\chi^2$  greater than the tabulated value at a significance level of 5% (3.84).

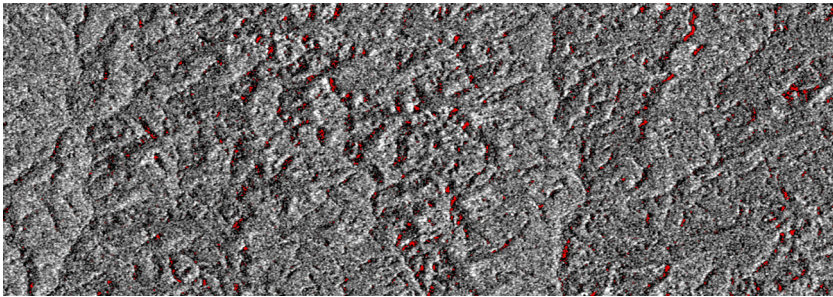


Fig. 10 – Shadow effects in the ANN classification over an enlarged portion of Sentinel-1 image acquired in 2019. Colour figure available online.

*Fig. 10 – Efeitos de sombreamento de relevo na classificação de imagens SAR pelo algoritmo ANN numa porção da imagem adquirida pelo satélite Sentinel-1, em 2019. Figura a cores disponível online.*

#### IV. DISCUSSION

The C-band backscattering coefficients of flooded areas in 2017, 2018, and 2019 varied from 10.68dB to 9.96dB in the VH polarization and from -7.08dB to -6.29dB in the VV polarization in the study area. These values are quite higher than those found by

Magalhães *et al.* (2022) for open water bodies in the Amazon River: -19dB in the VH polarization and -14dB in the VV polarization. Conde and Muñoz (2019) reported that backscattering intensity values of permanent water bodies are below -20dB. Moharrami *et al.* (2021) applied a threshold value of -14.9dB to the Sentinel-1 scenes to delineate flooded areas. Our higher values are probably due to the contribution of sparse shrubs and trees that we find in flooding areas in the surface backscattering process.

The accuracy assessment based on *Kappa* index, omission and commission errors showed overall accuracies of detecting flooded areas close to one another and higher than 90% for all three classifiers. These values are comparable to or higher than the accuracies obtained by other scientists who used Sentinel-1 SAR data for classifying flooded areas around the world. For example, Twele *et al.* (2016) used a processing chain approach to detect flood conditions in two test sites at the border between Greece and Turkey. They showed encouraging overall accuracies between 94.0% and 96.1%. Liang and Liu (2020), in their water and non-water delineation study using four different thresholding methods, reached overall accuracies ranging from 97.9% to 98.9% for a study area located in Louisiana State, USA. Siddique *et al.* (2022) evaluated RF and KNN algorithms applied to Sentinel-1 images from North India and concluded that C-band SAR data can detect changes in flood patterns over different land cover types with overall accuracies ranging from 80.8% to 89.8%.

The comparison between visual analysis and the selected ML classifiers based on three enlarged images (fig. 9) showed an overall underestimation of flooded areas for the ML classifiers. However, the *NcNemar*  $\chi^2$  test showed that the results from visual interpretation and ML classifiers were statistically equal from each other. The only exceptions were found for the KNN-11 applied in the scenes acquired in 2018 and 2019. The *McNemar* test also showed that ANN  $\times$  KNN-7 and ANN  $\times$  KNN11 did not differ statistically each other. These results indicate the existence of site-specific spatial heterogeneity within the study area. In other words, the overall statistical results found for the entire study area may differ depending on the local landscape conditions within the study area.

Figure 10 showed the presence of shadowing effects in the flood delineation, leading to higher commission errors. This effect has been reported widely in the literature (e.g., Chen & Zhao, 2022), even sometimes its presence in SAR images is used as an indicator of some targets, mainly deforestation (Bouvet *et al.*, 2018). Shadowing effects occur in SAR images because of their mandatory side-looking geometry. In other words, shadows in SAR images are related to areas in the terrain that cannot be reached by emitted radar pulses. As Sentinel-1 scenes are acquired in an almost north-south orbit (98.2° of inclination), most of the pixels classified as flooded in the upper lands are also oriented approximately north-south.

## V. CONCLUSION

The Sentinel-1 SAR images classified by ML algorithms showed good potential to map flooded areas in the Central Amazon. The three tested algorithms produced accuracies ranging from 92% to 97%. ANN and KNN-7 classifiers showed better potential than the KNN-11. Shadow effects appearing in non-flooded areas surrounding the flooded areas increased the commission errors.



The methodological approach used in this study may be suitable to map flooding areas in other regions of the Brazilian Amazon, but no broader generalizations can be made as the performance of the methods varies according to the local environmental and biophysical conditions.

As SAR images are quite sensitive to texture, the addition of textural attributes derived, for example, from the Gray Level Co-occurrence Matrix (GLCM), such as the angular second moment, dissimilarity, entropy, and variance in the classification procedure may improve the classification results. More recent studies have demonstrated the high performance of DL algorithms so they also should be tested to map flooded areas. U-Net and tensor flow are the ones that are becoming quite popular in the DL group of classifiers.

## ACKNOWLEDGEMENTS

The authors are grateful for financial support from CNPq. Special thanks to the research groups of the Laboratory of Spatial Information System at the University of Brasilia. Finally, the authors thank the anonymous reviewers who improved this research.

## ORCID ID

Ivo Augusto Lopes Magalhães  <https://orcid.org/0000-0003-4136-1972>

Osmar Abilio de Carvalho Junior  <https://orcid.org/0000-0002-0346-1684>

Edson Eyji Sano  <https://orcid.org/0000-0001-5760-556X>

## AUTHORS' CONTRIBUTIONS

**Ivo Augusto Lopes Magalhães:** Conceptualization; Methodology; Software; Validation; Formal analysis; Research; Writing. **Osmar Abílio de Carvalho Júnior:** Conceptualization; Methodology; Software; Resources; revision and editing; Visualization; Supervision; Fund acquisition. **Edson Eyji Sano:** Methodology; Validation; Data curation; Writing – revision and editing; Visualization.

## REFERENCES

- Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, 57(1), 1-20. <https://doi.org/10.1080/15481603.2019.1650447>
- Acharya, T. D., Subedi, A., & Lee, D. H. (2019). Evaluation of Machine Learning Algorithms for Surface Water Extraction in a Landsat 8 Scene of Nepal. *Sensors*, 19(12), 2758-2769. <https://doi.org/10.3390/s19122769>
- Aires, F., Papa, F., & Prigent, C. (2013). A long-term, high-resolution wetland dataset over the Amazon Basin, downscaled from a multiwavelength retrieval using SAR data. *Journal of Hydrometeorology*, 14(2), 594-607. <https://doi.org/10.1175/JHM-D-12-093.1>



- Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., & Sparovek, G. (2014). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6), 711-728. [10.1127/0941-2948/2013/0507](https://doi.org/10.1127/0941-2948/2013/0507)
- Alves, M. V. G., Chiavetta, U., Koehler, H. S., Machado, S. A., & Kirchner, F. F. (2013). Aplicação de k-nearest neighbor em imagens multispectrais para a estimativa de parâmetros florestais [Application of k-nearest neighbor on multispectral images to estimate forest parameters]. *Floresta*, 43(3), 351-362. <http://dx.doi.org/10.5380/rf.v43i3.18083>
- Atkinson, P. M., & Tatnall, A. R. L. (1997). Introduction Neural networks in remote sensing. *International Journal of Remote Sensing*, 18(4), 699-709. <https://doi.org/10.1080/014311697218700>
- Banks, S., White, L., Behnamian, A., Chen, Z., Montpetit, B., Brisco, B., ... & Duffe, J. (2019). Wetland Classification with Multi-Angle/Temporal SAR Using Random Forests. *Remote Sensing*, 11(6), 670. <https://doi.org/10.3390/rs11060670>
- Boateng, E. Y., Otoo, J., & Abaye, D. A. (2020). Basic tenets of classification algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A review. *Journal of Data Analysis and Information Processing*, 8(4), 341-357. <https://doi.org/10.4236/jdaip.2020.84020>
- Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., & Le Toan, T. (2018). Use of the SAR Shadowing Effect for Deforestation Detection with Sentinel-1 Time Series. *Remote Sensing*, 10(8), 1250. <https://doi.org/10.3390/rs10081250>
- Breiman, L. (2001). Random Forests. *Machine Learning*, (45), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., & Ihaka, R. (1984). *Nonlinear discriminant analysis via scaling and ACE. Technical Report*. Department of Statistics, University of California. <https://digitalassets.lib.berkeley.edu/sdtr/ucb/text/40.pdf>
- Brown, W. M., Gedeon, T. D., Groves, D. I., & Barnes, R. G. (2000). Artificial neural networks: A new method for mineral prospectivity mapping. *Australian Journal of Earth Sciences*, 47(4), 757-770. <https://doi.org/10.1046/j.1440-0952.2000.00807.x>
- Chen, Z., & Zhao, S. (2022). Automatic monitoring of surface water dynamics using Sentinel-1 and Sentinel-2 data with Google Earth Engine. *International Journal of Applied Earth Observations and Geoinformation*, 113, 103010. <https://doi.org/10.1016/j.jag.2022.103010>
- Conde, F. C., & Muñoz, M. M. (2019). Flood Monitoring Based on the Dstudy of Sentinel-1 SAR Images: The Ebro River Case Study. *Water*, 11(12), 2454. <https://doi.org/10.3390/w1122454>
- Congalton, R. G., & Green, K. A. (1993). A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering & Remote Sensing*, 59(5), 641-644. [https://www.asprs.org/wp-content/uploads/pers/1993journal/may/1993\\_may\\_641-644.pdf](https://www.asprs.org/wp-content/uploads/pers/1993journal/may/1993_may_641-644.pdf)
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, (20), 273-297. <https://doi.org/10.1007/BF00994018>
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27. <https://doi.org/10.1109/TIT.1967.1053964>
- DeVries, B., Huang, C., Armstron, J., Huang, W., Jones, J. W., & Lang, M. W. (2020). Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. *Remote Sensing of Environment*, 240, 111664. <https://doi.org/10.1016/j.rse.2020.111664>
- Du, Z., Ge, L., Ng, A. H.-M., Zhu, Q., Yang, X., & Li, L. (2018). Correlating the subsidence pattern and land use in Bandung, Indonesia with both Sentinel-1/2 and ALOS-2 satellite images. *International Journal of Applied Earth Observation and Geoinformation*, 67, 54-68. <https://doi.org/10.1016/j.jag.2018.01.001>
- European Space Agency. (2017). *Sentinel-1 SAR Technical Guide*. ESA. <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar>
- Fleischmann, A. S., Papa, F., Fassoni-Andrade, A., Melack, J. M., Wongchuig, S., Paiva, R. C. D., ... Collischonn, W. (2022). How much inundation occurs in the Amazon River basin? *Remote Sensing of Environment*, 278, 113099. <https://doi.org/10.1016/j.rse.2022.113099>
- Hird, J. N., DeLancey, E. R., McDerimid, G. J., & Kariyeva, J. (2017). Google Earth Engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping. *Remote Sensing*, 9(12), 1315. <https://doi.org/10.3390/rs9121315>
- Hirose, Y., Yamashita, K., & Hijjiya, S. (1991). Back-propagation algorithm which varies the number of hidden units. *Neural Networks*, 4(1), 61-66. [https://doi.org/10.1016/0893-6080\(91\)90032-Z](https://doi.org/10.1016/0893-6080(91)90032-Z)

- Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer. <https://doi.org/10.1007/978-1-4614-6849-3>
- Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons. <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118874059>
- Lee, J. S., & Pottier, E. (2009). *Polarimetric Radar Imaging: From Basic to Applications*. CRC Press. <https://doi.org/10.1201/9781420054989>
- Leeuw, J. D., Jia, H., Yang, L., Schmidt, K., & Skidmore, A. K. (2006). Comparing accuracy assessment to infer superiority of image classification methods. *International Journal of Remote Sensing*, 27(1), 223-232. <https://doi.org/10.1080/01431160500275762>
- Liang, J., & Liu, D. (2020). A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159, 53-62. <https://doi.org/10.1016/j.isprsjprs.2019.10.017>
- Magalhães, I. A. L., Carvalho Júnior, O. A., Guimarães, R. F., Gomes, & R. A. T. (2022). Sentinel-1 time series analysis on Central Amazon floods. *Mercator*, 21, e21019. <https://doi.org/10.4215/rm2022.e21019>
- Manandhar, R., Odeh, I. O., & Ancev, T. (2009). Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sensing*, 1(3), 330-344. <https://doi.org/10.3390/rs1030330>
- Matgen, P., Hostache, R., Schumann, G., Pfister, L., Hoffmann, L., & Savenije, H. H. G. (2011). Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies. *Physics and Chemistry of the Earth, Parts A/B/C*, 36, 241-252. <https://doi.org/10.1016/j.pce.2010.12.009>
- Mayer, T., Poortinga, A., Bhandari, B; Nicolau, A. P., Markert, K., Thwal, N. S., ... Saah, D. (2021). Deep learning approach for Sentinel-1 surface water mapping leveraging Google Earth Engine. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 2, 10005. <https://doi.org/10.1016/j.ophoto.2021.100005>
- Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), 2784-2817. <https://doi.org/10.1080/01431161.2018.1433343>
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, (12), 153-157. <https://doi.org/10.1007/BF02295996>
- Melack, J., & Hess, L. L. (2010). Remote sensing of the distribution and extent of wetlands in the Amazon Basin. In W. J. Junk, M. T. F. Piedade, F. Wittmann, J. Schöngart & P. Parolin (Eds) *Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management* (vol. 120, pp. 43-59). Springer. [https://doi.org/10.1007/978-90-481-8725-6\\_3](https://doi.org/10.1007/978-90-481-8725-6_3)
- Millard, K., & Richardson, M. (2013). Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR-SAR fusion using a Random Forest classifier. *Canadian Journal of Remote Sensing*, 39(4), 290-307. <https://doi.org/10.5589/m13-038>
- Mohammadimanes, F., Bahram, S., Masou, M., Mahdi, M., & Brian, B. (2018). An efficient feature optimization for wetland mapping by synergistic use of SAR intensity, interferometry, and polarimetry data. *International Journal of Applied Earth Observation and Geoinformation*, 73, 450-462. <https://doi.org/10.1016/j.jag.2018.06.005>
- Moharrami, M., Javanbakht, M., & Attarchi, S. (2021). Automatic flood detection using Sentinel-1 images on the Google Earth Engine. *Environmental Monitoring and Assessment*, 193, 248. <https://doi.org/10.1007/s10661-021-09037-7>
- Nemni, E., Bullock, J., Belabbes, S., & Bromley, S. (2020). Fully convolutional neural network for rapid flood segmentation in synthetic aperture radar imagery. *Remote Sensing*, 12(16), 2532. <https://doi.org/10.3390/rs12162532>
- Pham-Duc, B., Prigent, C., Aires, F., & Papa, F. (2017). Comparisons of global terrestrial surface water datasets over 15 years. *Journal of Hydrometeorology*, 18(4), 993-1007. <https://doi.org/10.1175/JHM-D-16-0206.1>
- Rosenqvist, J., Rosenqvist, A., Jensen, K., & McDonald, K. (2020). Mapping of maximum and minimum inundation extents in the Amazon Basin 2014-2017 with ALOS-2 PALSAR2 ScanSAR time-series data. *Remote Sensing*, 12(8), 1326. <https://doi.org/10.3390/rs12081326>
- Sant'Anna, S. J. S. (1995). *Avaliação de filtros redutores de speckle em imagens de radar de abertura sintética* [Evaluation of speckle-reducing filters in synthetic aperture radar images]. [Dissertação de Mes-

- trado, Instituto Nacional de Pesquisas Espaciais]. Repositório INEP. [https://oasisbr.ibict.br/vufind/Record/INPE\\_b028be22ec6357381466210a3e818ee4](https://oasisbr.ibict.br/vufind/Record/INPE_b028be22ec6357381466210a3e818ee4)
- Schlaffer, S., Matgen, P., Hollaus, M., & Wagner, W. (2015). Flood detection from multi-temporal SAR data using harmonic analysis and change detection. *International Journal of Applied Earth Observation and Geoinformation*, 38, 15-24. <https://doi.org/10.1016/j.jag.2014.12.001>
- Siddique, M., Ahmed, T., & Husain, M. S. (2022). An empirical approach to monitor the flood-prone regions of North India using Sentinel-1 images. *Annals of Emerging Technologies in Computing*, 6(4), 1-14. <http://dx.doi.org/10.33166/AETiC.2022.04.001>
- Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Rostan, F. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, 120, 9-24. <https://doi.org/10.1016/j.rse.2011.05.028>
- Tsyganskaya, V., Martinis, S., Marzahn, P., & Ludwig, R. (2018). Detection of temporary flooded vegetation using Sentinel-1 time series data. *Remote Sensing*, 10(8), 1286. <https://doi.org/10.3390/rs10081286>
- Twele, A., Cao, W., Plank, S., & Martinis, S. (2016). Sentinel-1-based flood mapping: A fully automated processing chain. *International Journal of Remote Sensing*, 37(13), 2990-3004. <https://doi.org/10.1080/01431161.2016.1192304>
- Wang, D., Wan, B., Qiu, P., Su, Y., Guo, Q., & Wu, X. (2018). Artificial mangrove species mapping using Pléiades-1: An evaluation of pixel-based and object-based classifications with selected machine learning algorithms. *Remote Sensing*, 10(2), 294. <https://doi.org/10.3390/rs10020294>