



ESJ Natural/Life/Medical Sciences

Evaluation Of Machine Learning Classification Methods For Rice Detection Using Earth Observation Data: Case Of Senegal

Fama Mbengue,

Laboratoire de Télédétection Appliquée-LTA, Institut des Sciences de la Terre, Université Cheikh Anta DIOP (UCAD) de Dakar, Sénégal
Département de Physique, Université Cheikh Anta DIOP (UCAD) de Dakar, Sénégal

Gayane Faye,

Laboratoire de Télédétection Appliquée-LTA Institut des Sciences de la Terre, Université Cheikh Anta DIOP (UCAD) de Dakar, Sénégal

Kharouna Talla,

Département de Physique, Université Cheikh Anta DIOP (UCAD) de Dakar

Mamadou Adama Sarr,

Université Gaston Berger (UGB) de Saint-Louis, Sénégal

André Ferrari,

Université cote d'Azur, OCA, UMR Lagrange, France

Modou Mbaye,

Omar Marigo,

Aissata Thiam,

Laboratoire de Télédétection Appliquée-LTA, Institut des Sciences de la Terre, Université Cheikh Anta DIOP (UCAD) de Dakar, Sénégal

Mamadou Semina Dramé,

Laboratoire de Physique de l'atmosphère et de l'océan-Simeon Fongang, Université Cheikh Anta Diop De Dakar, Senegal

Papa Sagne,

Laboratoire de Biostratigraphie-Sédimentologie, Département de Géologie, Université Cheikh Anta Diop de Dakar, Sénégal

Laboratoire de Télédétection Appliquée, Institut des Sciences de la Terre, Université Cheikh Anta Diop de Dakar, Sénégal

[Doi:10.19044/esj.2022.v18n17p214](https://doi.org/10.19044/esj.2022.v18n17p214)

Submitted: 24 November 2021

Accepted: 09 February 2022

Published: 31 May 2022

Copyright 2022 Author(s)

Under Creative Commons BY-NC-ND

4.0 OPEN ACCESS

Cite As:

Mbengue F., Faye G., Talla K., Adama Sarr M., Ferrari A., Modou M., Marigo O., Aissata T., Semina Dramé M., & Sagne P., (2022). *Evaluation Of Machine Learning Classification Methods For Rice Detection Using Earth Observation Data: Case Of Senegal* European Scientific Journal, ESJ, 18 (17), 214 <https://doi.org/10.19044/esj.2022.v18n17p214>

Abstract

Agriculture is considered one of the most vulnerable sectors to climate change. In addition to rainfed agriculture, irrigated crops such as rice have been developed in recent decades along the Senegal River. This new crop requires reliable information and monitoring systems. Remote sensing data have proven to be very useful for mapping and monitoring rice fields. In this study, a rice classification system based on machine learning to recognize and categorize rice is proposed. Physical interpretations of rice with other land cover classes in relation to the spectral signature should identify the optimal periods for mapping rice plots using three machine learning methods including Support Vector Machine (SVM), Random Forest (RF), and Classification and Regression Trees (CART). The database is composed of field data collected by GPS and high spatial resolution (10 to 30 m) satellite data acquired between January and May 2018. The analysis of the spectral signature of different land cover show that the ability to differentiate rice from other classes depends on the level of rice development. The results show the efficiency of the three classification algorithms with overall accuracies and Kappa coefficients for SVM (96.2%, 94.5%), for CART (97.6%, 96.5%) and for RF (98% 97.1%) respectively. Unmixing analysis was made to verify the classification and compare the accuracy of these three algorithms according to their performance.

Keywords: Rice agriculture, Senegal River Delta, Machine Learning, sentinel-2, Google Earth Engine

Introduction

Rice is one of the most important crops in the world, and it is the food staple for nearly 40% of the world's population (Lacharme, 2001). Senegal is one of the largest consumers of rice in West Africa. The average annual rice consumption per person is between 60 and 90 kg (Mendez Del Villar et al., 2011). Rice is grown mainly in the Senegal River valley in the north and in the Anambée basin in the south of the country. According to projections, production in the river valley should cover 57% of the country's needs, given the favorable conditions for double-cropping rice based on the SAED Final Report (2019).

Earth Observation data are increasingly becoming a real source of information for monitoring rice crops (Kuenzer & Knauer, 2013; Le Toan et al., 1997). Indeed, different methods of discrimination analysis related to rice varieties are presented in the scientific literature (Belder et al., 2004; Bouman et al.,

2007; Kshirsagar & Pandey, 1995; Kuenzer & Knauer, 2013; Radanielina et al., 2013; Singh et al., 2000; Tuong et al., 2005; Wassmann et al., 2004).

Research in this area can still be improved. Remote sensing promises scalable, low-cost, and unbiased estimates of rice area to support, augment, improve, or even replace survey and statistical methods (Gumma et al., 2014). However, there are technical challenges to the development of rice information systems at the national scale.

The aim of this study is to evaluate different machine learning methods for optimizing rice crop monitoring in northern Senegal using land observation data.

In Senegal, rice production systems are largely dominated by small-scale family farms. Therefore, there are three main types of rice production in Senegal:

- Traditional or rainfed lowland or upland rice production in the southern regions (Fatick, Ziguinchor, Sédhiou, Kolda, Tambacounda, and Kédougou).
- Irrigated rice cultivation in the Senegal River valley and in the Anambé basin.

Rice is an herbaceous plant with a round stem covered with flat sessile leaves and a terminal panicle. Under favorable and exceptional climatic conditions, the plant can grow for more than one year. The overall growth duration of a rice variety can be divided into 4 components: a basic vegetative phase, a phase of sensitivity to photoperiodism (if the variety is photoperiodic), a phase of sensitivity to temperature, and a reproductive phase from panicle initiation to maturity (Lancashire et al., 1991).

In the transplanted system, the rice plants are grown in a seedbed for about 20 days, with transplanting in a hillside configuration. Before transplanting, the rice field is flooded with water to depths ranging from 2 to 15 cm (Boschetti et al., 2014). This deliberate agronomic flooding is a key component of most remote sensing rice detection algorithms (Veloso et al., 2017). After transplanting, the recommended practice is to maintain the water level at about 3 cm and gradually increase it to 5-10 cm with increasing plant height. For rice established by direct seeding, rice seeds are sown at a higher density per unit area directly into moist soil or soil with a water level of 2-5 cm (Torbick et al., 2017). Due to the higher population density under direct seeding, tillers production is suppressed (Yoshida et al., 1971). The increase in leaf volume in the canopy is mainly due to the appearance and growth of leaves from the main culms. Keeping the water level low for the first 10 days after transplanting or for 21 days of direct seeding is recommended for crop management.

Materials and Methods

Presentation of the Study Area

The Senegal River Delta (**Figure 1**) extends from Dagana to the mouth of the river and covers an area of 394,530 ha. It is characterized by a Sahelian climate dominated by a short rainy season (July to October) and a long dry season (November to June). Rainfall is low and highly variable with annual totals of about 200 mm. (Touré, 2018). The average annual temperature is relatively high, sometimes exceeding 40°C in Podor department (ANDS, 2020).

The Senegal River delta is made up of a mosaic of morphological units ranging from dune formations, inter-dunes, levees, settling basins, mud flats, and littoral units (strips and beaches) (Michel and Sall). The ecosystems of the delta, like all ecosystems, are dependent on rainfall. It is a very dynamic area because many hydro-agricultural developments have been carried out there over the past decades for rice production. The area to be studied covers the rice plots located in the commune of Diama and Ronkhe in the department of Dagana.

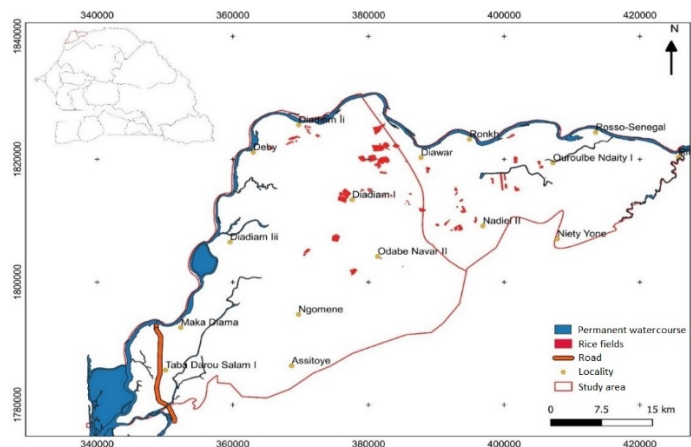


Figure 1. Map of study area

The Data

A total of 38 blocks of rice (one or more sticky fields), sown at different dates (in red in **Figure 1**), were digitized to serve as a training and test base (Table 1). These are field boundaries (or groups of fields) delineated in the field with a GPS (**Figure 2**). The rice field samples were divided into three blocks according to the date of sowing (see **Table 1**). For each land use type as well (Rice, soil bare, water, urban and other vegetation consisting of other crops and natural vegetation), training zones are labeled based on visual interpretation on Google Earth platform images.

Sentinel-2 optical data of level1 C/S2 are used in this study. These data were taken between January 1 and May 31, 2018, for a total of 66 images (Table 2). The access to these data and their pre-processing are made from the Google Earth Engine (GEE) platform. For each decade, a synthesis image (average image) is calculated (see **Table 2**).

Table 1. The Data

Block number	Number of parcels	Sowing dates	Total Parcels	Sown
Block 1	02	2018/02/04	46	
	06	2018/02/10		
	26	2018/02/11		
	01	2018/02/12		
	03	2018/02/13		
Block 2	07	2018/02/14	66	
	11	2018/02/15		
	02	2018/02/17		
	35	2018/02/19		
	08	2018/02/20		
	03	2018/02/21		
Block 3	05	2018/02/22	667	
	02	2018/02/24		
	121	2018/02/25		
	42	2018/02/26		
	03	2018/02/27		
	187	2018/02/28		
	144	2018/03/01		
	61	2018/03/02		
	23	2018/03/03		
56	2018/03/04			
30	2018/03/05			



Figure 2. Rice polygon sample

Table 2. Satellite and field data

Satellite images		66 Images Sentinel-2 Level-1C				
Month	January	February	March	April	May	
Dates of acquisition	//////////	2018/02/02(3)	2018/03/04(3)	2018/04/03 (2)	2018/05/03 (3)	
	2018/01/08(3)	//////////	2018/03/09(3)	2018/04/08 (1)	2018/05/08 (3)	
	//////////	2018/02/12(3)	2018/03/14(3)	2018/04/13 (3)	2018/05/13 (3)	
	2018 /01/18(3)	2018/02/17(2)	2018/03/19(3)	2018/04/18 (2)	2018/05/18 (3)	
	2018/01/23(3)	2018/02/22(3)	//////////	//////////	2018/05/23 (3)	
	2018/01/28(3)	2018/02/27(3)	2018/03/29(2)	//////////	2018/05/28 (3)	
Images divide into decades	j1 :1 st decad	f1: 1 st decad	ms1 :1 st decad	a1: 1 st decad	ma1: 1 st decad	
	j2: 2 nd decad	f2: 2 nd decad	ms2 :2 nd decad	a2: 2 nd decad	ma2: 2 nd decad	
	j3: 3 rd decad	f3: 3 rd decad			ma3: 3 rd decad	

Method

The method is based on the following phases:

1. Compute a number of spectral indices for each pixel (NDVI, NDWI and EVI):

NDVI (Normalized Difference Vegetation Index (Rouse Jr et al., 1974)) is the most commonly used vegetation index in remote sensing:

$$NDVI = \frac{NIR - RED}{NIR + RED} = \frac{B8 - B4}{B8 + B4} \quad (1)$$

The NDWI (Normalized Difference Water Index (Gao, 1996)) is an index that uses the mid-infrared band instead of red, which varies with the water content of the vegetation:

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR} = \frac{B8 - B12}{B8 + B12} \quad (2)$$

The Enhanced Vegetation Index (EVI) is designed to minimize saturation and background effects in NDVI (Huete et al., 2002):

$$EVI = 2.5 \times \left(\frac{(NIR - RED)}{(NIR + 6 \times RED - 7.5 \times BLUE + 1)} \right) \quad (3)$$

2. Extract from the Google Earth Engine (GEE) platform the pixel values (for all calculated bands and indices) of the control plots for all images (between January and May) using 20069 training polygons. For each block, the pixel values over the entire period and for all bands and indices are extracted for the spectral analysis.

3. Analyze the spectral profiles of the different classes at different phases of rice cultivation (January to May) to identify the most discriminating bands and the optimal period to better differentiate rice from other classes (bare soil, built-up, other vegetation and water). As a reminder, the spectral signature represents the measurement of energy in relation to the different wavelengths returned by a target.

Each object has its own spectral signature. Thus, this step consists of determining the most suitable period for classification and selection of the images of the most discriminating period but also the most relevant bands and/or indices.

4. Do the unmixing which consists, for each pixel, to estimate the contribution of each of the five classes. This linear spectral mixing model is based on the assumption that each pixel is a mixture of "pure" spectra. The pure spectra, called endmembers, are estimated from the values of the training areas.

5. Use GEE to perform supervised classification using the most relevant data. The first step is to manually create different training data. Using the geometry tools and the Sentinel-2 composite as a background, digitize the training polygons. This known sample set will be partitioned into training and test sets. For this purpose, 2/3 of the samples are used as training base. Finally, evaluate the accuracy of each of the three classifications using 1/3 of the field samples as a validation basis. Indeed, in this study, the field samples are divided into two bases: 70% as training base and 30% as test data (Genuer & Poggi, 2017). This makes it possible to estimate the errors of each model.

Three types of pixel-based classifications are tested in this study:

- Classification And Regression Trees (CART), introduced by Breiman et al. (1984), which builds tree-based predictors for both regression and classification. CART is a machine learning algorithm that can take non-linear patterns in the data. The general principle of CART is to recursively partition the input data X (X : pixel value) in a binary fashion, and then determine an optimal sub-partition for prediction. It uses in situ data to build a descriptive and predictive model of a relationship between a set of predictors and a categorical variable (Steinberg et al., 2012).
- The Random Forest classifier is a collection of random trees whose predictions are used to compute a mean (regression) or vote on a label (classification). The only parameter of the classifier is the number of trees (k). The RF creates different trees using a number of subsets of features. Each tree produces a classification result, and the classification model result depends on the majority of votes (Liu et al., 2012).
- Support vector machines (SVM) were initially proposed by Vapnik (1999) to solve classification and regression analysis

problems (Jha & Raha, 2013). SVM is a supervised learning technique that is trained to classify different categories of data from various disciplines. SVM creates a hyperplane or multiple hyperplanes in a high-dimensional space, and the best hyperplane among them is the one that optimally divides the data into different classes with the largest separation between classes (Schiilkop et al., 1995). The theory and detailed mathematical explanation of SVMs have been demonstrated in many previous studies (Ben-Hur & Weston, 2010; Cortes & Vapnik, 1995; Foody & Mathur, 2004). The basic linear SVM was used. All simulations were performed using the freely available LibSVM package (Chang & Lin, 2011).

This paper proposes a rice classification system based on machine learning to recognize and categorize rice. Physical interpretations of rice with other land cover classes in relation to the spectral signature should identify the optimal periods for mapping rice plots using three learning methods including Support Vector Machine (SVM), Random Forest (RF), and Classification and Regression Trees (CART). The database is composed of field collection and Sentinel-2 multispectral imager (MSI) Level-1C satellite data. Although the classification can correctly detect most rice-growing areas, certain types of vegetation cover can lead to classification errors. Wet or seasonal areas subject to drying followed by sudden vegetation growth can contribute to increased omission errors. The solution is to use multi-temporal optical images acquired at appropriate times, often outside the rice growing season, to exclude some non-rice growing areas. Furthermore, this study examines the suitability of sentinel-2 optical spatial data and the effectiveness of three machine learning methods in predicting the appropriate period for rice field mapping.

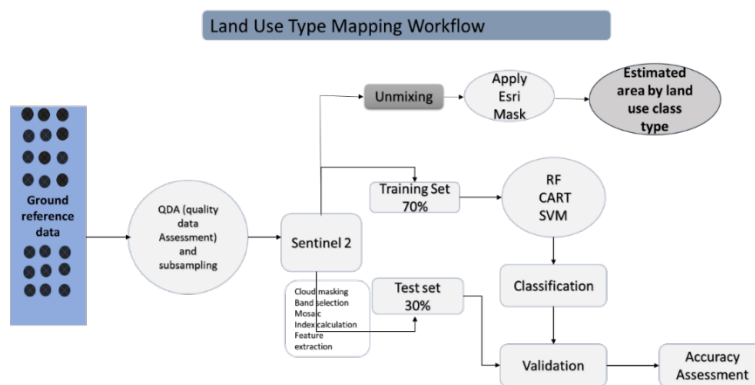


Figure 3. Flow chart of the method

Results and Discussions

Temporal Analysis of the Spectral Signatures of the Different Classes

The curves in **Figure 4** show the distributions of radiometric values for each land cover class for each decade from January to May 2018.

In January (the pre-sowing period), the reflectance of the rice fields merges with that of the urban and natural vegetation except for strips B2 to B5, where the natural vegetation has a lower response than the rice fields. This reflectance corresponds to the period when the fields are plowed and developed, and others are occupied by grasses. From February onwards, there is a decrease in rice reflectance, particularly for strips B7, B8, B8A, B11 and B12, with a wide distribution of values. This may be due to the fact that during this period, some rice fields are sown (early sowing) and not recommended, while others remain uncultivated. The same situation was observed in March and until early April. From the second dekad of April, the distribution of radiometric values of the rice fields remains low compared to the vegetation, with a spectral signature quite identical to that of the other vegetation, but with lower values. This phase is the reproductive phase where the temporal signature of the canopy is similar to that of rice. This shows that the development of rice goes hand in hand with that of the vegetation. However, weed control makes the rice plots more open and therefore less reflective than the other vegetated areas.

This difference in reflectance between rice and vegetation is reduced in the last decade of May, becoming almost zero in early June, corresponding to the maturity of the rice.

These curves show that during the vegetative period of rice, it is possible to discriminate rice with spectral bands B6, B7, B8, B8A, B11, and B12.

Indeed, the reflectance spectrum of the vegetation cover is the result of a complex relationship between its biophysical and biochemical attributes (Yang et al., 2007). As a result, the analysis of the separability of the radiometric values of rice fields with other land cover classes should help identify the optimal periods for mapping rice plots. The structure of a vegetation cover is not related to plant organs, but concerns the plant or stand. It takes into account both canopy structure parameters, such as leaf area index or leaf tilt angle, and the spatial organization of stands, their arrangement, density, as well as the rate of ground cover according to phenological stage. It should be noted that the life span of tropical rice varies from 110 - 120 days to 150 days (Jay & David, 2002; Le Toan et al., 1997; Nguyen, 2016), from germination to maturity in general.

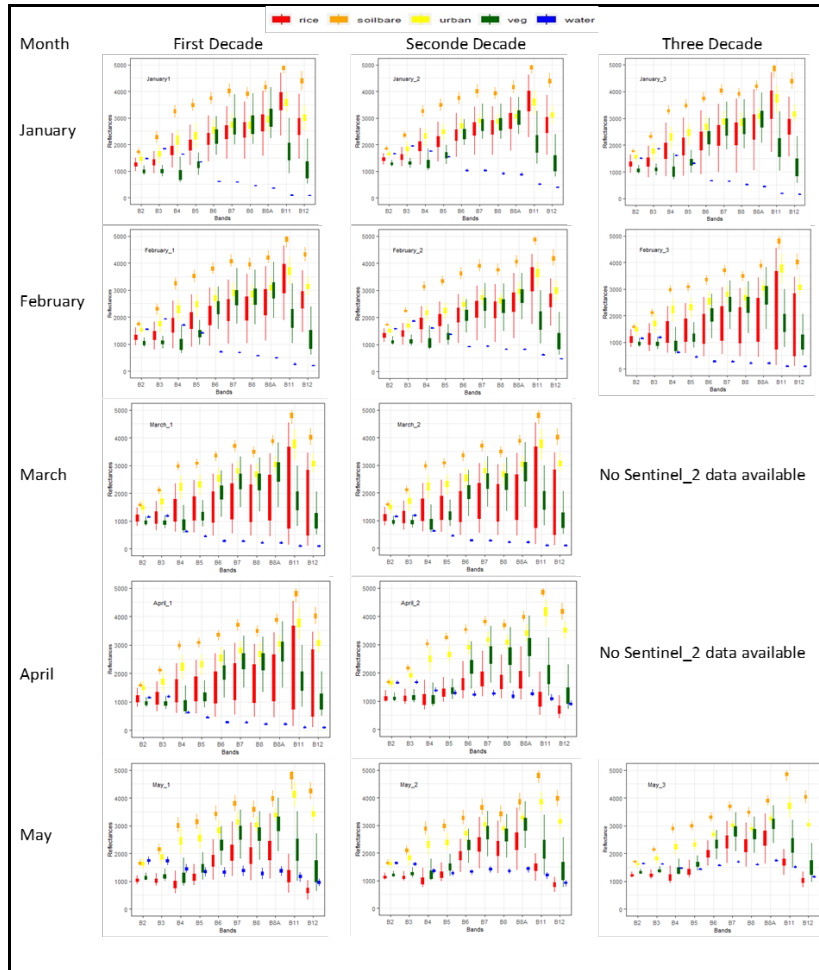


Figure 4. Distributions of radiometric values of each land cover class for each decade from January to May

Class Separability Analysis

The spectral analysis of the different land cover classes is not sufficient as a result to discriminate rice from other land cover classes. This is because rice discrimination appears to be quite difficult. One solution is to analyze the separability index of rice from other classes, given by the following formula:

Separability Index

$$= abs \left(\left((mean_{rice} - mean_{otherclass}) \right) \left(ecartype_{rice} - ecartype_{otherclass} \right) \right) \quad (4)$$

For two classes to be separable, their separability index must be greater than one (01), and the greater the separability index, the more easily the two classes can be separated.

Pre-sowing Analysis

In the pre-sowing period (January), plots planned for rice cultivation are indistinguishable from urban areas and other crops. With the exception of the NDVI and NDWI indices, the separability index between rice plots and other vegetation is greater than two (02). However, these two indices, like the NDVI, mix rice fields with bare soil. It was also noted that for band B4 and the EVI, rice-growing areas are mixed with water. This can be explained by the fact that the rice plots are flooded before the planting period (see **Figure 5**).

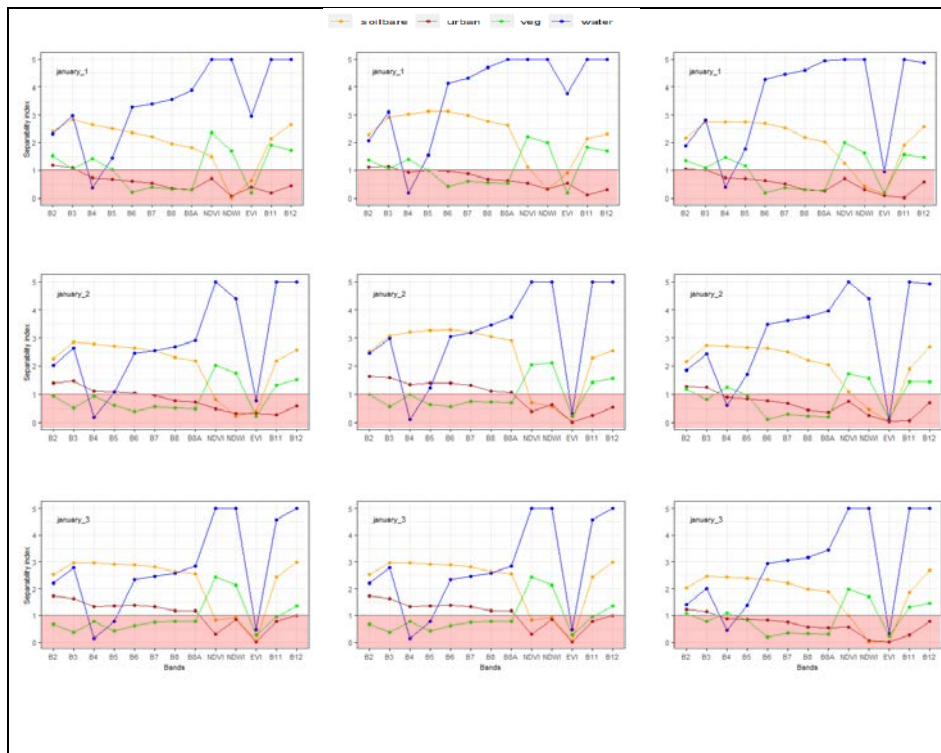


Figure 5. Index of separability between rice and other classes before sowing

Analysis During and After Sowing

- **Block 1:** Sowing date from February 2 to 14 (first and second dekad of February). During the sowing period, the observation was the same as before sowing, i.e., only the NDVI and NDWI indices make it possible to discriminate rice fields from other vegetation. However, from the first dekad after sowing (3rd dekad of February), all the separability indices between rice and other land uses are higher than 01 for bands B6, B7, B8, B8A, and NVDI. This situation continues until the third dekad of March. At the beginning of

April, it was observed that almost all the bands and indices discriminate rice from the other classes, except for band B5, which confuses it with water, and NDVI and EVI, which confuse rice with other vegetation. From the month of May, corresponding to the 7th dekad since the start of sowing in this block, the confusion between rice and other vegetation becomes significant and concerns almost all bands and indices except B2, B3, B4, and B5 (see **Figure 6**).

- **Block 2:** Sowing between February 15 and 24 (2nd and 3rd dekad of February). The curves of the first dekad of February, which belongs to the period before sowing, are identical to those of January. From the third dekad of February, corresponding to the end of sowing for this block, it was noted that there was a positive evolution of rice separability with the separability indices of bands B6, B7, B8, B8A, and the NDVI higher than 01. This situation is maintained throughout March and until the first dekad of April, when all the bands from B2 to B8A and the NDVI discriminate well between rice and the other classes, except for water in band B5. Beyond this period, it is no longer possible to separate rice from other vegetation (see **Figure 6**).

- **Block 3:** Sowing between February 25 and March 5 (between the 3rd dekad of February and the 1st dekad of March). The observation before and during sowing remains the same as that observed on the first two blocks. However, unlike the first two blocks, only the first dekad of April here is favorable for differentiating rice from the other classes which are from bands B6, B7, B8, and B8A. Unfortunately, the lack of data during the third dekad of March due to clouds means that it is not possible to analyze the situation in the second dekad after the end of planting for this block.

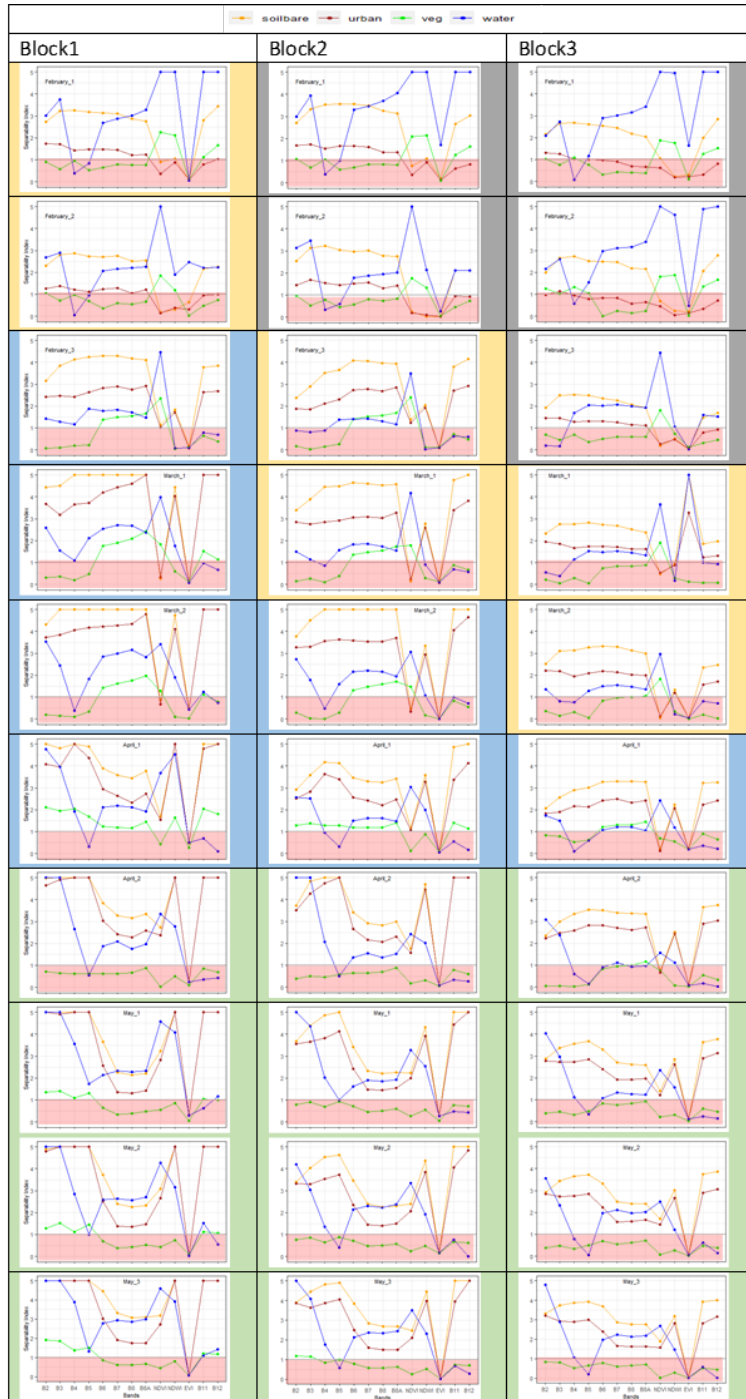


Figure 6. Separability index between rice and other classes before sowing (gray) during sowing (yellow) and after sowing (blue and green). The blue is the optimal period to discriminate.

It can be concluded that during the pre-sowing period (Figure 4), there is no band or index to discriminate rice-growing areas from other land-use classes. Mapping the area sown before the sowing period is therefore not an easy task using Sentinel-2 data. In summary, however, analysis of Figure 5 shows that the second dekad of April is the best time to map rice plots in this area of the Delta. The images acquired during this dekad will be used for classification according to the three selected classifier models (SVM, RF, CART).

Results of the Classification

The comparison of the three Machine Learning methods shows the result with different accuracies for each land use class (**Table 3**). This phase of validation of the 30% of the tested samples helped to obtain the results of precisions and kappa coefficients of each model.

Table 1. Comparing the accuracies of the three methods

Algorithm	Rice (0)	Soil bare (1)	Water (2)	Other Vegetation
CART	91,33	99,22	99,78	98,04
RF	90,88	99,70	99,78	98,81
SVM	81,25	99,76	1	97,90

For the RF method (Azzari & Lobell, 2017; Huang et al., 2017; Parente et al., 2019; Teluguntla et al., 2018; Zhang et al., 2018), it was noted that there were some confusion between the rice class and other vegetation (9100 square meter of rice is considered vegetation). There are also confusions between rice and other classes such as bare soil (3200 square meter of bare soil are classified as rice). There is also confusion between water and vegetation due to the presence of aquatic vegetation in the area. The results show that the RF algorithm gives an overall accuracy of 98% with a kappa index of 97.1% (**Figure 7**).

With the CART method, as with RF, there is confusion between the different classes with 9500 square meter of rice being confused with vegetation and 2200 square meter with bare soil. The CART algorithm gives an overall accuracy of 97.6% and a kappa index of 96.5% (Figure 7). It should be noted that while CART handles missing values well in prediction, random forests, which are unpruned sets of trees, essentially lose this property (Genuer & Poggi, 2017).

For the SVM method, there was a strong confusion between the rice class and other vegetation with 19400 square meter of rice considered as other vegetation.

SVMs are particularly attractive in the field of remote sensing because of their ability to generalize well even with limited samples, which is common in remote sensing applications. However, they also suffer from parameter

assignment problems that can significantly affect the results obtained. The SVM classifier has an overall accuracy of 96.2% and a kappa index of 94.5% (Figure 7).

The classifiers (CART and RF) correctly predict rice at 91.33% and 90.88% respectively (Table 3). Previous studies have suggested that the number of random drill decision trees is generally proportional to the accuracy of the classifier (Rodriguez-Galiano et al., 2012). Foody and Mathur (2006) proposed to focus on mixed pixel training samples rather than the more tedious conventional pure pixel acquisition assuming an SVM classifier. The results of this study suggest that the Random Forest classifier performs as well as SVMs in terms of classification accuracy and training time. The performance of Random Forest improves as the number of trees increases. However, the performance improves when NDVI, NDWI, and EVI bands (Ferrant et al., 2017; Kuenzer & Knauer, 2013; Xiao et al., 2006) are added to the selected sentinel-2 bands for both SVM and RF.

Previous classification results indicate that RF is reasonably suitable for classifying such data, as in some cases (Ferrant et al., 2017; Hong Son & Thai-Nghe, 2019), and it performed better than other classifiers because it has less error risk. The highest confusion was noted between rice and other vegetation in all three methods. This is due to the similarity of rice crops to natural vegetation. It is manifested by very similar reflectance values in relation to their shape. Thus, it is difficult to differentiate between vegetation, sugarcane crops and rice in the north, especially at this time of year.

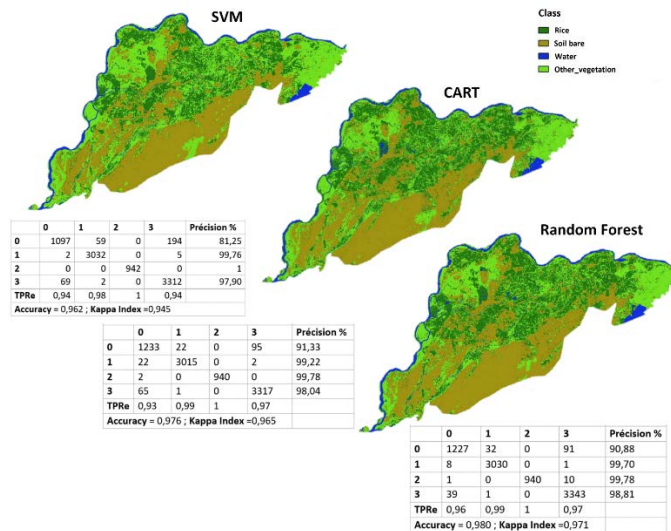


Figure 6. Results classification of the three methods

Analysis of the Classification

To better analyze the results of the classification, the unmixing of the pixels was performed. This method available on the GEE platform consists of calculating the contribution of each class in the value of each pixel (see **Figure 8**).

For the pixels classified as rice, there is a low contribution of bare soil, water, and buildings between 0 and 30% compared to that of vegetation and other crops, which can be as high as 80%. The contribution of rice in the value of pixels classified as other vegetation is quite significant, even exceeding 80%. For pixels with bare soil and water, the contribution of rice is very low, sometimes less than 10%. These results clearly show that the presence of vegetation in the rice fields, on the one hand, and the resemblance between rice and other crops or vegetation at certain periods of the rice crop cycle, on the other hand, have an impact on the quality of the classification. For this reason, deep learning (DeepLearning) could improve the results obtained, although the results obtained remain very satisfactory.

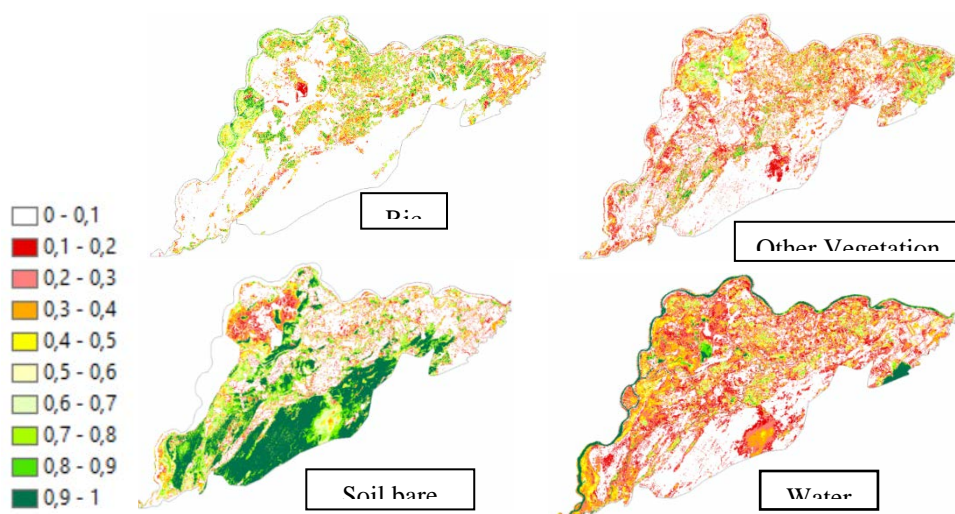


Figure 7. The contribution of each land cover class

Conclusion

The sentinel-2 data performed well with all three algorithms. The CART method gives a good accuracy of rice with a value of 91.33%. This study proposes an approach for rice identification and classification with image processing algorithms and machine learning methods. Adding more classes to recognize and identify rice and comparing these models show that the proposed approach works well for this study. Based on the results of this study, it would be interesting to synergize the sentinel-1 and sentinel-2 data to

improve large-scale rice mapping in the Senegalese context. It should be remembered that the size of rice fields is small compared to the spatial resolution of Sentinel, which may be a limitation to the use of these data. However, a combination of optical-radar and Deep learning methods could help improve current methods of estimating rice area in the Sahelian context.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References:

1. ANDS (2020). Situation économique et sociale régionale de Saint-Louis 2017-2018, ANDS
2. Azzari, G. & Lobell, D.B. (2017). Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. *Remote Sensing of Environment* 202, 64–74.
3. Belder, P., Bouman, B.A.M., Cabangon, R., Guoan, L., Quilang, E.J.P., Yuanhua, L., Spiertz, J.H.J., & Tuong, T.P. (2004). Effect of water-saving irrigation on rice yield and water use in typical lowland conditions in Asia. *Agricultural water management* 65, 193–210.
4. Ben-Hur, A. & Weston, J. (2010). A user's guide to support vector machines, in: *Data Mining Techniques for the Life Sciences*. Springer, pp. 223–239.
5. Boschetti, M., Nutini, F., Manfron, G., Brivio, P.A., & Nelson, A. (2014). Comparative Analysis of Normalised Difference Spectral Indices Derived from MODIS for Detecting Surface Water in Flooded Rice Cropping Systems. *PLoS ONE* 9, e88741. <https://doi.org/10.1371/journal.pone.0088741>
6. Bouman, B.A.M., Humphreys, E., Tuong, T.P., & Barker, R. (2007). Rice and water. *Advances in agronomy* 92, 187–237.
7. Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth. International Group 432, 151–166.
8. Chang, C.-C. & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)* 2, 1–27.
9. Cortes, C. & Vapnik, V. (1995). Support-vector networks *Machine learning* (pp. 237–297), Vol. 20. Boston, MA: Kluwer Academic Publisher.

10. Ferrant, S., Selles, A., Le Page, M., Herrault, P.-A., Pelletier, C., Al-Bitar, A., Mermoz, S., Gascoin, S., Bouvet, A., Saqalli, M., Dewandel, B., Caballero, Y., Ahmed, S., Maréchal, J.-C., & Kerr, Y. (2017). Detection of Irrigated Crops from Sentinel-1 and Sentinel-2 Data to Estimate Seasonal Groundwater Use in South India. *Remote Sensing* 9, 1119. <https://doi.org/10.3390/rs9111119>
11. Foody, G.M. & Mathur, A. (2004). Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment* 93, 107–117. <https://doi.org/10.1016/j.rse.2004.06.017>
12. Foody, G.M. & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote sensing of environment* 103, 179–189.
13. Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* 58, 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
14. Genuer, R. & Poggi, J.-M. (2017). Arbres CART et Forêts aléatoires, Importance et sélection de variables. arXiv:1610.08203 [math, stat].
15. Gumma, M.K., Thenkabail, P.S., Maunahan, A., Islam, S., & Nelson, A. (2014). Mapping seasonal rice cropland extent and area in the high cropping intensity environment of Bangladesh using MODIS 500m data for the year 2010. *ISPRS Journal of Photogrammetry and RemoteSensing* 91,98–113. <https://doi.org/10.1016/j.isprsjprs.2014.02.007>
16. Hong Son, N. & Thai-Nghe, N. (2019). Deep Learning for Rice Quality Classification, in: 2019 International Conference on Advanced Computing and Applications (ACOMP). Presented at the 2019 International Conference on Advanced Computing and Applications (ACOMP), IEEE, NhaTrang, Vietnam, pp.92–96. <https://doi.org/10.1109/ACOMP.2019.00021>
17. Huang, J., Li, Y., Fu, C., Chen, F., Fu, Q., Dai, A., Shinoda, M., Ma, Z., Guo, W., & Li, Z. (2017). Dryland climate change: Recent progress and challenges. *Reviews of Geophysics* 55, 719–778.
18. Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83, 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)

19. Jay L. Maclean & David Charles Dawe (2002). Rice Almanac: Source Book for the Most Important Economic Activity on Earth - Google Livres [WWW Document].
20. Jha, J. & Ragha, L. (2013). Intrusion detection system using support vector machine. *International Journal of Applied Information Systems (IJAIS)* 3, 25–30.
21. Kshirsagar, K.G. & Pandey, S. (1995). Diversity of rice cultivars in a rainfed village in the Orissa state of India. Using diversity: enhancing and maintaining resources on-farm. International Development Research Center (IDRC), Regional Office for South Asia, India. p 55–65.
22. Kuenzer, C. & Knauer, K. (2013). Remote sensing of rice crop areas. *International Journal of Remote Sensing* 34, 2101–2139. <https://doi.org/10.1080/01431161.2012.738946>
23. Lacharme, M. (2001). Données morphologiques et cycle de la plante 22
24. Lancashire, P.D., Bleiholder, H., Boom, T.V.D., Langelüddeke, P., Stauss, R., Weber, E., & Witzzenberger, A. (1991). A uniform decimal code for growth stages of crops and weeds. *Ann Applied Biology* 119, 561–601. <https://doi.org/10.1111/j.1744-7348.1991.tb04895.x>
25. Le Toan, T., Ribbes, F., Li-Fang Wang, Floury, N., Kung-Hau Ding, Jin Au Kong, Fujita, M., & Kurosu, T. (1997). Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. *IEEE Trans. Geosci. Remote Sensing* 35, 41–56. <https://doi.org/10.1109/36.551933>
26. Liu, Y., Wang, Y., & Zhang, J. (2012). New machine learning algorithm: Random forest, in: *International Conference on Information Computing and Applications*. Springer, pp. 246–252.
27. Mendez Del Villar, P., Bauer, J.M., Maiga, A., & Ibrahim, L. (2011). Rice crisis, market developments and food security in West Africa. Ministry of Foreign and European Affairs, West Africa, 61.
28. Michel, P. & Sall, M. Dynamique des paysages et aménagement de la vallée alluviale du Sénégal 21.
29. Nguyen, X.T. (2016). Modélisation de l'impact des rizières et de l'irrigation sur le régime hydrologique de la rivière Cong au Vietnam. (PhD Thesis). Université du Québec, Institut national de la recherche scientifique.
30. Parente, R., Rong, K., Geleilate, J.-M.G., & Misati, E. (2019). Adapting and sustaining operations in weak institutional

- environments: A business ecosystem assessment of a Chinese MNE in Central Africa. *Journal of International Business Studies* 50, 275–291.
31. Radanielina, T., Ramanantsoanirina, A., Raboin, L.-M., & Ahmadi, N. (2013). Determinants of rice varietal diversity in the region of Vakinankaratra (Madagascar). *Cahiers Agricultures* 22, 442–449. <https://doi.org/10.1684/agr.2013.0648>
 32. Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J.P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 67, 93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>
 33. Rouse Jr, J.W., Haas, R.H., Schell, J.A., & Deering, D.W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. Third Earth Resources Technology Satellite-1 Symposium: Volume 1; Technical presentations, section B, SC Freden, EP Mercanti, and MA Becker, Eds., NASA Special Publ. NASA-SP-351-VOL-1-SECT-B, A 20, 309–317.
 34. Schiilkop, P.B., Burgest, C., & Vapnik, V. (1995). Extracting support data for a given task, in: *Proceedings, First International Conference on Knowledge Discovery & Data Mining*. AAAI Press, Menlo Park, CA. pp. 252–257.
 35. Singh, V.P., Tuong, T.P., & Kam, S.P. (2000). Characterising rainfed rice environments: an overview of the biophysical aspects. *Characterising and understanding rainfed environments*. Los Baños, Philippines, IRRI 3–32.
 36. Steinberg, D., Golovnya, M., & Polosukhin, I. (2012). Text Mining Using STMTM, CART®, and TreeNet® from Salford Systems: Analysis of 16,000 iPod Auctions on eBay, in: *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*. Academic Press, pp. 413–416.
 37. Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., & Huete, A. (2018). A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing* 144, 325–340.
 38. Torbick, N., Chowdhury, D., Salas, W., & Qi, J. (2017). Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2. *Remote Sensing* 9, 119. <https://doi.org/10.3390/rs9020119>

39. Touré, M.A. (2018). "Climate variability and ecosystem dynamics in the Senegal River Delta from the 1950 to the 2010."
40. Tuong, T.P., BAM, B., & Mortimer, M. (2005). More Rice, Less Water—Integrated Approaches for Increasing Water Productivity in Irrigated Rice-Based Systems in Asia—. *Plant Production Science* 8, 231–241.
41. Vapnik, V.N. (1999). An overview of statistical learning theory. *IEEE transactions on neural networks* 10, 988–999.
42. Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F., & Ceschia, E. (2017). Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment* 199, 415–426. <https://doi.org/10.1016/j.rse.2017.07.015>
43. Wassmann, R., Hien, N.X., Hoanh, C.T., & Tuong, T.P. (2004). Sea level rise affecting the Vietnamese Mekong Delta: water elevation in the flood season and implications for rice production. *Climatic change* 66, 89–107.
44. Xiao, X., Boles, S., Frohling, S., Li, C., Babu, J.Y., Salas, W., & Moore, B. (2006). Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment* 100, 95–113. <https://doi.org/10.1016/j.rse.2005.10.004>
45. Yang, C.-M., Cheng, C.-H., & Chen, R.K. (2007). Changes in spectral characteristics of rice canopy infested with brown planthopper and leafhopper. *Crop science* 47, 329–335.
46. Yoshida, S., Forno, D.A., Cock, J.H. (1971). Laboratory manual for physiological studies of rice. *Laboratory manual for physiological studies of rice*.
47. Zhang, H., Li, Y., Zhu, J.-K. (2018). Developing naturally stress-resistant crops for a sustainable agriculture. *Nature plants* 4, 989–996.