Post-processing of Paddy Rice Fields from MapBiomas' Land use Land Cover. A Time-Weighted Dynamic Time Warping approach on Google Earth Engine

Abstract

This study evaluates the application of Time-Weighted Dynamic Time Warping (TW-DTW) for post-processing MapBiomas Land Use Land Cover data, focusing on paddy rice classification in the Araranguá hydrographic region, southern Brazil. A multi-source fusion framework was implemented, integrating temporal Sentinel-1 SAR (VV/VH) and Sentinel-2 optical (NIR, SWIR1, NDVI) imagery within Google Earth Engine to extract phenological signatures for agricultural classification. TW-DTW performs shape matching and temporal alignment of multi-dimensional time series, enabling pixel-wise comparison with reference phenological profiles. Validation against MapBiomas' deep learning-based irrigation dataset resulted Jaccard similarity coefficient of 81.99%. After thresholding, overlap metrics resulted of 88.96%. Computational demands limited the analysis to a single-year period, and classification bias was observed within wetland areas. Despite these constraints, TW-DTW achieved accuracy comparable to deep learning approaches while requiring 87% fewer training samples. These results highlight its potential for operational crop monitoring in data-limited contexts, offering an efficient and scalable complement to resource-intensive deep learning methods.

1. Introduction

The rising concern is at the establishment of a resilient biosphere for all beings. As human beings, the common census rises up on fundamental biological rights for its own survival: 1. Plenty of drinkable water & 2. Nutritional and enough food. Resilience of water demands a multifaceted approach from maintenance of springs to preservation of forests, as vital land-use class settlements. Increasing of world's population brings challenges for the nutritional food bullet point, and, as a vital sector over the global consumption chain, agriculture faces adaptation issues on its production, in order to enable effectiveness of yields and resilient food practices.

An important step towards a resilient and adapted agriculture agenda, also depends on good quality land-use and land-cover (LULC) datasets [1]. This methodology corroborates with the Sustainable Development Goals (SDG), including poverty and zero hunger. [2].

By classifying the diverse land management, LULC datasets provides fundamental insights the dynamic complex-system. Earth Observation satellites are the only source that provides a continuous and consistent set of information about Earth's land and oceans [3]. Due the revisit time of satellites, multiple images over a temporal series can be achieved at the same place. These observations can be organized in regular time intervals, to that each measure from sensor is mapped into a three-dimensional array in space-time.

At the data analysis perspective, many supervised and unsupervised algorithms to analysis time series have been developed over the last years. Therefore, the ability of producing quality LULC datasets is not enough to classify crop types. Different methods and approaches have been develop to classify crops using time series analysis. Gomes et al. (2016) [4] proved that time series images perform better to classify crop types than single-date mapping methods, especially for complex farming areas. One of the persisting bottlenecks is crop-specific seasonality, which requires time series to characterize the different phenological phases of crop development [5].

Through the last few years, EO domain has made some advances towards time-series analysis, including parcel boundary delineation extent [6], patterns of vegetation dynamics were used to map double cropping, single cropping, forest and pasture [3]. Temporal features from time series were used to classify cropland and pasture fields using Landsat data [7]. In addition, mapping crop types and classifying them follow two big fields, as a pixel-based analysis and object-based analysis.

The inclusion of irregular plant dates, spatial complexity, intercropping and complex terrain results in challenges for crop mapping in remote sensing. For instance, similar crops can exhibit different backscatter while different crops may exhibit similar backscatter [5], where the inter-crop similarity and the intra-crop heterogeneity can only be resolved in the temporal domain to account for the evolvement of spectral signatures and scattering mechanisms [5].

Research on time series data mining showed that methods based on dynamic time warping (DTW) have achieved significant results in many applications and seems to tackle the challenge of leveraging phenological information into a classification scheme. DTW is a method that compares temporal signature of a known event to an unknown event [3]. The algorithm compares two time series and finds their optimal alignment, providing as result a dissimilarity measure. [3]. Due the nature of the algorithm, DTW provides a robust distance measure when comparing time series even if they are irregularly sampled. However, Maus et al. (2013) argues that DTW is not suited per se for remote sensing time-series because it disregards the temporal range when finding the best alignment between two time series. The assumption is that a good time-series land-cover classifier should balance between shape matching and temporal alignment [3]. Therefore, Maus et al. introduces a refinement of the DTW algorithm by proposing time-weighted dynamic time warping (TW-DTW), a adaption which account the temporal constraint. This improvement creates extra costs using logistic weights in case the seasonal difference between the sample pattern and its match when the period is too large. The same feature controls the time warping and makes the time-series alignment dependent on the seasons.

On the application of TW-DTW on RS, Maus et al. used optical imagery to identify single cropping, double cropping, forest and pasture from MODIS data. Other studies contributed to examine crop classification. Belgiu and Csillik (2017) [6] investigate pixel-based and object-based TWDTW using Sentinel-2 optical imagery, identifying that OBIA outperformed pixel-based. The study also compared TWDTW with Random Forest (RF) machine-learning derived method and found out comparable results between both methods. However, TWDTW proved to be less sensitive in relation to the training samples, which indicates a degree of matter for mapping areas with imitated training samples points.

Gella and Belgiu (2021) [7] investigated Sentinel-1 and investigated the accuracy with derived indices and polarimetry decomposition features. Their research demonstrated that merging six different grain crops with similar leaf geometry improved accuracy. Moharana et Al. (2021) [8] investigated and concluded that parcel based TWD-TW approach conditioned by temporal signatures of RVI has effectively delineated croplands.

The time-weighted implementation by Maus et Al. was develop in R [8] and opened different applications on the domain, however, TWDTW is a computational expensively method which requires computer significance [3], therefore, an implementation on cloud-based platforms such as Google Earth Engine (GEE), make it possible to retrieve and calculate the time-weighted dynamic time warping for LULC purposes.

This study aims to apply TW-DTW method to post-process the Land Use Land Cover collection 8 of MapBiomas, the annuals network providers of LULC in Brazil. The study focus on improving the agricultural class LULC by mapping paddy rice crops over a hydrographic region in southern Brazil.

2. Materials and Methods

2.1. Study Area

The study area is Araranguá region in southern Brazil, located approximately at 49º52''S, 28º92'W, within the state of Santa Catarina. This region experiences a subtropical climate, with an average annual temperature of 23°C and an average annual precipitation of 1200 mm.

The region is well known for its high production of flooded rice crops (paddy rice), which has a significant positive impact on the local economy. However, this intensive cultivation also causes negative environmental effects, such as diffuse agrochemical pollution affecting the main rivers in the basin. The current study delimitate its Area of Interest (AOI) according to the guidelines of Brazil's Water Agenc. Considering the hierarchical composition of watersheds. Thus, the definition of hydrographic region encompass the assembly of higher hierarchical watershed levels. Gathering a total of 12 hydrographical regions within the state of Santa Catarina, on the south of Brazil. (see Fig. 1).

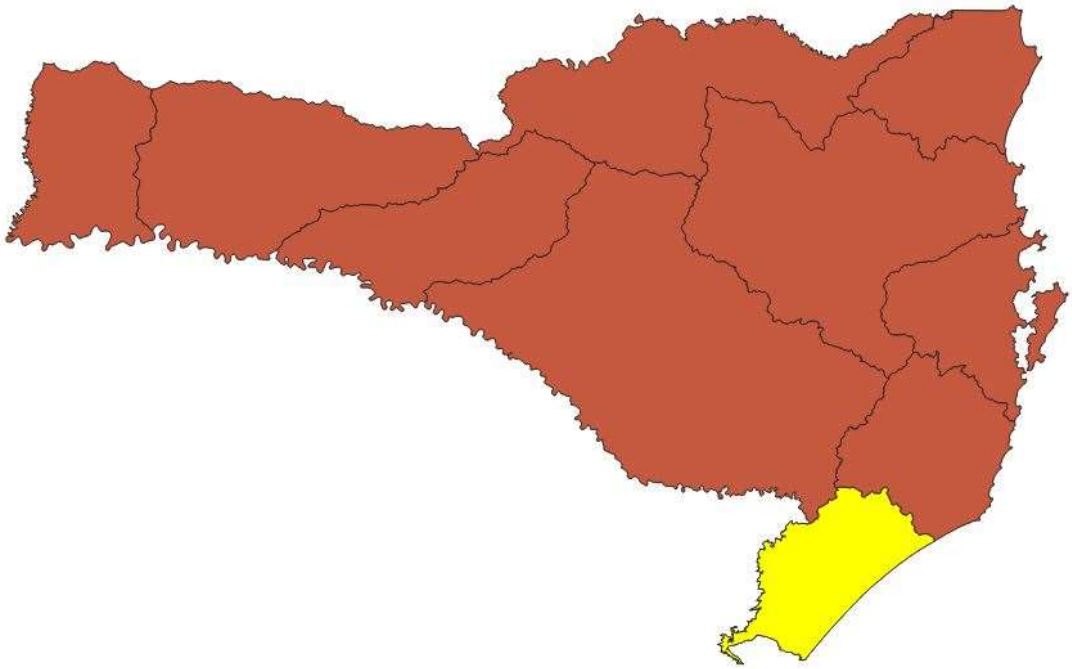


Figure 1: The 10th hydrographical region of Santa Catarina State represented by the highlighted feature.

The total area of the study region is approximately 48,814.00 km². The landscape is diverse, ranging from mountainous terrain in the western part, forming the plateau ("planalto") region of Santa Catarina, to a flatter, coastal plains influenced by the nearby sea.

Agricultural activity in the region is predominantly focused on paddy rice cultivation, with fields extending along the rivers throughout the basin. Additionally, livestock such as, chicken production, pig farming and forestry are also significant economic activities. The forestry sector mainly involves the cultivation of pine (Pinus elliottii) and eucalyptus (Eucalyptus spp.). Many small farms practice family-based agriculture, usually on small plots with a variety of crops such as lettuce, tomatoes, fruits, sugarcane and grains such as corn. In the AOI's western part, the landscape is more rugged and banana plantations are extremely frequent, covering the hills and surrounding areas, supplying local food processing industries and factories.

The figure 2 highlights the land-use land-cover of the region. Showing case the extension of the agricultural land on the hydrographical region. Most of the forest region is concentrated on western mountain part. The biggest city of the AOI is on the north, called Criciuma with a

population of 200.000 habitants, and followed by Araranguá with 80.000 habitants. Other cities within the AOI does not reach a total population of 20.000. Rangeland are sparse and mostly located close to the coast, having an influence of the sea and therefore a flora very characteristic of coastal areas.

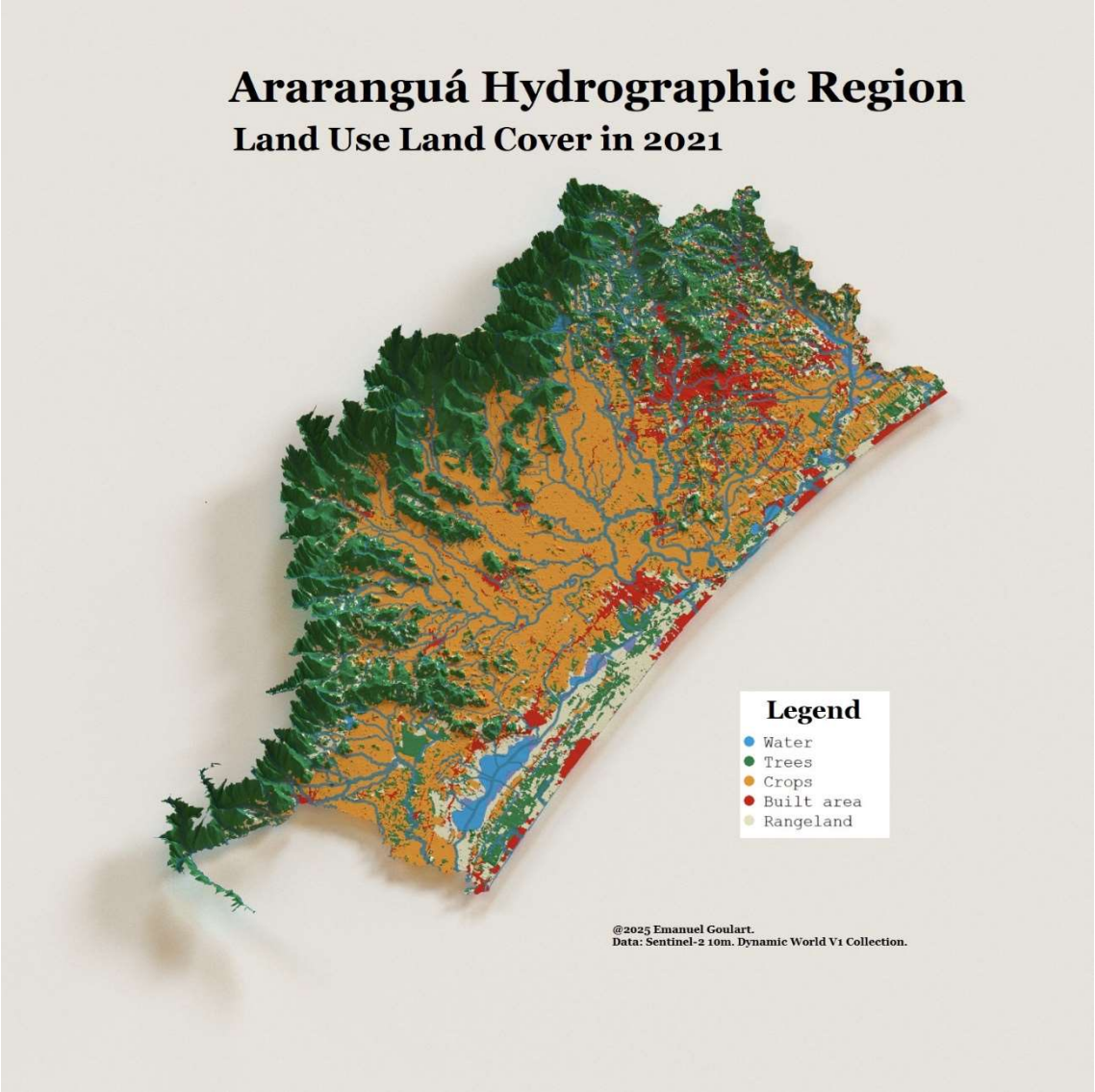


Figure 2- Land Use Land Cover of the region of study. Author.

2.2. MapBiomas – Land Use Land Cover

MapBiomas is a collaborative network composed of NGOs, universities, laboratories, and startups that conduct annual land cover and land use (LULC) mapping, as well as monthly monitoring of surface water, fire scars, and deforestation reports. MapBiomas operates in partnership with institutions focused on different biomes and cross-cutting thematic areas. In collaboration with Google Earth Engine, MapBiomas distributes its products as a GEE-powered platform.

As part of its collections, the legacy LULC maps are primarily produced from LANDSAT historical images with a spatial resolution of 30 meters. However, MapBiomas has recently released a beta collection of LULC maps derived from Sentinel-2 imagery at a finer spatial

resolution of 10 meters, known as Collection 2 10 meters. A technical note associated with this collection highlights its limitations and provides guidance for careful usage, especially concerning agricultural details and comparisons with other MapBiomas products at different spatial resolutions. This Sentinel-2 derived LULC collection spans the period from 2016 to 2023 and covers the entire country.

The collection includes five main classes: Forest, Herbaceous, Farming, Non-Vegetated, and Water. Each main class is further divided into subclasses; for example, Pasture, Forest Plantation, and Mosaic of Uses fall under the main class 'Farming'. To facilitate mask creation, the subclasses have been reclassified into their respective main classes, with the exception of Farming, which has been subdivided into Pasture, Agriculture, and Forest Plantation.

The figure 3 illustrates the reclassification scheme used in the MapBiomas beta LULC version.

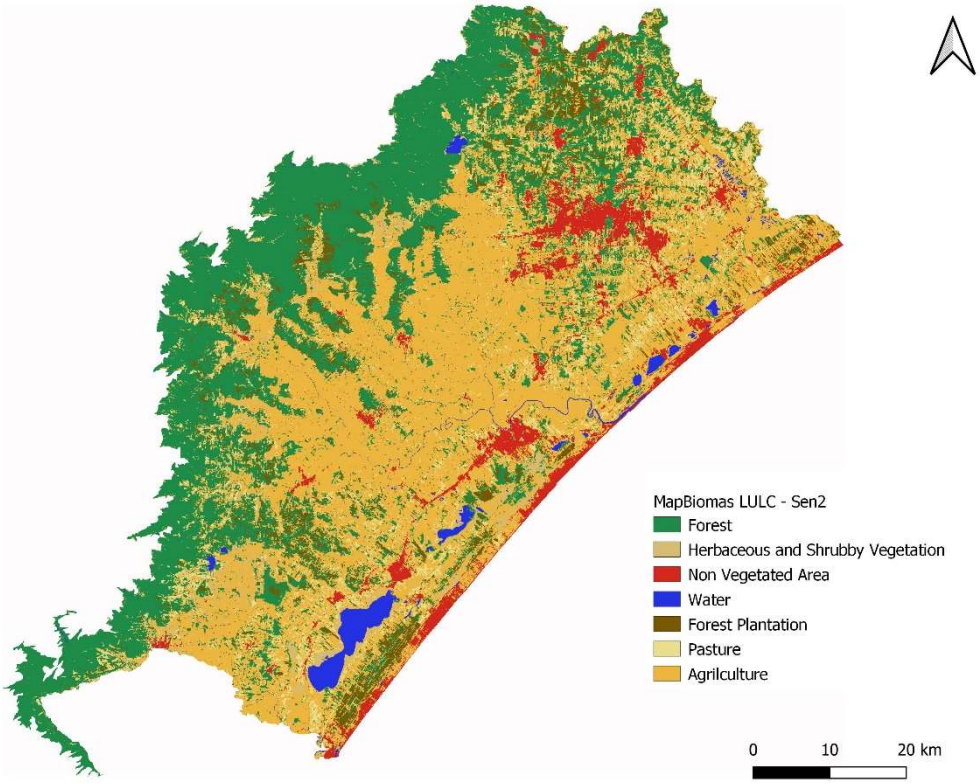


Figure 3 - Land Use Land Cover colecction of Mapbiomas. Author.

Additionally, MapBiomas offers a specialized module focused on irrigated lands, enabling the retrieval of paddy rice areas within regions of interest. The extent of paddy rice cultivation can be quantified within specific hydrographical regions.

The module is based on flooded field detection using Deep Learning (DL) techniques, specifically a neural network based on the U-net architecture. For the Santa Catarina region, 140 images chips were used for training, validation and testing, with a split ratio of 70%, 20% and 10% respectively.[13] the methodology is based on Deep-Learning methods [13].

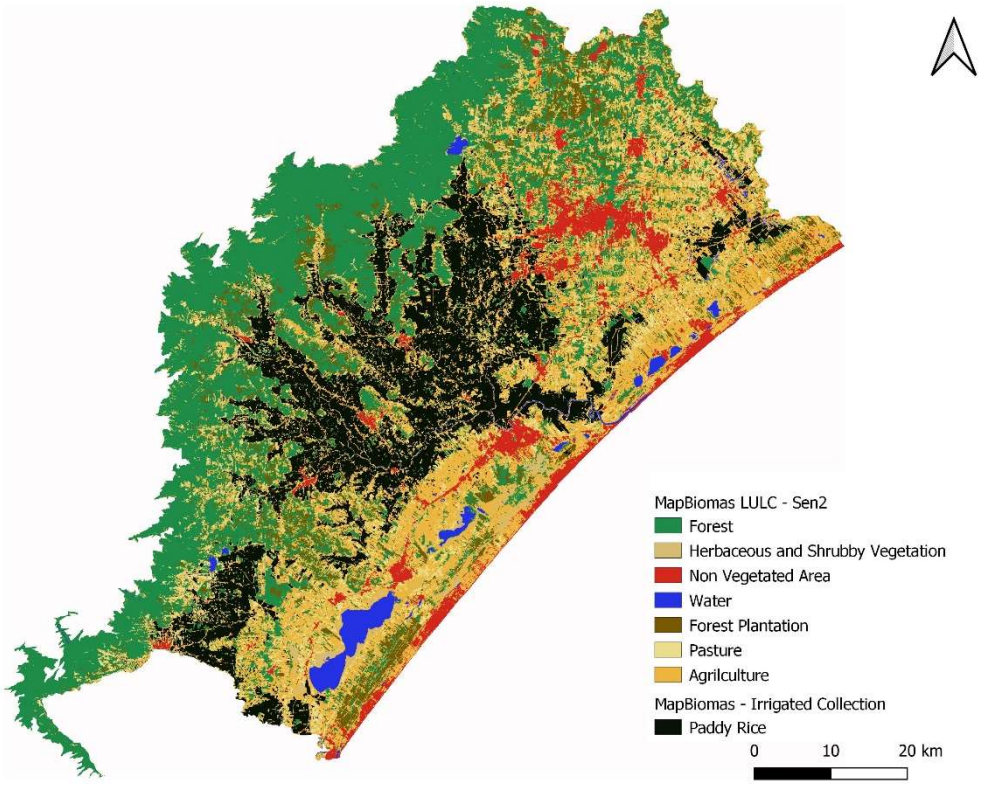


Figure 4- Flooded Rice classifcation of MapBiomas from the collection Irrigation. Author.

The figure 5 illustrates a chart containing the percentage of each land use and land cover type for the AOI.

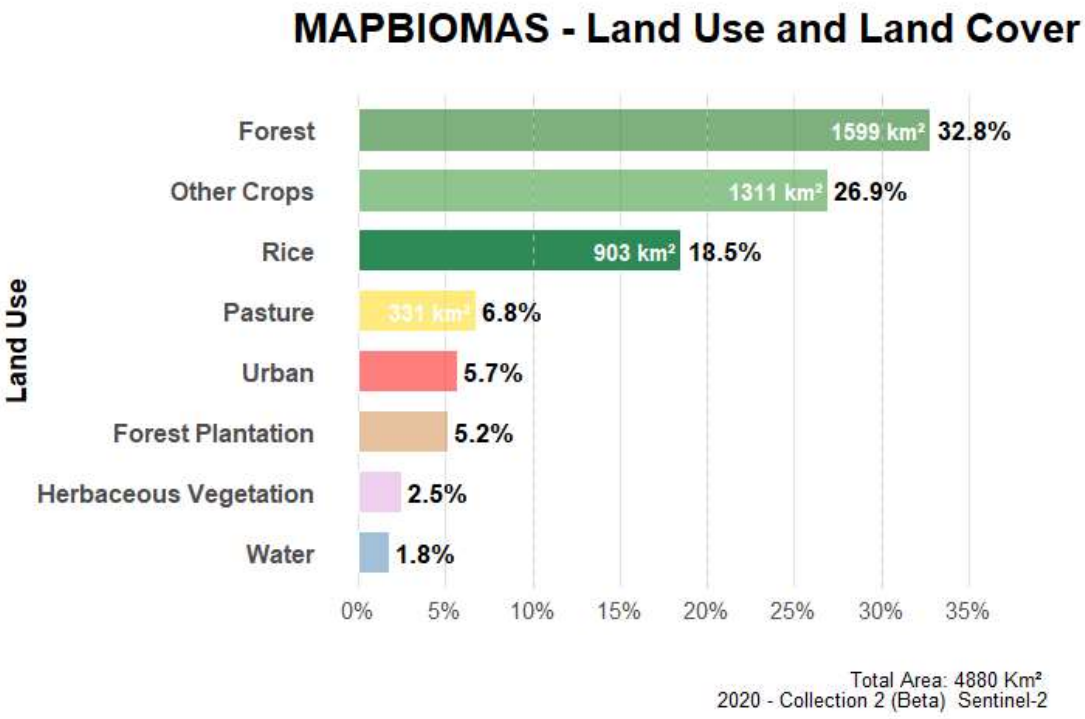


Figure 5 - Land Use of the 10th hydrographical region of Santa Catarina. Author.

The chart highlights that 25% of the land in the region is classified as forest, primarily distributed across the higher altitudes where a sequence of mountains is found. Farming land use

follows, with Pasture and Agriculture together accounting for 52% of the region's land use. This underscores the importance of farming and its significant contribution to the region's gross domestic product. Paddy rice crops make up 18.5% of the total land area, and when considering only the Farming class, rice cultivation represents nearly 41% of all agricultural fields. The region heavily depends on rice production, with the majority of this output exported to other states. Urban areas within the region occupy only about 5% of the total land area.

2.3. Training and validation samples

Manual labelling generated the training samples. A binary classification between Rice and Non-Rice was used to fine-tuning the Agricultural LULC classification. In total, 26 samples were acquihired (13=Rice; 13=Non Rice), everything labelled as a feature collection in Google Earth Engine.

Training Samples	
Class	Train
Rice	13
Non Rice	13

Table 1 - Samples for training the TW-DTW

2.4. Feature Space

This study utilized Sentinel-1 and Sentinel-2, Collection 2 Level 2 data from the year 2020. The dataset combines Sentinel-2 optical imagery with Sentinel-1 radar data, both retrieved from Google Earth Engine (GEE). For Sentinel-1, pre-processing followed the steps outlined by Mullisa et al. [15], which include procedures such as speckle filtering to prepare the Sentinel-1 SAR backscatter as Analysis-Ready Data in GEE. Sentinel-2 was pre-processed using the GEE module developed by the Soil Watch Project [14]. This module implements a masking procedure that removes clouds using morphological techniques and shadows by calculating the sun azimuth angle and identifying likely shadow pixels within cloudy images. This approach ensures that only high-quality pixels are retained for analysis.

Due to Sentinel-1 data availability constraints within the study region, the application of the TW-DTW method with the fused dataset was limited to the year 2020.

The feature space considered for TW-DTW in relation to the target includes:

- Sentinel 2: Band 08 - NIR
- Sentinel 2: Band 12 – SWIR1
- Sentinel 2 derived index: NDVI - Normalized Difference Vegetation Index
- Sentinel 1: VH – Horizontal Backscattering
- Sentinel 1:VV - Vertical Backscattering

2.5. Phenology Rice Paddy Patterns

Temporal phenology patterns of the target clases were computed using NDVI [9], generated from 10-meter resolution red and near-infrared Sentinel-2 spectral bands.

Rice phenology, as discussed by Bariani [12], indicates that the seedling stage occurs at the end of September. The crop then progresses toward its peak around mid-January. At this state, NDVI values typically reach their maximum, reflecting the full development of the crop canopy.

Following flowering, the rice enters the maturation phase, culminating in harvest when the seeds reach maximum weight, a timing strongly influenced by prevailing weather conditions.
The figure 6 illustrate the rice phenology for the southern Brazil part.

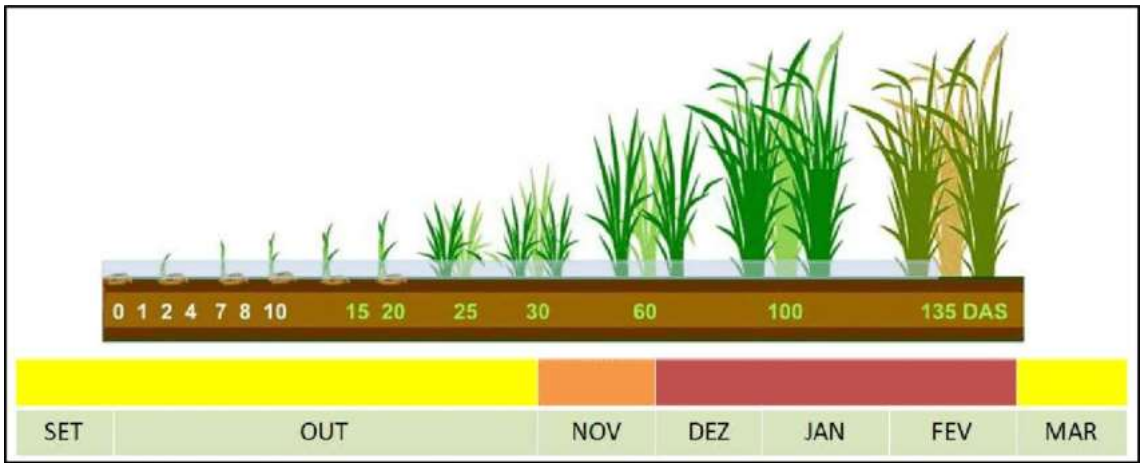


Figure 6 - Rice phenology. Bariani [12]

2.6. TW-DTW

The TW-DTW algorithm was by Maus et Al. (2016) [3], and the approach rely on the computation of the distance matrix between a data sequence at a given pixel location over the time series of a specific crop. The distance matrix D, with a size of the length of the sequence (M) and length specific crop temporal profile (N), the Euclidean distance is calculated by:

$$D[i,j] = tw_{i,j}\sqrt{\sum([SITS_{r,c}]_i - TPF_{j)}{}^2} \quad (1)$$

Where: D[i,j] is the distance at row I and column j of the distance matrix D.

$[SITS_{r,c}]_i$ is a value at row *r* and column *c* in time-series at time *I*.

TPFj is the temporal profile of a specific crop at a time *j*.

$tw_{i,j}$ is a logistic weight that accounts for the absolute differences between times *I* and *j,* which is computed as:

$$w_{i,j} = \frac{1}{1+e^{-a([i-j]-B)}} \quad (2)$$

Where: I is the time of observation in $[SITS_{r,c}]_i$ at the *ith* position in a time series, j if the time of observation in the temporal profile at the *jth* position. *B,* is the maximum time lag for the warping window to search for a match, which is constant for all crops, and *a* is a user-defined constant to control the steepness of the slope [5].

Willian Oullette gives the implementation of the algorithm in Google Earth Engine and the code can be find here: https://github.com/wouellette/ee-dynamic-time-warping

The GEE implementation allows the user to select between a time-weighted approach and a time-constrained approach. The distance can be calculated using either Euclidean Distance or Angular Distance, and there are adjustable parameters to control the weight of the logistic regression and the length of the temporal series.

Importantly, the GEE TW-DTW algorithm outputs two bands: one is the classification map, and the other is the dissimilarity score map. The classification map is generated using a "hard

decision" rule, where for each pixel, the class corresponding to the minimum dissimilarity score is assigned as the classified label.

However, this approach introduces an important nuance: the dissimilarity score band shows the minimum distance associated with the assigned class but does not explicitly indicate which class produced that minimum score. This ambiguity can be problematic if a user wants to analyze or threshold the dissimilarity scores directly without knowing their class origin.

To resolve this, one must use the classification band as a mask to isolate the dissimilarity scores for each class. By masking the dissimilarity score band with the classification labels, it becomes possible to determine which class corresponds to the given score at each pixel.

Recognizing this nuance is crucial because misinterpreting the dissimilarity scores without relating them back to the assigned classes can lead to incorrect conclusions when using these scores for further analysis or thresholding.

The table 2 expose the parameters chosen:

Time-weighte Dynamic Time Warping	Parameters
Constrain Type	Time-constrained
Distance measure	Euclidean Distance
Number of Patterns to Iterate over	12
Number of bands	['VV','VH','NDVI', ,'SWIR1','NIR']
Length of the image TS	12
Length of the Reference patterns	12
Beta	45 days
Alpha	0.1

Table 2 – TW-DTW parameters choice.

2.7. OTSU Thresholding

The TW-DTW method returns a band containing dissimilarity scores. However, not all of these scores accurately reflect the presence of paddy rice fields. To improve classification, Otsu thresholding is applied.

Otsu's thresholding is an image classification algorithm that automatically determines an optimal intensity threshold to separate image pixels into two classes: foreground and background. It does this by minimizing the intra-class variance — that is, the weighted sum of variances within the two classes formed by the threshold.

The algorithm exhaustively searches through all possible threshold values and selects the one that results in the smallest combined variance within each class, effectively maximizing the separability between foreground and background pixels. This makes Otsu's method particularly useful for segmenting images with bimodal histograms, where pixel intensities cluster around two distinct peaks. [16]

The implementation of OTSU is given by:

$$\sigma_{\omega}^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t) \quad (3)$$

Where: ω_0 represents the probabilities of the two classes separated by a threshold (t)

σ_0^2 represents the variance of the class 0.

2.8. Masked methods

A series of mask were applied over the TW-DTW for focusing on agriculture LULC class. The mask is given by classes' presents at the MapBiomas 10-m Beta Collection (see section 2.2). The list classes (see table 3) is first grouped into five major classes: Forest, Farming, Herbaceous and Shrubby Vegetation, Non-Vegetated Area and Water.

Mapbiomas Class Number	Class Name	Masked
1	Forest	True
2	Herbaceous and Shrubby Vegetation	True
4	Non Vegetated Area	True
5	Water	True
9	Forest Plantation	True
15	Pasture	True
18	Agriculture	False

Table 3 - LULC Classes used for mask

2.9. Post-Processing

The image was post-processed using pixel connectivity analysis to eliminate isolated rice pixels and reduce the "salt and pepper" noise effect. A connected component labelling method with an eight-connected neighbourhood and a kernel radius of 1 was applied to identify groups of connected pixels. Only connected components with a size greater than seven pixels were retained, effectively removing small isolated pixel clusters and improving the spatial coherence of the rice field classification.

The figure 7 exemplifies how the present study is post-processing the agriculture class and paddy rice classification of the Land Use and Land Cover collection from MapBiomas.

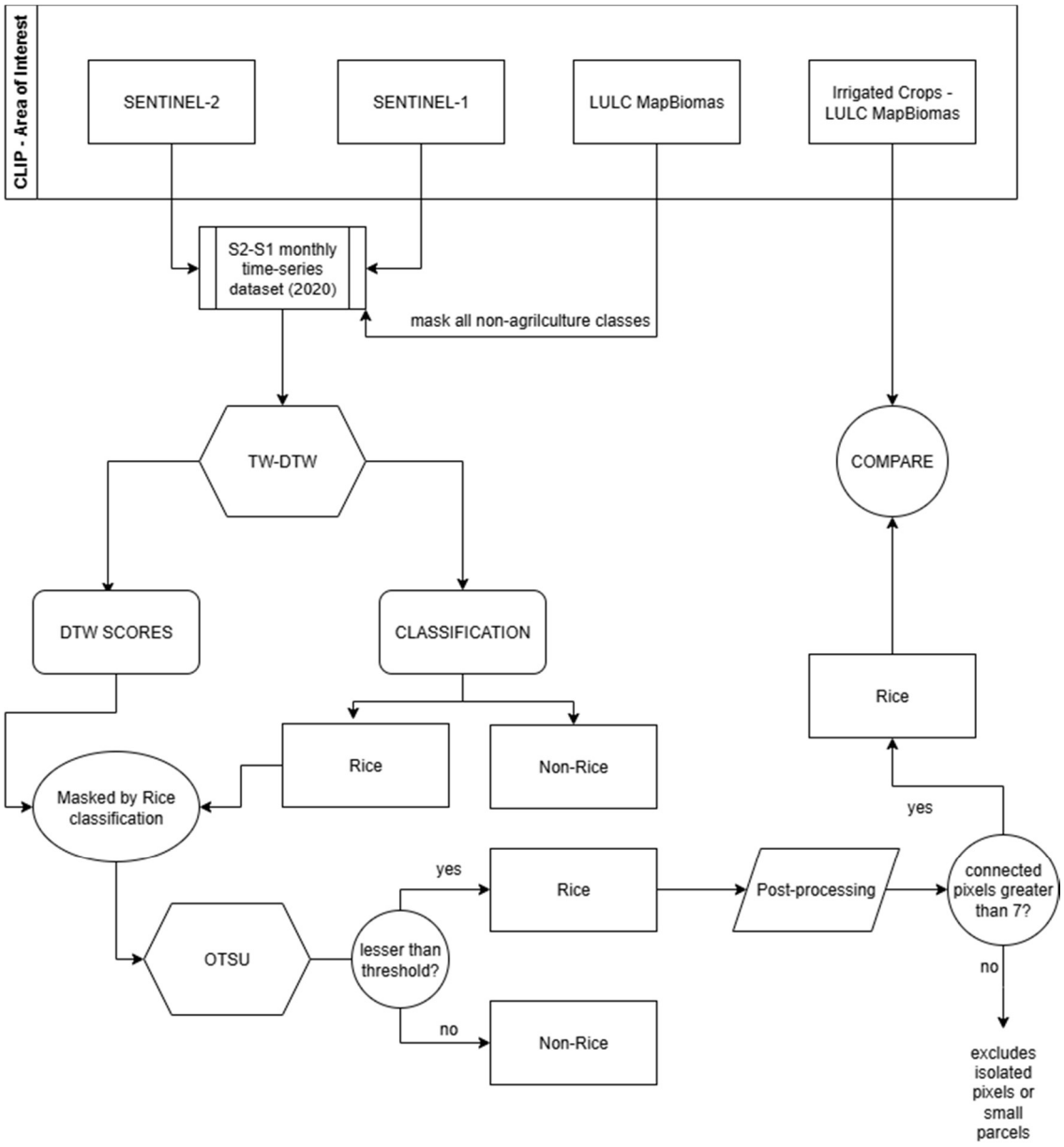


Figure 7 - Methodology diagram.

2.10. Accuracy Assessment

To verify the results, the TW-DTW were compare pixel-wise with the Irrigation Collection (see section 2.2). The number of matching pixels from the predicted TW-DTW algorithm was assessed with the Irrigation Collection.

In addition, a visual and spatial assessment was made to verify the inference.The assessment is based on Jaccard Index, given by:

$$jaccard = \frac{A \cap B}{A \cup B} \text{ [4]}$$

Where Jaccard Index is the ratio between the intersection and the union of the polygons.

Similarity Coefficient (Dice Coefficient), given by:

$$Dice\ Coefficient = \frac{2(A \cap B)}{A + B}$$

Overlap Coefficient, given by:

$$Overlap = \frac{A \cap B}{\min(A, B)}$$

3. Results

3.1. Phenology Analysis

The samples used on the TW-DTW algorithm were analysing to understand its dynamics and phenology. The figure 8 illustrate the backscatter and the spectral response of samples. An average of the two classes(Rice and Non Rice) was used to plot the time series.

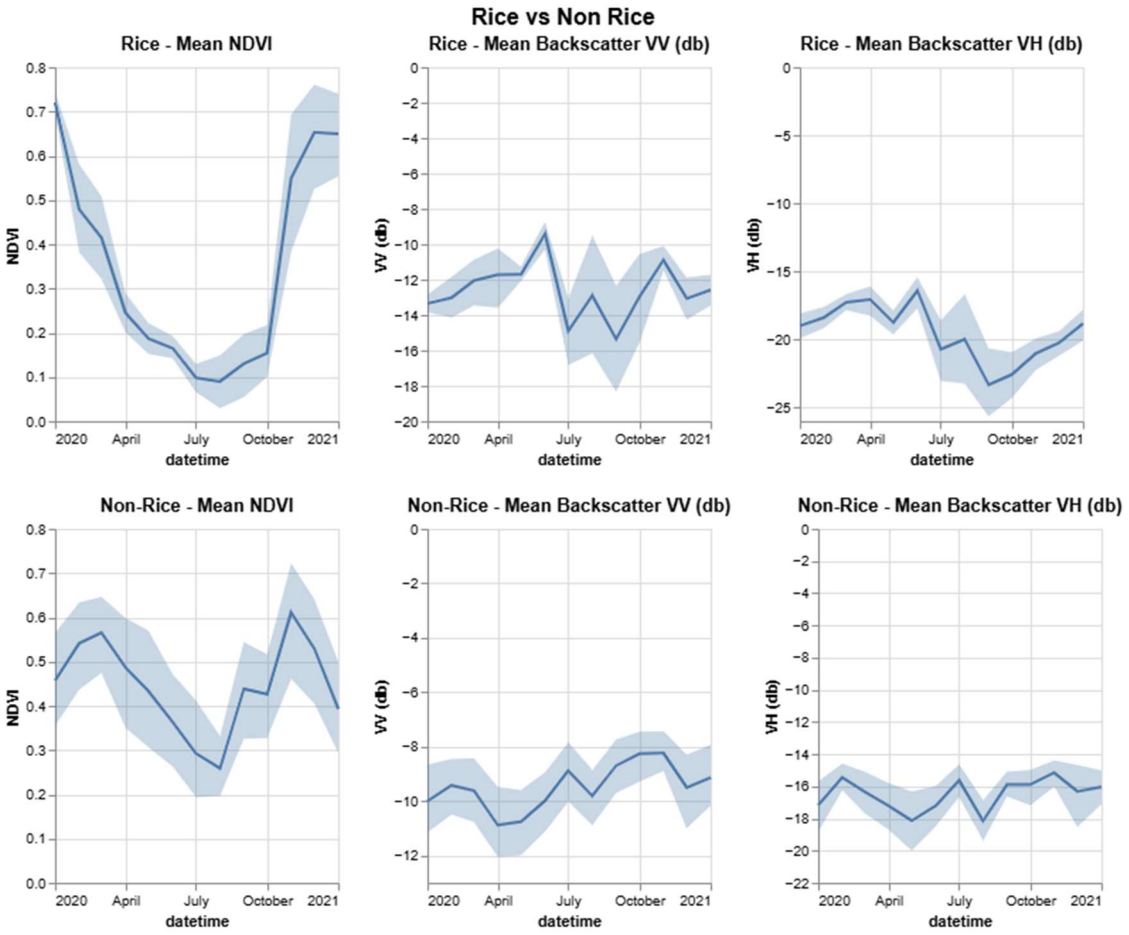


Figure 8 - Backscattering and spectral time series average for the samples points.

As discussed in Section 2.5, the phenology and traditional crop management of paddy rice begin in early September, when seeds are planted in inundated fields with a typical water depth of about 10 cm. The time-series spectral response of rice samples follows a very distinct pattern, reflecting the early seedling stage and rice sprout growth. NDVI gradually increases and reaches its peak in January, consistent with the described phenology [12]. The period of July and August corresponds to soil preparation, during which the soil is mostly exposed (bare soil), a condition clearly captured by the NDVI time series.

In contrast, the Non-Crop samples exhibit greater variability in NDVI, as they encompass multiple crop types. Unlike rice, the NDVI for Non-Crop samples does not reach the same minimum values during the bare soil period. Instead, several NDVI peaks are observed in: February, September, and December, highlighting considerable intra-class variability. This variability likely affects the TW-DTW classification performance, as the method has known limitations when dealing with high intra-class variability [2,3].

The backscatter coefficients for VV and VH polarizations do not show a clear temporal pattern but do overlap with the phenological stages. During the rice seedling stage in flooded fields, the VH backscatter reaches its minimum value, consistent with the typical low backscatter of water surfaces. Both VV and VH backscatter exhibit a peak in June, which can be attributed to soil preparation activities following harvest.

For Non-Rice samples, backscatter values display higher intra-class variability with no consistent pattern. While some peaks are noticeable, their causes are unclear due to the heterogeneity of crop types within the Non-Rice class. On average, backscatter values tend to be higher for Non-Rice samples compared to rice fields, suggesting potential differentiation of crop types detectable through radar sensing.

3.2. Multi-temporal NDVI composition

This multi-temporal NDVI composite illustrates how rice fields exhibit different greenness peaks throughout the year. The composite image is displayed using Red, Green, and Blue channels, where Red corresponds to March, Green to January, and Blue to June. Figure 9 highlights the crop field patterns within the area of interest.

The green channel clearly reveals the spatial extent of paddy rice fields, with January representing the peak growth month, consistent with the phenological cycle of rice cultivation in southern Brazil as described by Barreto [12] and observed at the section 3.1. This seasonal dynamic is further emphasized in the animated GIF provided in the Supplementary Materials, which visually demonstrates rice planting beginning in early September and harvesting occurring in late March.

During the winter period in the study area, soil preparation activities, such as tilling and revolving are prevalent, which is reflected by the low NDVI values indicating a gap in crop greenness. Black pixels in the composite represent masked land cover classes classified as Forest, Water, and Non-Vegetated areas (including urban regions).

The spatial distribution of the paddy rice fields in this NDVI composite also visually corresponds well with Figure 3, which presents the flooded rice classification developed by MapBiomas using deep learning models. This comparison highlights the effectiveness of time-series analysis in capturing phenological patterns specific to crop fields.

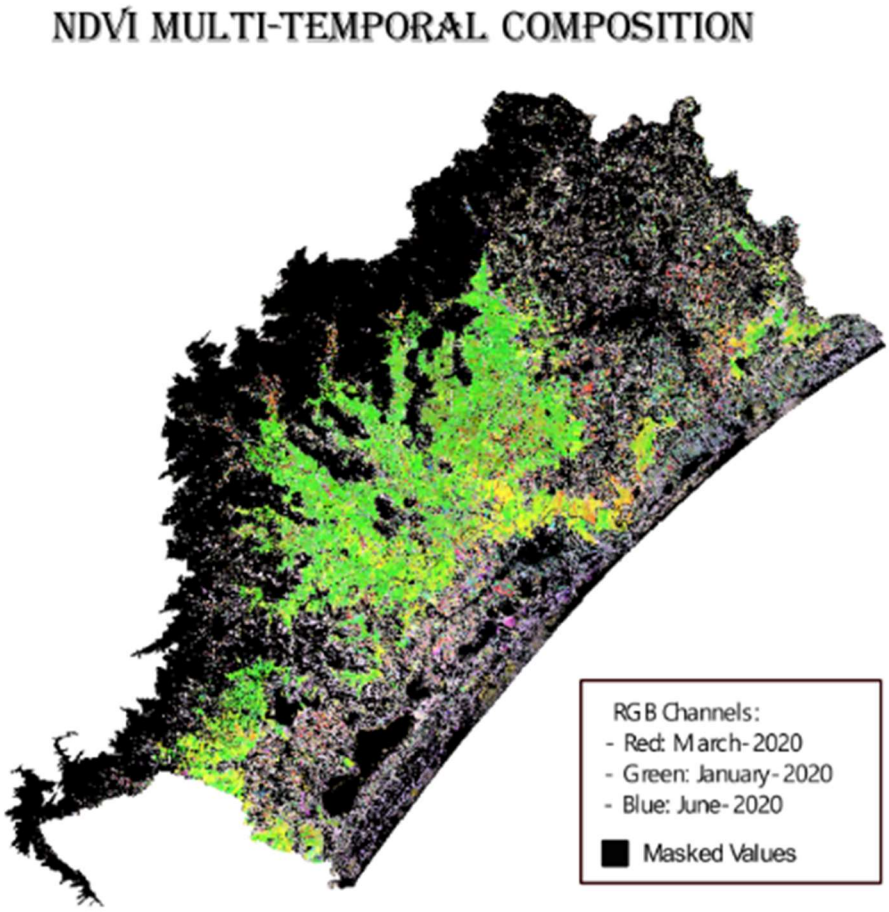


Figure 9 - Multi-temporal composition. Author.

3.3. Classified Classes with TW-DTW

Figure 10 presents the classification output of the TW-DTW method. The spatial distribution of the rice crops closely matches the patterns observed in the multi-temporal NDVI composite (Section 3.2) and the irrigated crop areas identified by the MapBiomas classification.

Rice cultivation is primarily concentrated along the Rio Araranguá river, which is part of the Araranguá watershed. Most paddy fields follow the river channels, reflecting the intensive water use during the seedling stage.

However, Figure 10 also reveals a noticeable salt-and-pepper effect, with many isolated rice pixels scattered throughout the map. This highlights the need for further post-processing to remove such noise and improve the spatial coherence of the classification.

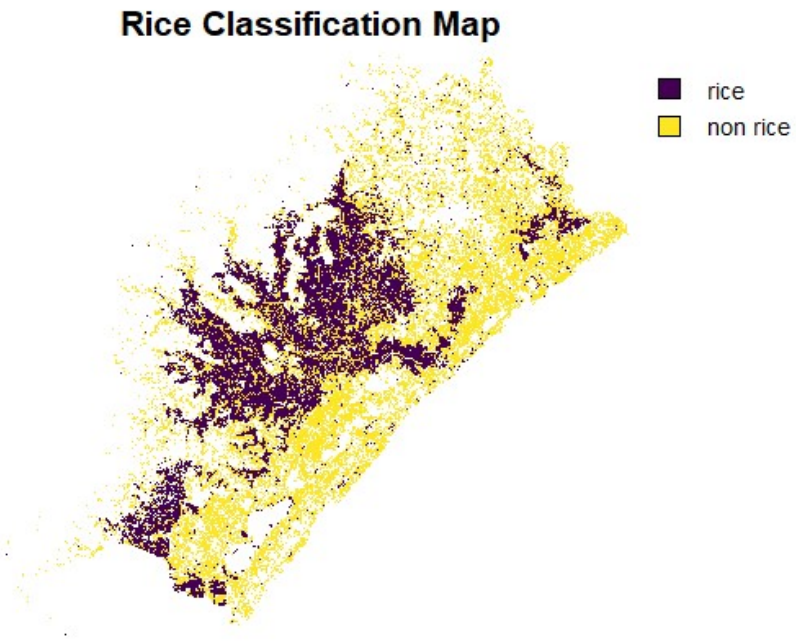


Figure 10 – Classification output of TW-DTW

DTW Dissimilarity Results

Figure 11 shows the dissimilarity score, which represents the DTW distance results produced by the TW-DTW algorithm for the classified rice and non rice class.

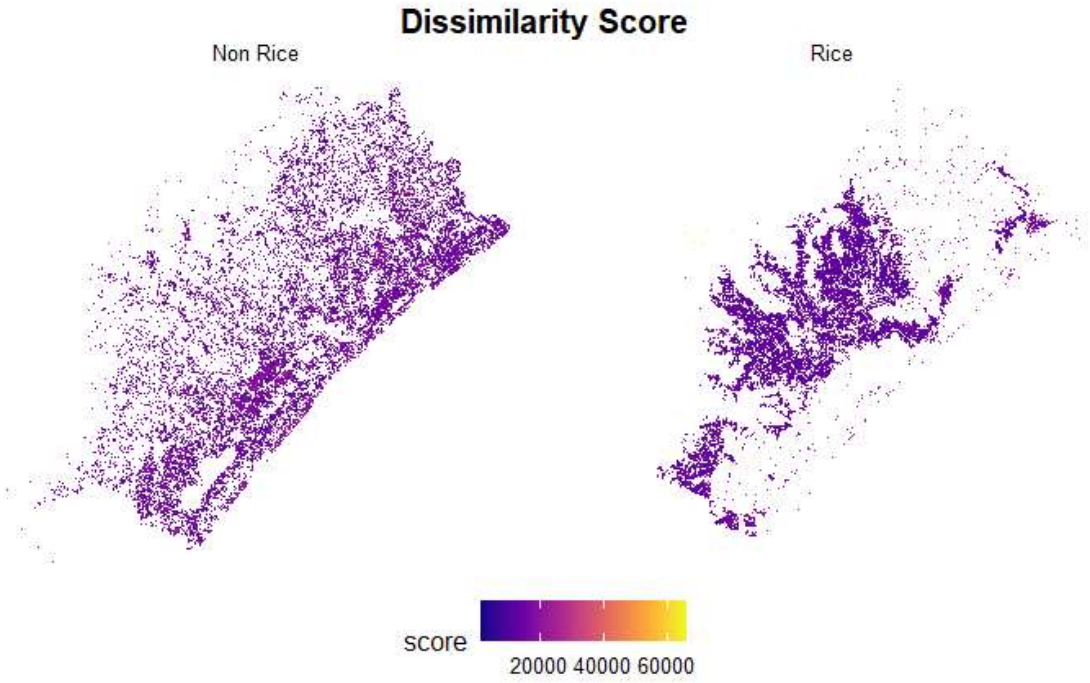


Figure 11 - TW-DTW, Dissimilarity Score

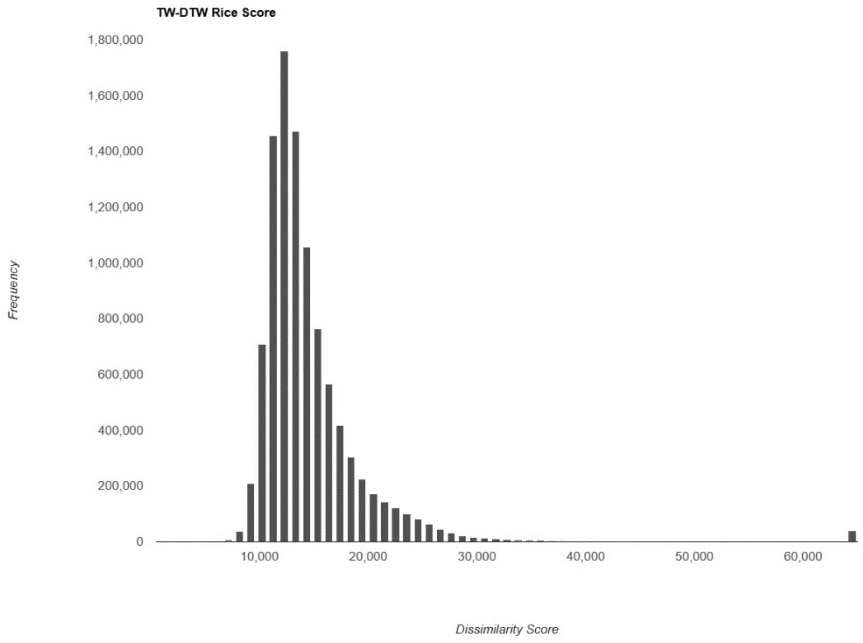
These dissimilarity scores represent the distance between each pixel's time series and the reference rice time series. Generally, a lower distance indicates a higher similarity to the target

class, while a higher distance suggests that the pixel is less likely to belong to that class. In the dissimilarity map, paddy rice fields tend to have scores ranging from 0 to about ~20,000. Beyond the 20,000 threshold, the number of pixels with such high dissimilarity scores decreases substantially, indicating fewer samples with lower similarity. This suggests that the TW-DTW algorithm effectively uses these scores to assign classes, and thresholding serves as a deterministic way to fine-tune classification by separating likely rice pixels from others.

As shown in Figure 12, the histogram of rice dissimilarity scores exhibits a distribution that appears to approximate either a Poisson or bimodal distribution. A formal statistical test should be conducted to determine the exact distribution; however, such a test was not performed in this study.

Notably, the rice dissimilarity scores display a lower standard deviation compared to the Non-Rice scores, indicating less variability within the rice class. The rice histogram also shows a pronounced tail and a relatively small secondary peak around a dissimilarity score of 60,000. This feature is absent in the Non-Rice histogram, which, while exhibiting greater variability, is confined to a narrower score range compared to rice.

These differences suggest distinct underlying characteristics in the dissimilarity score distributions between rice and Non-Rice classes, which may have implications for thresholding and classification accuracy.



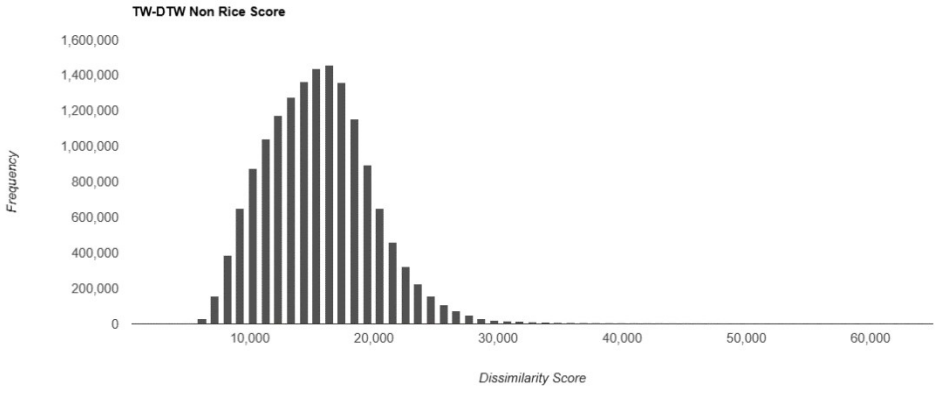


Figure 12 - Histogram of the dissimilarity score.

3.4. OTSU

OTSU was the method employed for retrieving the thresholding the DTW score. The result of the OTSU was 24,957.78.

It is important to side note that the arbitrary threshold was a choice over recursively analysis of threshold that best suits and generalize the rice crops extension for the AOI.

The table 4 shows the rice classification threshold by the OTSU value.

Method	Threshold
OTSU	24957.78
Arbitrary (based on expert knowledge)	33000

Table 4 - OTSU threshold and arbitrary threshold

Following the analysis, the table 5 represents the assessment of the predicted results and the applied thresholds in comparison with the Rice Irrigation collection from MapBiomas;

Asset	Area (km²)	Intersection (km²)	Union (km²)	Jaccard Index	Overlap	Dice Coefficient	Absolute Error (km ²)
Irrigation MapBiomas – Rice	9029.62	-	-	-	-	-	-
TWDTW – Rice (All prediction)	8583.86	7403.13	9029.62	81.99%	86.26%	0.8406	445.76
OTSU Threshold (24.957)	8306.36	7389.53	9029.62	81.84%	88.96%	0.8525	723.26
Arbitrary Threshold – (33.000)	8521.98	7400.48	9029.62	81.96%	86.83%	0.8432	507.64

Table 5 – Comparison of TWDTW classification and thresholds limits with the Irrigated Rice MapBiomas classification.

The dice coefficient represents a reasonable value of similarity between the predicted TW-DTW pixel-classification with the compared collection. The area of prediction has a mean absolute error of 558km² over a total area of 48,814.00 km².

Discussion

The central aim of this study was to evaluate the accuracy of TW-DTW to post-process MapBiomas' LULC of Irrigation Agriculture. Particularly, Irrigated rice class, at southern Brazil. The question was whether TW-DTW could refine LULC outputs when compared with the original MapBiomas collection.

TW-DTW's strength lies in capturing temporal patterns and providing a robust distance metric for time series. When compared to the classification of UNet-based MapBiomas model, results shows a degree of similarity, by spatially matching. For the inference of TW-DTW, it was observed that for low values of dissimilarity score, pixels seem to capture well temporal characteristic and represent robustly the classification of rice crops.

Although this work did not explore the TW-DTW feature space in depth, its outputs still reproduced rice patterns with fidelity. Jaccard index showed an 81.99% overlap with the irrigated rice class from the MapBiomas Irrigation Collection. Considering the area of 8583km² and 9029km² for TW-DTW and Irrigation collection respectively, which represents an absolute error of 446 km². Showing a likely underestimate of the total paddy rice area within area of study.

The TW-DTW implementation in GEE produces two outputs: a classification map and a dissimilarity score. The class assignment is, per pixel, the minimum dissimilarity value from the all classes. In this case, the distribution presented at figure 12, shows a peak of high dissimilarity score (>45000), indicating that thresholding based on dissimilarity scores could improve performance.

OTSU threshold results, increased both Overlap index and Dice coefficient unless Jaccard index. The assessment with Jaccard index does not shows any evidence of improvement when applying a threshold. Although Otsu may not represent the most optimal method for thresholding dissimilarity scores, it remains a robust technique with effective generalization capabilities for other areas. Figure 7 indicates that dissimilarity scores do not follow a bimodal distribution as required by the Otsu method for achieving optimal separation between foreground and background

However, when comparing threshold results with arbitrary thresholding from multiple tests, the score value of 33,000 captures most paddy rice areas with sensitivity that Otsu was unable to achieve. While Otsu represents a more computer vision field, treating scores as a computer vision problem raises discussions within the domain.

One limitation of TW-DTW thresholding is that it can only adjust the distribution of true positives and false positives. Pixels misclassified as other classes (false negatives) remain unrecovered, as they lie outside the classification scope considered in this study.

Additionally, the computational expense of the method warrants emphasis. To achieve these results, temporal scope was constrained to a single year (2020) rather than multiple consecutive years. This imposes limitations on the method since many crop types may exhibit phenological cycles exceeding one year, exposing TW-DTW's fragility in capturing complete growth cycles. For instance, crops that do not exhibit well-defined growth-decline patterns

observable from above (e.g., passion fruit, dragon fruit) remain a limitation, as DTW may interpret their temporal signatures as similar features to bushes or other vegetation types.

The results suggest that TW-DTW achieved an 81% similarity level when compared with the deep learning U-Net-based Irrigation Collection from MapBiomas. Given that deep learning methods demand larger sample populations than TW-DTW, they may not represent the most suitable approach for data-scarce regions. Furthermore, TW-DTW, utilizing only 13% of the training dataset compared to U-Net's requirements, achieved comparable similarity results.

Specifically, the Irrigation Collection indicates 18.44% paddy crop coverage within the AOI, compared to 17.88% derived from TW-DTW classification.

Post-processing the results using pixel connectivity and a focal mean helped reduce the salt-and-pepper effect, removing many isolated false-positive pixels that did not correspond to rice fields. The high number of such isolated false positives suggests a recurring similarity in their time-series patterns across the chosen spectral bands. Figure 13 indicates a concentration of false positives near lakes and rivers, particularly over wetland vegetation. This bias may stem from seasonal hydrological patterns: in southern Brazil, heavy precipitation at the end of winter and beginning of spring (late August to early September) coincides with the seedling stage of paddy rice, when farmers flood parcels. This temporary water signal may be misinterpreted by the TW-DTW algorithm as indicative of rice cultivation.

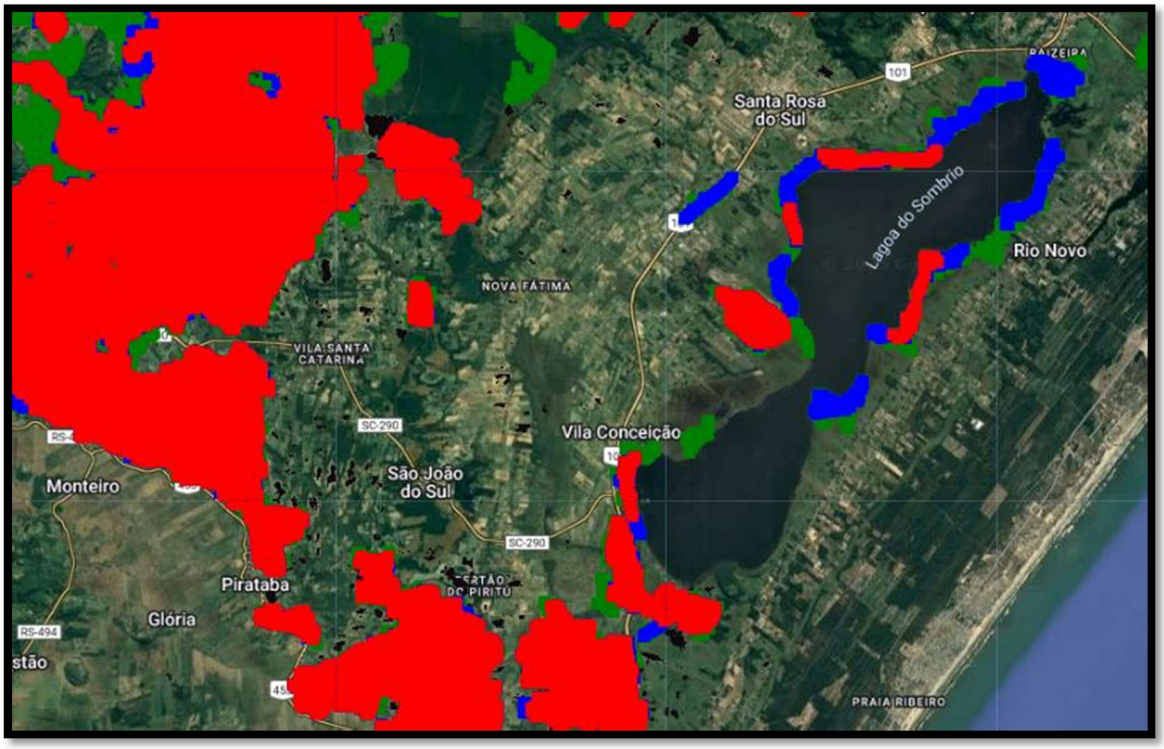


Figure 13 – Both blue and red colors are inference results from TW-DTW. The color blue, underneath the red, represents the TW-DTW classification results without thresholding. Whereas the blue represents the threshold classification

The thresholding strategy aimed to reduce false-positive pixels from the TW-DTW classification. Figure 14 illustrates this improvement: blue areas indicate false positives in the full TW-DTW output, while green highlights the results after applying the Otsu threshold, showing enhanced classification sensitivity.

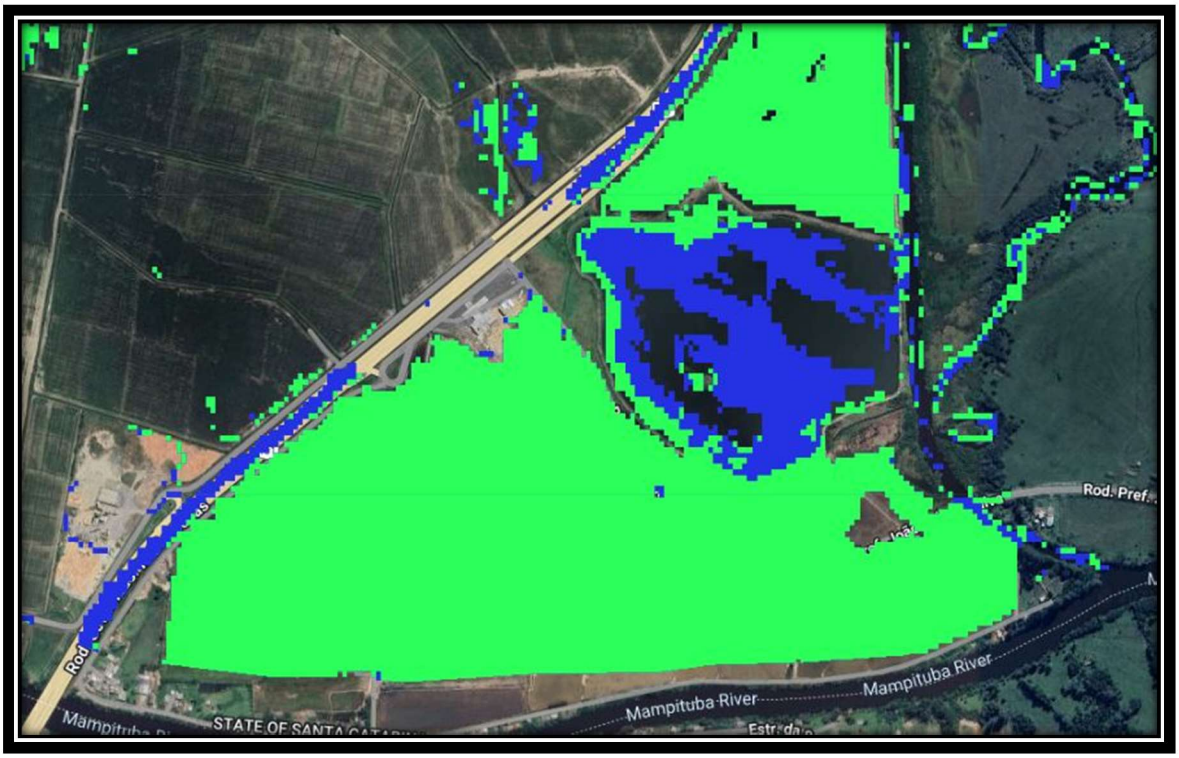


Figure 14 - Difference of the Rice classification for the hard-classification with DTW and the post-processed Rice classes with OTSU thresholding.

Preliminary results from Section 3.1 indicate that TW-DTW is effective for crop-type classification, particularly in regions with limited data availability, absent or incomplete reference datasets, and no labeled training/testing parcels. In such contexts, TW-DTW emerges as a promising method for identifying crop types with reliance on the temporal characteristics of crops.

Limitations and Future Developments

TW-DTW exhibits numerous vulnerabilities with substantial room for improvement. For instance, only polarimetric bands VV and VH, along with optical bands SWIR1 and NIR and the NDVI index, were tested for inference. Alternative band fusion approaches could be explored, incorporating indices that better preserve temporal feature characteristics. It remains uncertain whether additional bands or indices would contribute important information regarding temporal similarity scores and further improve results.

An alternative approach, as proposed by Moharana et al. [8], involved evaluating optimal backscatter parameters before algorithm execution. They partitioned data for training and validation, discovering that individual polarization and ratio-based indices performed better in mapping field samples. Additionally, simple cross-ratios (VV/VH) failed to sub-classify certain crop types. The current study did not attempt to optimize bands and select optimal backscatter and spectral signatures, representing a limitation.

Google Earth Engine's computational processing limitations expose the method's computational expense and suggest that alternative approaches could prevent reaching GEE's maximum memory allowances for limited cloud-based resources.

Multi-year TW-DTW could be incorporated by classifying paddy rice crops across sequential years to increase confidence in paddy rice parcel identification.

Sensitivity to time-warping parameters warrants exploration. The study by Moharana et al. [8] investigated sensitivity and determined that classification is sensitive to alpha and beta values, achieving optimal results with a maximum time delay of 120 days (Beta). This study did not test different parameter impacts on TW-DTW, imposing current limitations on accuracy results.

Future developments should integrate radar-based satellite scattering mechanisms related to crop development, such as polarimetric decomposition features, entropy, anisotropy, embedding and anglemeasurements.

Either the dissimilarity score could be used as a temporal tensor or combined architecture on deep learning models or TW-DTW could be represent a feature of a DL model.

Conclusion

This study focused on paddy rice field classification using TW-DTW to post-process Brazil's annual Land Use Land Cover dataset provided by MapBiomas, comparing results with MapBiomas' deep learning methodology for creating paddy rice classification LULC datasets. Implemented within Google Earth Engine, the study also incorporated thresholding of TW-DTW results, achieving 88% overlap in comparing with MapBiomas collection.

Arbitrary thresholding appears to better delineate biased and false-positive areas but proves more challenging to generalize across other regions, whereas Otsu demonstrates superior generalization capacity with greater adaptability for other areas. TW-DTW exhibits strength in capturing temporal features and pixel-based classification through similarity analysis while requiring remarkably 87% fewer samples for regional classification.

The methodology demonstrates significant potential for operational crop monitoring applications, particularly in regions characterized by limited training data availability. The algorithm's ability to achieve comparable accuracy with substantially reduced sample requirements positions it as a valuable tool for agricultural surveillance in data-constrained environments.

However, several limitations must be acknowledged, including computational constraints that restricted analysis to single-year periods, potential wetland misclassification bias, and the need for enhanced parameter optimization. Future research should prioritize multi-temporal analysis, improved spectral-polarimetric feature integration, and comprehensive sensitivity assessments of algorithm parameters.

The integration of TW-DTW with existing LULC products represents a promising technique for enhancing agricultural monitoring capabilities while capturing phenological seasonal pattern.

References

[1] S. Fritz et al., “The need for improved maps of global cropland,” Eos Trans. Amer. Geophys. Union, vol. 94, no. 3, pp. 31–32, 2013.

[2] United-Nations, 2015b. Resolution adopted by the General Assembly on 25 September 2015. In: Washington: United Nations: UN General Assembly.

[3] V. Maus, G. Câmara, R. Cartaxo, A. Sanchez, F. M. Ramos, and G. R. de Queiroz, “A Time-Weighted Dynamic Time Warping Method for Land-Use and Land-Cover Mapping,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 8, pp. 3729–3739, Aug. 2016, doi: 10.1109/JSTARS.2016.2517118.

[4] Gómez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: a review. ISPRS J. Photogramm. Remote Sens. 116, 55–72.

[5] G. W. Gella, W. Bijker, and M. Belgiu, “Mapping crop types in complex farming areas using SAR imagery with dynamic time warping,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 175, pp. 171–183, May 2021, doi: 10.1016/j.isprsjprs.2021.03.004.

[6] Persello, C., Tolpekin, V.A., Bergado, J.R., de By, R.A., 2019. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. Remote Sens. Environ. 231 https://doi.org/10.1016/j. rse.2019.111253.

[7] Müller, H., Rufin, P., Griffiths, P., Siqueira, A.J.B., Hostert, P., 2015. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. Remote Sens. Environ. 156, 490–499

[8] S. Moharana, B. V. N. P. Kambhammettu, S. Chintala, A. S. Rani, and R. Avtar, “Spatial distribution of inter- and intra-crop variability using time-weighted dynamic time warping analysis from Sentinel-1 datasets,” *Remote Sensing Applications: Society and Environment*, vol. 24, p. 100630, Nov. 2021, doi: 10.1016/j.rsase.2021.100630.

[9] Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150. Doi: 10.1016/0034-4257(79)90013-0

[10] A. Mullissa et al., “Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine,” Remote Sensing, vol. 13, no. 10, p. 1954, May 2021, doi: 10.3390/rs13101954.

[11] K. Rolão et al., "Análise de rentabilidade entre o cultivo de arroz irrigado e cultivo de arroz sequeiro," 2019, doi: 10.22533/at.ed.3231916054.

[12] C. Bariani, N. M. Bariani, and G. Neto, "Sensoriamento remoto aplicado ao monitoramento dos principais estádios fenológicos do arroz irrigado no Sul do Brasil," 2017, doi: 10.7127/iv-inovagri-meeting-2017-res2560757.

[13] E. Weber, K. S. M. Santos, and P. D. P. Teixeira Junior, "Irrigation - Appendix C9," MapBiomas, Collection 9, Aug. 2024. [Online]. Available at: https://brasil.mapbiomas.org/wp-content/uploads/sites/4/2024/08/Irrigation-Appendix-C9.docx.pdf

[14] SoilWatch, "ee-dynamic-time-warping," GitHub. [Online]. Available: httpxas://github.com/wouellette/ee-dynamic-time-warping. Accessed: Aug. 9, 2025.

[15] A. Mullissa et al., "Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine," Remote Sens., vol. 13, no. 10, p. 1954, 2021, doi: 10.3390/rs13101954.

[16] N. Otsu, "A threshold selection algorithm from graylevel histograms," IEEE Trans. Syst. Man Cybern., vol. SMC-9, no. 1, pp. 62–66, Jan. 1979. doi: 10.1109/TSMC.1979.4310076

Supplementary Materials

- GEE code, R scripts at GITHUB:
  - https://github.com/emanuel-gf/tw-dtw-rice-classification
- GIF of Monthly NDVI values:
  - https://github.com/emanuel-gf/tw-dtw-rice-classification/blob/main/GIF-NDVI.gif