Journal of Geographic Information System, 2026, 18(1), 12-27
https://www.scirp.org/journal/jgis
ISSN Online: 2151-1969
ISSN Print: 2151-1950

Efficiency of Classification Algorithms in Monitoring Land Use through Remote Sensing in Sissili Province of Burkina Faso

Oumar Kaboré, Wennepinguere Virginie Marie Yaméogo

Natural Resources Management Department, Institute of Environment and Agricultural Research (INERA), Ouagadougou, Burkina Faso
Email: oumarkabore@hotmail.com

How to cite this paper: Kaboré, O. and Yaméogo, W.V.M. (2026) Efficiency of Classification Algorithms in Monitoring Land Use through Remote Sensing in Sissili Province of Burkina Faso. *Journal of Geographic Information System*, **18**, 12-27.
https://doi.org/10.4236/jgis.2026.181002

Received: November 26, 2025
Accepted: January 27, 2026
Published: January 30, 2026

Abstract

Monitoring of natural resources is a major challenge that remote sensing tools help to facilitate. The Sissili province in Burkina Faso is a territory that includes significant areas dedicated for the preservation of forest resources. The development of satellite image processing tools offers an opportunity for better monitoring of these resources. The aim of this research is to evaluate the performance of classification algorithms in order to determine which one is most suitable for assessing and monitoring land use in the Sissili province. The methodology used is based on the comparative application of three classification algorithms: Maximum Likelihood, Random Forests, and Support Vector Machine. These algorithms were tested on land use units in the Municipality of Sissili, to determine their respective performance. Performance measures were based on the production of the confusion matrix for each algorithm, the calculation of overall accuracy, and the calculation of Kappa coefficient for the three algorithms. The results show that the Random Forest algorithm is the most effective, with a Kappa coefficient of 0.92 and an overall accuracy of 95.14%. This algorithm is followed by SVM with a Kappa coefficient of 0.73 and a maximum overall accuracy of 90.37%. The least effective algorithm for classifying land use units is Maximum Likelihood, with a Kappa coefficient of 0.36 and an overall accuracy of 52.95%. These results clearly demonstrate the superior effectiveness of machine learning algorithms, specifically Random Forest, in classifying land use units in the Sissili province.

Keywords

Algorithm, Classification, Land Use, Sissili

1. Introduction

Land resources are subject to various pressures, both natural and human, on a global scale [1] [2]. In this sense, the Sahel region, particularly West Africa, is facing environmental upheavals, and the analysis of land use dynamics allows us to understand the resulting territorial changes [3]-[5]. This spatial analysis is supported by remote sensing, an appropriate means for monitoring and managing natural resources [6]. Indeed, it contributes to the analysis of the Earth's surface, a fundamental necessity in order to better appreciate the changes occurring there [7] [8]. This ever-growing dynamic, along with the increasing interest in using increasingly large datasets, has motivated the emergence of new classification methods [9]. Indeed, the development of certain algorithms based on artificial intelligence has greatly improved methods for processing spatial data and offers new perspectives for planning and resource management through various applications [10]-[12]. In recent decades, particularly since the early 2000s, classification algorithms have undergone rapid evolution with the work of [13], gradually tending to supplant classical models such as Maximum Likelihood, which is still used in some work. Among the classification algorithms that have emerged and are based on artificial intelligence, it is worth noting, for example, ordinary machine learning algorithms such as Random Forests and Support Vector Machines which are relevant to the present work [14]. There are also deep learning algorithms based on artificial neural networks. This article focused on machine learning models and a classical model based on Maximum Likelihood, in search of the best model for monitoring land use in the Sissili province of Burkina Faso. The use of appropriate algorithms is justified because accuracy in thematic mapping production is one of the objectives within the framework of landscape analysis. In Burkina Faso, numerous studies have been carried out on land use and land cover classification algorithms such as Maximum Likelihood in particular [15]-[19]. The Sissili province is home to numerous forestry sites that serve as areas for the conservation of vegetation and biodiversity. Monitoring these areas requires appropriate tools and methods. Remote sensing, through classification algorithms whose effectiveness is increasingly proven, makes significant contributions to the management of natural resources. However, the effectiveness of algorithms is determined and strongly influenced by the context in which they are applied. This is why choosing a suitable algorithm must be considered in the monitoring of natural resources. It was necessary to compare a conventional algorithm with two types of machine learning algorithms in order to guide methodological choices in the case of land use monitoring in the Sissili Province.

2. Materials and Methods

2.1. Study Area

The study area is the Sissili province located in the south of Burkina Faso, between latitudes 2˚48'27" and 1˚24'49" west, and longitudes 10˚58'36" and 11˚55'27" north. It is bordered to the east by the provinces of Nahouri and Ziro, to the west by those of Sanguié, Balé, and Ioba. To the south, it is bordered by the territory of Ghana (Figure 1). This province straddles two climatic zones, namely the Sudano-Sahelian climate zone (600 to 900 mm/year on average) and the Sudanian climate zone (more than 900 mm/year on average). The vegetation consists mainly of wooded savanna, shrubland, grassland, as well as riparian formations and gallery forests that border waterways. Significant protected areas occupy this province, and their management requires rigorous monitoring over time and space. Agriculture and livestock farming are the main activities that occupy most of the population and contribute to creating a certain landscape dynamic.

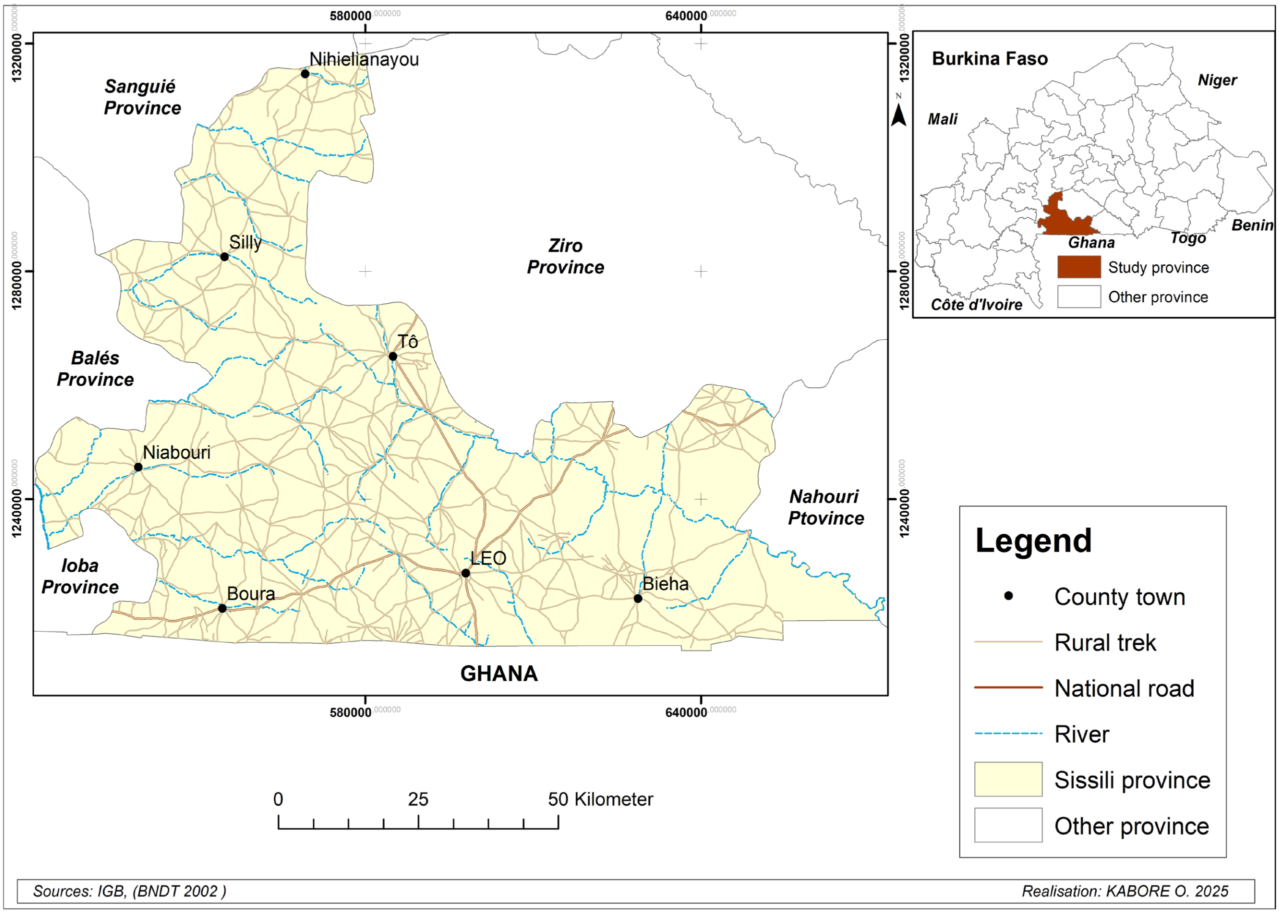


Figure 1. Location of the study area.

2.2. Characteristics of the Remote Sensing Images

The image scene used to analyze the accuracy of classification algorithms comes from the Landsat 9 satellite, and the image was captured using OLI and TIRS sensors. It is scene 195_052 with a resolution of 30 m × 30 m, which is sufficient to map land use at the scale of the province of Sissili. This image scene underwent pre-processing by the supplier, including geometric and radiometric corrections. The quality of these treatments at "L2SP" level is considered good (Table 1).

Table 1. Characteristics of the remote sensing image.

Sensor and satellite type	Image reference	Resolution	Acquisition date	Processing level
OLI-TIRS de Landsat 9	195_052	30 m × 30 m	04/12/2024	L2SP

2.3. Image Scene Preprocessing

The radiometric corrections performed by the supplier allowed obtaining ground reflectance values, but these values cannot be used directly for processing within the scope of this research. Recalibration is therefore necessary to obtain pixel values of the image ranging between 0 and 1. A tool was then used in QGIS to force so that all bands, including band 6, could fall within this range of value. Following this processing, a color composition was applied by combining bands 5, 4, and 3. This color composition was applied to the entire scene, followed by an extraction operation using a mask to restrict the image to the boundaries of the Sissili province.

For choice of classification algorithms among the machine learning algorithms, the Random Forest model and the Support Vector Model (SVM) were compared. The RF algorithm is a collection of decision trees that form a neural network. These trees are obtained from a sample of the entire training dataset. Each tree in its structure is composed of nodes containing input data, so-called hidden nodes that receive information in the form of signals from preceding nodes, and that measure the weight of each piece of information received, then perform a weighted linear combination of all this information before making a decision regarding the class to assign. At the end of the process, an arbitration is made, and the designated class is the one that would have received the highest number of votes. As for the SVM algorithm, it is a machine learning algorithm that allows the generation of a hyperplane to delimit more or less homogeneous sets of entities. The hyperplane is obtained from an equation whose nature varies depending on the homogeneity or heterogeneity of the terrain, as the support vector points are not always perfectly separated. Indeed, in the real world, elements often do not show a clear boundary. In this case, a so-called slack variable is calculated to allow a less rigid margin of the hyperplane [20]. The Maximum Likelihood algorithm is one of the classical methods and has also been included in the comparison. It is based on a probabilistic method that allows a pixel to be assigned to a class. A calculation of the probability of the pixel belonging to a class is then carried out, and the pixel is assigned to the class with the highest probability [21].

2.4. Supervised Image Classification According to Three Algorithms

The image classification was carried out according to the three chosen algorithms in order to allow for the comparison of their performance in reproducing land use patterns as close as possible to reality. Classification using the Maximum Likelihood algorithm was carried out in ArcGIS and required the use of the image's color composition as well as preselected training sites. In total, 241 training sites were chosen and recorded in a vector file in *.SHP format. Furthermore, classification using the Random Forest algorithm was applied in QGIS, using training data. The layer resulting from the satellite image's color composition was used, along with the layer containing the training sites. These layers of data first allowed the creation of the training model. The creation of this training model required to configure certain parameters in particular the field in the training sites layer table that contains the numerical values corresponding to each land cover unit, the type of algorithm used, the number of classes, the number of trees to generate, and many other parameters. Regarding the parameters used specifically for Random Forest, the number of variables used (*mtry*) was 10, and the number of trees to be generated (*ntree*) was set to 100. This model allows obtaining accuracy parameters that make it possible, if necessary, to adjust the input data to improve classification accuracy. The model was saved in *.txt format and was then used to train the classification. As for the support vector algorithm, the same steps as for random forests are necessary (using the same training data, creating a training model, training the classification). However, differences exist regarding the parameter settings. Indeed, in the case of the SVM algorithm, a Gaussian radial basis function was used as the SVM kernel type, and the Nu support vector classification option was chosen as the SVM model type.

To avoid biases in data analysis, a second sample of 100 points, different from the one used for training the classification, was integrated into the evaluation of classification accuracy using the semi-automatic classification tool in QGIS.

2.5. Methods for Evaluating Classification Accuracy

Classification accuracy was first evaluated through the production of the confusion matrix, which was generated in QGIS, using the color composite as well as the vector layer of training sites. The latter was then rasterized and cross compared with that of the colored composition considered as the reference layer. The comparison is then made pixel by pixel. At the end of this operation, a confusion matrix is generated in various formats, thereby allowing the calculation of classification errors for each class, as well as the overall classification accuracy and the Kappa coefficient. These data were generated both for the Maximum Likelihood algorithm and for automatic training algorithms such as Random Forests and Support Vector Machines. The values obtained were then compared to assess the performance of each.

3. Results and Analyses

3.1. Mapping of Land Use Units According to the Three Chosen Algorithms

Depending on their respective performances, the land use unit data vary from one algorithm to another, for the same reference image, the same training sites, the same territory (**Figure 2**).

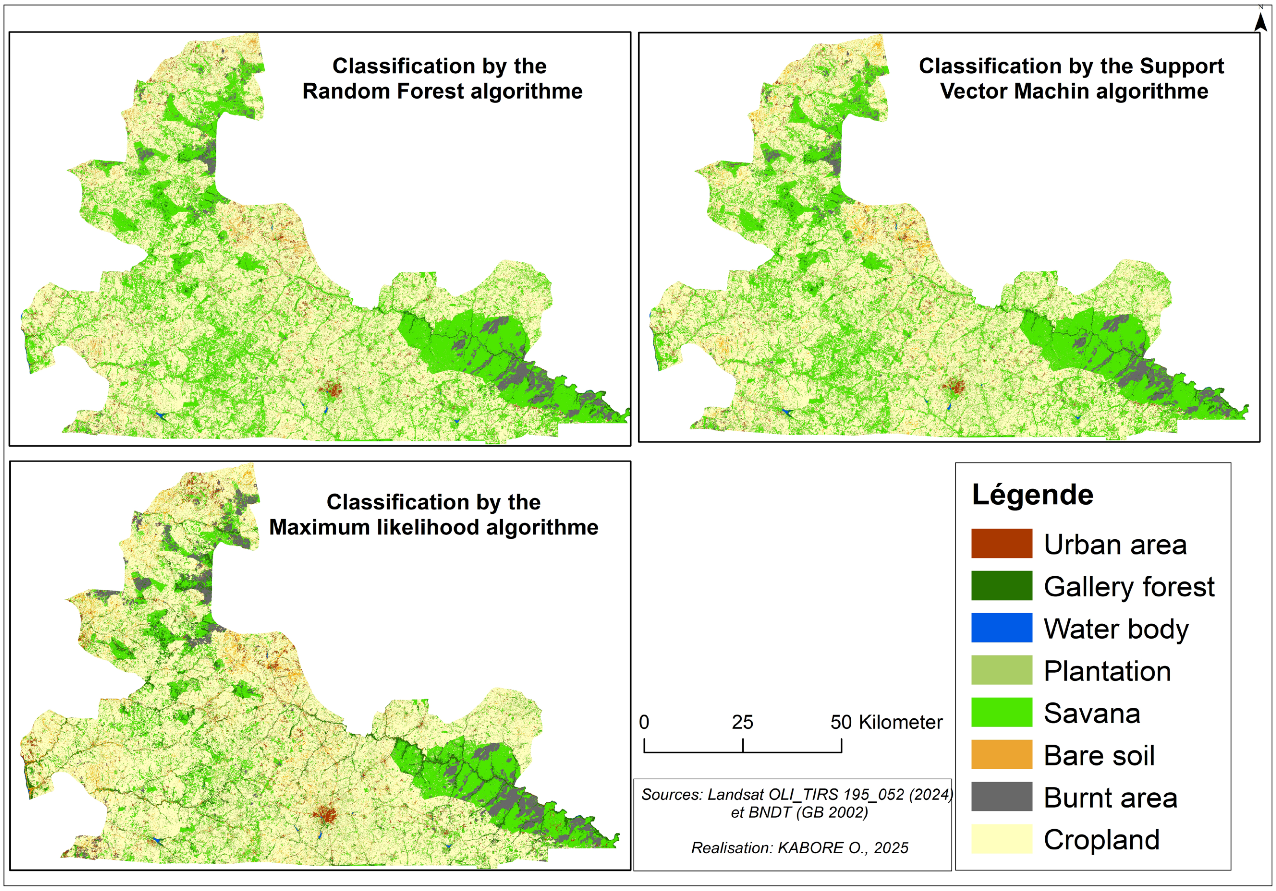


Figure 2. Land use units according to classification algorithms.

According to the classification results from the random forests algorithm, crop areas occupy the largest part of the provincial territory, covering 56.2%, compared to 2.7% for burned areas, 30.3% for savanna, which is the second most extensive unit, 1.14% for gallery forest, 0.07% for water bodies, 1.8% for bare soils, 5.2% for plantations, and 2.2% for settlements (Table 2).

Table 2. Proportions of land use units based on classification algorithms.

Unit	Land use rate using RF (%)	Land use rate using SVM (%)	Land use rate using ML (%)
Cropland	56.28	56.16	60.93
Burnt area	2.80	2.83	4.25
Savanna	30.30	30.50	17.78
Gallery Forest	1.14	1.16	4.95
Water body	0.07	0.07	0.09
Bare soil	1.86	1.87	3.91
Plantation	5.25	5.17	5.44
Urban area	2.30	2.24	2.66

RF: Random Forests, **SVM**: Support Vector Machine, **ML**: Maximum Likelihood.

In the case of the Support Vector Machines algorithm, crop areas cover 56.16% of the territory, burned areas occupy 2.83%, savanna ranks second at 30.50%, gallery forest represents 1.16%, water bodies 0.07%, bare soils 1.87%, plantations 5.17%, and settlements cover 2.24%.

Regarding the Maximum Likelihood algorithm, crop areas occupy 60.93%. % of the province's area, which represents the largest occupied surface, while burned areas cover 4.24% of the surfaces, savanna 17.7%, gallery forest 4.94%, water bodies 0.08%, bare soils 3.91%, plantations account for 5.44%, and 2.65% of the territory is occupied by urban settlements.

3.2. Evaluation of the Accuracy of Classification Algorithms

The accuracy of the classification algorithms was measured using the confusion matrix data, which allowed for the calculation of overall accuracy as well as the Kappa coefficient.

Table 3 represents the confusion matrix produced by the Random Forests algorithm. The values along the diagonal indicate correctly classified pixels. The greatest confusion was observed between savanna and plantations, which are both units occupied by woody plants even though they comprise different species, leading to significant confusion between these classes.

Table 3. Confusion matrix for the random forests algorithm.

Classes	Cl	Ba	Sa	Gf	Wb	Bs	Pla	Ua	Total
CL	20,234	32	806	46	0	63	145	320	21,646
Ba	0	5814	30	10	0	0	0	0	5854
Sa	254	297	40,250	1391	0	4	24	65	42,285
Gf	0	136	795	596	0	0	0	0	1527
Wb	0	0	0	0	124	0	0	0	124
Bs	839	0	0	2	0	268	0	100	1209
Pla	846	16	1173	100	0	0	1268	5	3408
Ua	30	31	6	7	0	13	0	1737	1824
Total	22,203	6326	43,060	2152	124	348	1437	2227	77,877

Cl: Cropland, **Ba**: Burnt area, **Sa**: Savanna, **Gf:** Gallery forest, **Wb**: Water body, **Bs**: Bare soil, **Pla**: Plantation, **Ua**: Urban area.

Table 4 shows the confusion matrix obtained by the Support Vector Machines algorithm. In this table, the values on the diagonal indicate correctly classified pixels. The most significant confusion was observed between savanna and plantations, as was the case with the Random Forest algorithm. This confirms the confusion between these two units, likely due to their proximity, since they are all areas occupied by woody plants.

Table 4. Confusion matrix for the support vector machine algorithm.

Classes	Cl	Ba	Sa	Gf	Wb	Bs	Pla	Ua	Total
CL	20,214	7	1089	82	0	59	119	273	21,843
Ba	0	6028	39	3	2	0	0	8	6080
Sa	115	132	38,215	1182	0	4	15	66	39,729
Gf	0	77	2704	796	0	0	0	0	3577
Wb	0	0	0	0	122	0	0	0	122
Bs	1071	0	0	0	0	259	0	106	1436
Pla	772	0	972	82	0	0	1303	4	3133
Ua	31	82	41	7	0	26	0	1770	1957
Total	22,203	6326	43,060	2152	124	348	1437	2227	77,877

Cl: Cropland, **Ba**: Burnt area, **Sa**: Savanna, **Gf:** Gallery forest, **Wb**: Water body, **Bs**: Bare soil, **Pla**: Plantation, **Ua**: Urban area.

Table 5 represents the confusion matrix produced following the application of the Maximum Likelihood algorithm. As with the other tables, the values on the diagonal indicate correctly classified pixels. The other cells represent various cases of class confusion. The most significant confusion concerns the savanna and riparian formations, unlike the other two algorithms. This confusion could stem from the fact that both entities are occupied by woody plants that, in some locations, are very similar from a radiometric point of view. However, the magnitude of the confusions between entities shows a closer similarity between the two automatic training algorithms compared to the Maximum Likelihood algorithm, which yielded dissimilar results.

Table 5. Confusion matrix using the Maximum Likelihood algorithm.

Classes	Cl	Ba	Sa	Gf	Wb	Bs	Pla	Ua	Total
Cult	20,339	9	1279	137	0	11	54	112	21,941
ZB	0	6226	125	5	0	0	0	1	6357
Sa	100	22	38,072	338	0	0	12	0	38,544
Fg	2	55	2971	1493	0	0	0	2	4523
Pe	0	0	0	0	124	0	0	0	124
Sn	787	0	13	64	0	320	3	63	1250
Pla	924	0	567	103	0	1	1367	3	2965
Agl	51	14	33	12	0	16	1	2046	2173
Total	22,203	6326	43,060	2152	124	348	1437	2227	77,877

Cl: Cropland; **Ba**: Burnt area; **Sa**: Savanna; **Gf:** Gallery forest; **Wb**: Water body; **Bs**: Bare soil; **Pla**: Plantation; **Ua**: Urban area.

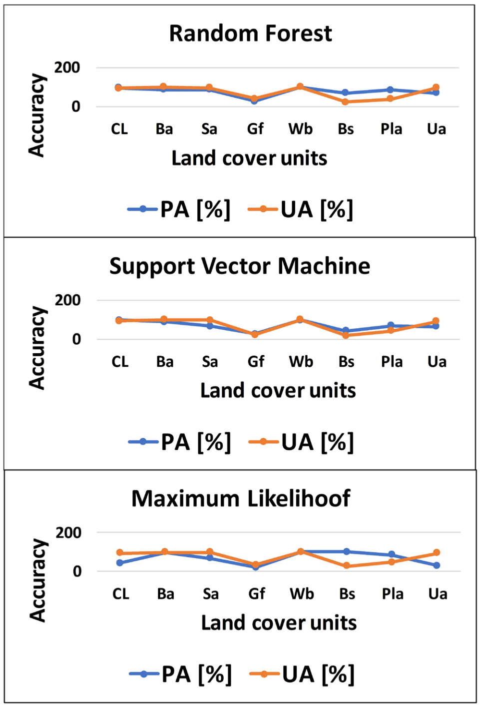
Regarding the types of errors considered by the confusion matrix, there are commission errors that determine user accuracy, and omission errors that determine producer accuracy. Commission error occurs when, in a classified image, a pixel is identified as belonging to class A while the reference image shows it belongs to class B. Class A has been overestimated. On the other hand, an omission error occurs when, on a reference image, a pixel is identified as belonging to class A, while in the classified image, it is indicated that this pixel belongs to class B. The pixels of class B were not sufficiently represented, so they were overlooked. In Table 6, the Random Forest algorithm shows producer accuracies ranging from 25.71% to 100%, corresponding respectively to the gallery forest and Water body units whereas user accuracies range from 22.16% to 100%. The Kappa coefficient, on the other hand, ranges from 0.21 to 1. Regarding the Support Vector Machine algorithm, producer accuracies range from 25.46% to 98.28%, while user accuracies vary between 18.03% and 100%. The Kappa coefficient ranges from 0.17 to 1. Finally, the Maximum Likelihood algorithm shows values ranging from 27.49% to 100% for producer accuracies, whereas user accuracies range from 25.60% to 100%. The Kappa coefficient is between 0.13 and 1. Overall, the Random Forest algorithm and Maximum Likelihood are the highest producer accuracy at 100%, while the Maximum Likelihood has the weakest producer accuracy at 25.46%.

Table 6. Classification accuracy for each land use unit based on the algorithm used.

Random Forest								
Measurement Type	Cl	Ba	Sa	Gf	Wb	Bs	Pla	Ua
PA [%]	94.92	85.75	86.45	25.71	100.00	69.26	83.20	67.75
UA [%]	93.47	99.31	95.18	39.03	100.00	22.16	37.20	95.23
Kappa hat	0.91	0.99	0.94	0.38	1.00	0.21	0.36	0.95
Support vector machine								
Measurement Type	Cl	Ba	Sa	Gf	Wb	Bs	Pla	Ua
PA [%]	98.22	88.83	66.09	25.46	98.28	41.85	67.43	63.44
UA [%]	92.54	99.14	96.18	22.25	100.00	18.03	41.58	90.44
Kappa hat	0.68	0.99	0.95	0.21	1.00	0.17	0.40	0.90
Maximum Likelihood								
Measurement Type	Cl	Ba	Sa	Gf	Wb	Bs	Pla	Ua
PA [%]	41.33	97.36	65.46	18.67	100.00	99.83	84.63	27.49
UA [%]	92.69	97.93	98.77	33.00	100.00	25.60	46.10	94.15
Kappa hat	0.80	0.97	0.98	0.30	1.00	0.13	0.45	0.93

Pa: Producer's accuracy, **Ua**: User's accuracy, **Cl**: Cropland, **Ba**: Burnt area, **Sa**: Savanna, **Gf**: Gallery forest, **Wb**: Water body, **Bs**: Bare soil, **Pla**: Plantation, **Ua**: Urban area.

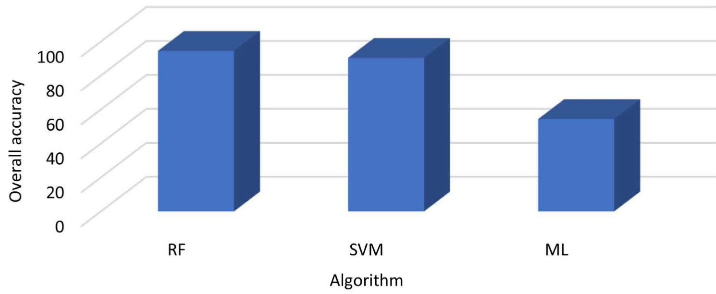
Figure 3 allows us to observe the different variations in accuracies according to producers and user accuracies depending on the different land cover units and the type of algorithm used. The curves show that the values corresponding to machine learning algorithms have more similar trends than those of the Maximum Likelihood algorithm, except for the case of user accuracies where all the curves are very similar. Indeed, regarding the three algorithms (random forest, support vector machine, and maximum likelihood), the highest user accuracy values concern crop areas, burned areas, savannas, and water bodies. On the other hand, the lowest values concern gallery forests, bare soils, and water bodies. The values concerning urban areas occupy an intermediate position. Regarding the accuracy values according to the producer, they show a similar variation only for the two machine learning algorithms. For these two methods, the highest values are recorded for croplands, savannas, water bodies, plantations, and urban areas. On the other hand, the lowest values are mainly recorded for gallery forests and bare soils. However, concerning Maximum Likelihood, the curve of accuracy values according to the owner has a completely different appearance from those of the other two algorithms. The sometimes contrasting evolution of these different curves also shows that accuracies in the classification of land cover units demonstrate some variability depending on the classification algorithms. In general, however, it appears that machine learning algorithms have values that are closer together than those maximum likelihoods.



Pa: Producer's accuracy, **Ua**: User's accuracy, **Cl**: Cropland, **Ba**: Burnt area, **Sa**: Savanna, **Gf:** Gallery forest, **Wb**: Water body, **Bs**: Bare soil, **Pla**: Plantation, **Ua**: Urban area.

Figure 3. Accuracy trends according to producer and user by classification algorithms and based on land use units.

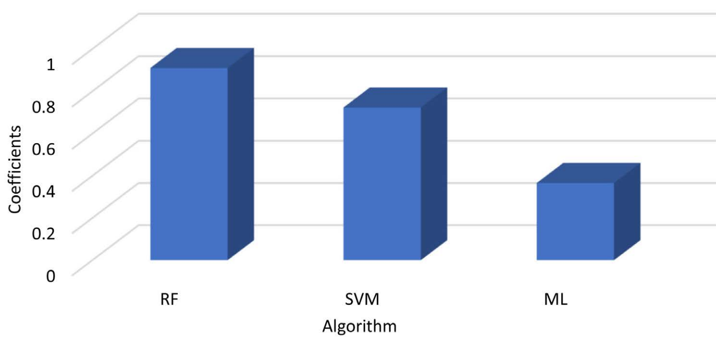
Overall accuracies were also calculated for each type of algorithm. Figure 4 shows this variability. Indeed, the highest overall accuracy is that of the Random Forest algorithm with a rate of 95.14%, followed by the Support Vector Machine algorithm with an accuracy of 90.37%. The Maximum Likelihood algorithm has the lowest overall accuracy, at 52.95%. Therefore, Random Forest is the most efficient classifier in terms of overall accuracy compared to the other algorithms.



RF: Random Forests, **SVM**: Support Vector Machine, **ML**: Maximum Likelihood.

Figure 4. Variation of overall accuracy depending on the algorithm used.

The Kappa coefficient was also calculated for the three algorithms. Figure 5 illustrates the variability of this coefficient depending on the classification methods. Indeed, the Random Forest algorithm has a coefficient of 0.92, representing the highest value compared to the coefficients in the other two cases. The second highest coefficient has a value of 0.73 and was assigned to the Support Vector Machine method. The algorithm with the lowest Kappa coefficient is Maximum Likelihood, with a value of 0.36. These values show that the Random Forest algorithm performs best in classifying land use units according to the Kappa coefficient values. This ranking of the algorithms is the same as that obtained by the overall classification accuracy, which confirms the performance order of the different algorithm methods. However, observation of the figures shows that the differences in values between the algorithms are not the same depending on the type of measurement. In fact, the differences between the Kappa coefficients are greater between the algorithms compared to the differences in values regarding overall accuracy.



RF: Random Forests, **SVM**: Support Vector Machine, **ML**: Maximum Likelihood.

Figure 5. Variation of the Kappa coefficient according to the type of algorithm used.

4. Discussion

Numerous studies have demonstrated the importance of using remote sensing tools through the processing of satellite images for the monitoring and management of natural resources [22] [23].

However, the extraction of geospatial information faces a problem of algorithm selection, for example in the case of classifying landscape units. This is all the more significant because the performance of classification algorithms is often influenced by context. Many classification methods exist, making it not easy to find the one suitable for a specific study. Considerable efforts must be acknowledged for practitioners and researchers in the search for even more accurate classification methods [24]-[26]. Support Vector Machines (SVM) are a supervised learning classification method, based on defining a margin between entities such that this margin is the distance to the boundary of the nearest points. Its advantage is that it allows achieving good results even in situations where the number of training samples is small. The work carried out on land use dynamics and peasant perceptions in the southwest of Niger, particularly in the Kori Ouallam watershed, has also used this algorithm because of its robustness. The Random Forest algorithm provides good classification results. Used in Canada for comparison with Robust Classification Methods, the results indeed showed an overall accuracy level of 99% for this algorithm [27]. It is used as a non-parametric model, meaning it is not based on an initial assumption of the statistical distribution of the data. These two algorithms, are grouped under automated training classification methods. The Maximum Likelihood algorithm is a classical method still used among more recent methods known as automatic training. It has been recently used in studies in the Middle Casamance basin, in the agropastoral area of the Korahane commune in Chad, and many other studies in Africa and elsewhere. It is a probabilistic method whereby a pixel is assigned to a class if that class has the highest probability. Various studies have shown the potential performance of these algorithms in different contexts using different methods. In the commune of Cocody the Maximum Likelihood algorithm was able to achieve an accuracy of 91.27% with using a GPS receiver to collect ground truth data. The accuracy measurements of the algorithm's classification, just as was the case in the present research, often focused on overall accuracy as well as the Kappa coefficient. This is the case with the work, who, in the context of analyzing the thematic accuracy of GEOBIA classification, were able to achieve overall accuracy and a Kappa index ranging respectively between 86% and 96% for a set of classifiers including Random Forests.

In the context of mapping a protected area in northwestern Morocco using automatic training algorithms, accuracy assessment values for these classifiers were applied to two types of images. According to the results obtained, the Random Forest algorithm allowed for a Kappa coefficient between 0.45 and 0.92 depending on the type of images, and an accuracy overall ranging between 53.03% and 93.06%. For the Support Vector Machine method, the Kappa coefficient is between 0.39 and 0.83, and the overall accuracy index between 47.8% and 86.08%.

These results clearly show that while the effectiveness of an algorithm is linked to the context, notably the local characteristics of the environment, it is also necessary to consider that their performance is also related to the choice of images used. Furthermore, these results also confirm those found in the context of the present study, which was able to show better performance of the Random Forest algorithm compared to the Support Vector Machine method. Furthermore, the principle of neural networks is based on an objective approach. Indeed, each decision tree containing nodes is generated from a sample drawn from the training dataset. The randomness in the choice of the subset, along with the large number of trees, determines the robustness of this algorithm, allowing it to avoid bias in generalization and overlearning. Random forest is also capable of integrating a wide variety of explanatory variables, which may especially appear in areas where the landscape shows a certain diversity. The comparison of different machine learning algorithms for soil classification in the Sine Saloum in Senegal revealed the ability of Random Forests to capture complex relationships between the studied variables [28]. Studies have also shown this algorithm's ability to manage nonlinear relationships and complex interrelationships in the analysis of ecological systems [29]. Moreover, in satellite image processing, the maximum likelihood algorithm is based on the assumption that the reflectance values of each defined class follow a normal distribution. However, this requires that a large number of ROIs be defined for each class [30]. This algorithm aims to find the parameter that best approximates the value of population variable from a sample. The absence of the data normal distribution could lead to this algorithm, poor performance compared to others [31].

One of the shortcomings of this research is that it is based on a single image, whereas the RF algorithm has the capacity to handle a large number of images in contexts of significant landscape diversity, which could have improved the quality of the results, even if increasing the number of images does not necessarily imply better accuracy of the results. Moreover, the images obtained at different times and dates could have improved the scientific quality of the results.

In Burkina Faso, various studies have been carried out mainly based on the Maximum Likelihood algorithm. The results obtained in the present study allowed showing the performance differences between the Random Forest, Support Vector Machine, and algorithms Maximum likelihood. This comparison highlighted the greater efficiency of the Random Forest method in classifying land use units in the Sissili province of Burkina Faso. Indeed, for this high-performance algorithm, the results in terms of overall accuracy and Kappa coefficient are 95.14% and 0.92, respectively.

5. Conclusion

Mapping landscapes on the Earth's surface is an important means of monitoring and managing natural resources. In Burkina Faso, the Sissili province is home to numerous protected areas, notably forest management sites, whose sustainability depends on proper monitoring of their development. The processing of remote sensing products can contribute significantly to this monitoring. However, extracting information contained in remote sensing images involves classification, which requires using various methods. The choice of a classification algorithm is sometimes delicate because its effectiveness is influenced by local context determined by particular biophysical characteristics, which often shape the landscape structure. In the context of this study, classification algorithms such as Random Forest, Support Vector Machines, and the classical algorithm based on the Maximum Likelihood principle were used. Likelihood-based methods were tested and compared to determine the classifier best suited for mapping and monitoring natural resources in the Sissili province using Landsat OLI-TIRS type images. The results show that automatic training algorithms provide better results than the Maximum Likelihood method. However, among the automatic training algorithms, the random forest method is the most effective. These results could inspire future work aimed at using remote sensing products for monitoring natural resources in the Sissili province and similar contexts.

Conflicts of Interest

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

References

[1] Chotte, J.L., Dangles, O. and Sabrié, M.L. (2023) Atteindre un monde neutre en termes de dégradation des terres. Sciences De la Durabilité: Comprendre, coconstruire, transformer: Réflexion collective, 2, 104-107.

[2] Blanfort, V., Demenois, J. and Hrabanski, M. (2025) L'agriculture et les systèmes alimentaires du monde face au changement climatique. Enjeux pour les Suds, Versailles, éditions Quæ, 416 p.

[3] Noma Adamou, S., Daoudi, L. and Abdourhamane Touré, A. (2024) Dynamique d'occupation des sols et perception paysanne au Sud-ouest du Niger: Cas du bassin versant du Kori Ouallam. *European Scientific Journal*, *ESJ*, **20**, 163. https://doi.org/10.19044/esj.2024.v20n35p163

[4] Gomis, C. (2025) Dynamiques spatio-temporelles des formes d'occupation du sol dans le bassin de la Moyenne Casamance au prisme des changements climatiques. *African Scientific Journal*, **3**, 1178-1200.

[5] Souleymane, A.D., Aboubacar, A., Nassirou, A.M. and Garba, I. (2025) Dynamique spatio-temporelle d'occupation du sol dans la zone agropastorale de la commune de Korahane, centre Nord de la république du Niger. *European Scientific Journal*, *ESJ*, **21**, 111. https://doi.org/10.19044/esj.2025.v21n14p111

[6] Acharki, S., Frison, P.L., Amharref, M., Khoj, H. and Bernoussi, S. (2021) Complémentarité des images optiques Sentinel-2 avec les images radar Sentinel-1 et Alos-Palsar-2 pour la cartographie de la couverture végétale: Application à une aire protégée et ses environs au Nord-Ouest du Maroc via trois algorithmes d'apprentissage automatique. *Revue Française de Photogrammétrie et de Télédétection*, **223**, 143-158. https://doi.org/10.52638/rfpt.2021.599

[7] Messner, F., Corbonnois, J. and Tchitouo Ntenzou, F.S. (2018) Analyse de l'incertitude et de la précision thématique de classifications GEOBIA d'une image

WorldView-2. *Revue Française de Photogrammétrie et de Télédétection*, No. 216, 19-37. https://doi.org/10.52638/rfpt.2018.310

[8] Macarringue, L.S., Bolfe, É.L. and Pereira, P.R.M. (2022) Developments in Land Use and Land Cover Classification Techniques in Remote Sensing: A Review. *Journal of Geographic Information System*, **14**, 1-28. https://doi.org/10.4236/jgis.2022.141001

[9] Scorne, E. (2015) Learning with Random Forests. Statistics [math.ST]. Université Pierre et Marie Curie (Paris 6), 165 p. https://tel.archives-ouvertes.fr/tel-01250221

[10] Akoguhi, P.N., Dibi, H.N., Godo, M.H., Adja, G.M. and Kouamé, F.K. (2022) Évaluation des méthodes de classifications dirigées (spectrale et orientée objet) sur les images satellitaires à THRS: Cas de la cartographie du tissu urbain de la commune de Cocody et d'Attécoubé (Abidjan, Côte d'Ivoire). *VertigO*, **22**, 1-32.

[11] Kipterer, J.K., Boitt, M.K. and Mundia, C.N. (2025) Comparative Analysis of Machine Learning Algorithms for Optimal Land Use and Land Cover Classification: Guiding Method Selection for Resource-Limited Settings in Tiaty, Baringo County, Kenya. *Journal of Geoscience and Environment Protection*, **13**, 393-414. https://doi.org/10.4236/gep.2025.134021

[12] Hosseini, L. and Shaghaghi Kandovan, R. (2017) Hyperspectral Image Classification Based on Hierarchical SVM Algorithm for Improving Overall Accuracy. *Advances in Remote Sensing*, **6**, 66-75. https://doi.org/10.4236/ars.2017.61005

[13] Beiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. https://doi.org/10.1023/A:1010933404324

[14] Chowdhury, M.S. (2024) Comparison of Accuracy and Reliability of Random Forest, Support Vector Machine, Artificial Neural Network and Maximum Likelihood Method in Land Use/Cover Classification of Urban Setting. *Environmental Challenges*, **14**, Article ID: 100800. https://doi.org/10.1016/j.envc.2023.100800

[15] Tapsoba, A., Kokou, K. and Sanou, Z. (2023) Dynamique spatiale d'occupation du sol à la périphérie du barrage de Bagré au Burkina Faso. *Revue Ivoirienne des Sciences et Technologie*, **41**, 66-82. http://www.revist.ci

[16] Tinto, B., Sanou, K.M., Lompo, M. and Adouabou, B.A. (2024) Analyse de la Dynamique d'Occupation des Terres dans la Réserve de Biosphère Transfrontalière du W/Burkina Faso dans un Contexte de Terrorisme. Analyse de la Dynamique d'Occupation des Terres dans la Réserve de Biosphère Transfrontalière du W/Burkina Faso dans un Contexte de Terrorisme. Collection Pluraxes/Monde.

[17] Zoungrana, A., Visser, M., de Cannière, C., Ouédraogo, P.C., Bationo, B.A. and Traoré, S. (2023) Influence des changements agraires sur la dynamique paysagère autour des aires protégées du complexe Pô-Nazinga-Sissili au Burkina Faso. *Tropicultura*, **41**, 2280.

[18] USAID (2024) Cartographie de l'occupation des terres de la commune de Dédougou de 2017 à 2023. 26 p.

[19] Sama, A., Ramdé, Z., Ouédraogo, P.C., Sikuzani, Y.U., Salomon, W., Dipama, J., *et al.* (2023) Pression anthropique et dynamique de l'occupation du sol autour du corridor n°1 du complexe écologique Pô-Nazinga-Sissili au Burkina Faso. *Tropicultura*, **41**, 2200. https://doi.org/10.25518/2295-8010.2200

[20] Grimene, C. (2018) Classification de l'occupation des terres agricoles par télédétection satellitaire. Mémoire de Master, Gestion Agricole et Territoires. Institut Agronomique Méditerranéen de Montpellier et Université Paul Valéry de Montpellier III. 64 p.

[21] Ahmad, A. and Quegan, S. (2012) Analysis of Maximum Likelihood Classification on

Multispectral Data. *Applied Mathematical Sciences*, **6**, 6425-6436.

[22] Moskolaï, W.R. (2022) Exploitation des algorithmes d'apprentissage profond et des séries temporelles d'images satellitaires pour la prédiction de la déforestation. Thèse de doctorat, Université bourgogne Franche-Comté université de bourgogne, France université de Maroua, Cameroun. Ecole doctorale n°37-SPIM: Sciences pour l'Ingénieur et Microtechniques Unité de Formation Doctorale: Sciences de l'Ingénieur. 124 p.

[23] Genserbe, B.M., Assoma, V.T., Kouamé, K. and N'Guessan B.V.H. (2023) Machine Learning appliquée aux images MSI Sentinel-2A pour la cartographie des ressources agropastorales dans le Fitri au nord du Tchad. *Revue RAMReS-Sciences Appliquées et de l'Ingénieur*, **5**, 57-64.

[24] El Kharki, O., Mechbouh, J., Ducrot, D., Rouchdi, M. and Ngono, J.M. (2015) Panorama sur les méthodes de classification des images satellites et techniques d'amélioration de la précision de la classification. *Revue Française de Photogrammétrie et de Télédétection*, **210**, 23-38. https://doi.org/10.52638/rfpt.2015.259

[25] Matsaguim Nguimdo, C.A. and Tiomo, E.D. (2020) Forêt d'arbres aléatoires et classification d'images satellites: Relation entre la précision du modelé d'entrainement et la précision globale de la classification. *Revue Française de Photogrammétrie et de Télédétection*, **222**, 3-14. https://doi.org/10.52638/rfpt.2020.477

[26] Bernard, I. (2018) Comparaison des performances prédictives de deux algorithmes construisant des forêts aléatoires par HEC MONTRÉAL, Sciences de la gestion Option: Intelligence d'affaires. Mémoire présenté en vue de l'obtention du grade de maîtrise ès sciences en gestion (M.Sc.). 68 p.

[27] Harris, J.R., He, J.X., Rainbird, R. and Behnia, P. (2014) A Comparison of Different Remotely Sensed Data for Classifying Bedrock Types in Canada's Arctic: Application of the Robust Classification Method and Random Forests. *Geoscience Canada*, **41**, 557-584. https://doi.org/10.12789/geocanj.2014.41.062

[28] Ndiaye, M., Boissy, R., Faye, M. and Kouao, N.S.R. (2025) Évaluation comparative des algorithmes d'apprentissage automatique pour la classification des types de sols à partir de caractéristiques physico-chimiques: Application de Random Forest, XGBoost, SVM et KNN. *VertigO*, **24**, 16 p. https://doi.org/10.4000/13raq

[29] Emma, P. (2025) Apport du LiDAR aéroporté à la modélisation d'habitat de l'avifaune forestière wallonne: Une approche comparative par Random Forest et sélection adaptative de variables. Master Bioingénieur en gestion des forêts et des espaces naturels, Université de Liège, 92 p.

[30] Tiomo Dongfack, E., Matsaguim, N.C.A. and Tsafack Ngoufo, S. (2021) Evaluation de l'impact de la correction atmosphérique sur la précision globale de la cartographie de l'occupation du sol par télédétection. *Revue Scientifique et Technique Forêt et Environnement du Bassin du Congo*, **16**, 12-21.

[31] Kuma, J.K. (2019) Estimation par la méthode du Maximum de Vraisemblance: Éléments de Théorie et pratiques sur Logiciel. Centre de Recherches Economiques et Quantitatives (CREQ). Congo-Kinshasa. https://hal.science/cel-02189969v1