INITIATIVE ON Aquatic Foods

CGIAR

Small reservoirs in the Northern regions of Ghana and their vulnerability to drying

Ebenezer K. Siabi, Komlavi Akpoti, Sander J. Zwart

December 2023

Author affiliations

Ebenezer K. Siabi[1], Komlavi Akpoti[1], Sander J. Zwart[1]

[1] International Water Management Institute, Accra, Ghana

Suggested citation.

Siabi, E. K.; Akpoti, K.; Zwart, S. J. 2023. *Small reservoirs in the northern regions of Ghana and their vulnerability to drying*. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Initiative on Aquatic Foods. 37p.

Acknowledgements

This work was carried out with support from the CGIAR Aquatic Foods Initiative (AqFI). We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund.

Cover photo: Giulia Zane (IWMI)

About the CGIAR Aquatic Foods Initiative (AqFI)

The CGIAR Initiative on Aquatic Foods aims to tackle systemic challenges to the sustainability and resilience of aquatic food systems, including data gaps that lead to the exclusion of the sector from wider food and nutrition policies and programs and limited research investment. Working closely with research partners in fisheries and aquaculture, civil society, industry, and governments, the Initiative contributes to the reduction of greenhouse gas emissions from the production of aquatic foods. It enhances ecological and social resilience through the development and dissemination of improved fish strains, better management practices, integrated fish-rice production systems, and fish-friendly irrigation systems. Learn more about AqFI here**:** https://www.cgiar.org/initiative/aquatic-foods/

CONTENTS

1. INTRODUCTION

1.0 BACKGROUND AND CONTEXT

Aquatic foods are a crucial component of the global food system, offering numerous benefits ranging from nutrition to economic stability. However, these systems face a range of stressors, including overharvesting, pollution, climate change impacts, and governance challenges. These stressors threaten the sustainability of aquatic ecosystems and impact the livelihoods and food security of communities dependent on these resources. The Resilient Aquatic Food Systems (AqFS) Initiative, a collaboration between World Fish and the International Water Management Institute (IWMI), represents a concerted effort to address these challenges.

1.1 AQFS INITIATIVE AND ITS FOCUS IN GHANA

The AqFS Initiative operates across six regions and 11 countries, focusing on systematic challenges such as data gaps, gender imbalances, and water mismanagement. The Initiative's goal is to leverage aquaculture for multi-use purposes, combating food insecurity and promoting sustainable practices. Specifically, in Ghana, the Initiative aims to integrate AqFS into multifaceted water management plans and strengthen data-driven strategies to transform AqFS in the face of climate change.

1.2 PROJECT OBJECTIVES IN NORTHERN GHANA

In Northern Ghana, the AqFS Initiative focuses on introducing fish cage culture in small reservoirs, collaborating with the Fisheries Commission and the Water Research Institute (CSIR-WRI). The overarching objectives are to enhance the multifunctionality of water bodies, improve food security, and empower women and youth through aquaculture business development. The Initiative also explores scaling successful business models to communities around inland valleys, focusing on identifying and characterizing small reservoirs for potential development.

1.3 METHODOLOGICAL APPROACH: MACHINE LEARNING AND REMOTE SENSING

This technical report presents a comprehensive analysis of the dynamics of small reservoirs in Northern Ghana, using cutting-edge machine learning techniques and remote sensing data processed through the Google Earth Engine (GEE). The approach involves mapping small reservoirs, assessing their maximum extent, and analyzing their monthly water availability, particularly during the dry season. This geospatial analysis aims to perform a suitability analysis based on biophysical and socioeconomic characteristics, enabling a more informed selection process for aquaculture development.

1.4 RELEVANCE AND IMPLICATIONS

The findings of this report are crucial for understanding water availability and the potential for aquaculture in small reservoirs. By providing detailed insights into the dynamics of these water bodies, the report supports the broader goals of the AqFS Initiative in enhancing food security, promoting sustainable aquaculture practices, and empowering local communities in Northern Ghana. The methodologies and insights derived from this study also contribute to addressing the broader challenges aquatic food systems face globally, as highlighted by various researchers and initiatives.

2. METHODOLOGY

2.1 DATA ACQUISITION

We retrieved multispectral imagery over the five northern regions of Ghana (Upper West, Upper East, Northeast, Northern, and Savannah) for the study period. These provide high resolution ranging from 10-20m images of the selected regions at 5-day intervals when combining images from Sentinel-2A and Sentinel 2B, which began operating in June 2015 and March 2017, respectively. We selected images from November to April for each dry season between 2018 and 2023. However, the acquisition of images was done monthly throughout the study period. The image collection within each month was clipped to the boundary of the study area to reduce storage issues and processing time. This resulted in a total of 30 months of different images. The number of images attained varied across the months. For instance, 190, 202, and 169 images were obtained for November 2018, 2019, and 2022, respectively.

2.2 IMAGE PRE-PROCESSING

In tropical regions, frequent and dense cloud cover can significantly impede the accuracy of satellite imagery analysis. This is a critical issue when assessing water bodies, as clouds can lead to misinterpretation of water pixels. Clouds and water bodies often share similar spectral properties, confusing image processing algorithms. This spectral similarity can result in overestimating or underestimating water extent, directly impacting the accuracy of studies on water resource management and ecological assessments.

Additionally, optical remote sensing imagery, commonly used to monitor the spatial and temporal distribution patterns of inland waters, faces limitations due to cloud contamination. This contamination often results in low-quality images or missing data, further complicating the analysis. Selecting cloud-free scenes or combining multi-temporal images to produce a cloud-free composite image can partially overcome this issue, but it often comes at the cost of reduced monitoring frequency. This trade-off highlights the need for advanced and more robust image processing techniques to effectively differentiate between clouds and water bodies, especially in tropical environments where cloud cover is a persistent challenge.

Therefore, we applied a cloud mask to all images using the quality assurance band (QA60) to eliminate cirrus and thick clouds from the dataset. However, we found that this technique could not remove clouds in severely affected regions in the study area for some specific months. We found that images obtained for April for each dry year contained some level of clouds even after applying the cloud mask. Figure 1 compares the image collection of April 2018 before and after applying the cloud mask from the QA60 band. Areas within the enlarged part of the study area contained cloud footprints and shadows even after applying the cloud mask (Figure 2). There are several other methods of creating cloud-free images from image collections. Whereas some methods apply machine learning algorithms such as random forests, support vector classifiers, and stochastic gradient descent (see Hollstein et al. [1]) for cloud removal, others use manual methods to create cloud-free images that are as artifact-free as possible [2]. Appendix 1 reveals the various methods tested in this study and examples of their outputs. However, after applying some of these methods, four major issues were consistently persistent. These include: 1. Footprints of clouds were not entirely removed, 2. Some transparent clouds were not removed; 3. Images with more than 75% clouds were not corrected, 4. Some water pixels were mistaken as shadows or clouds and removed.

We therefore applied the Sentinel-2 Cloud Masking with the s2cloudless algorithm in Google Earth Engine (GEE). Overall, this approach was superior in correcting the major issues observed compared to the previous techniques. The s2cloudless represents an automated cloud-detection algorithm explicitly designed for Sentinel-2 imagery [3]. This algorithm relies on a gradient-boosting technique and was created by the EO Research team at Sinergise. It is openly available under the MIT License and can be accessed via the link: https://github.com/sentinel-hub/sentinel2-cloud-detector. The algorithm was trained using an extensive global dataset and is a mono-temporal algorithm, disregarding spatial context, thereby allowing its execution at any resolution. GEE offers users precomputed s2cloudless cloud probability maps and masks, covering the entire Sentinel-2 archive [4].

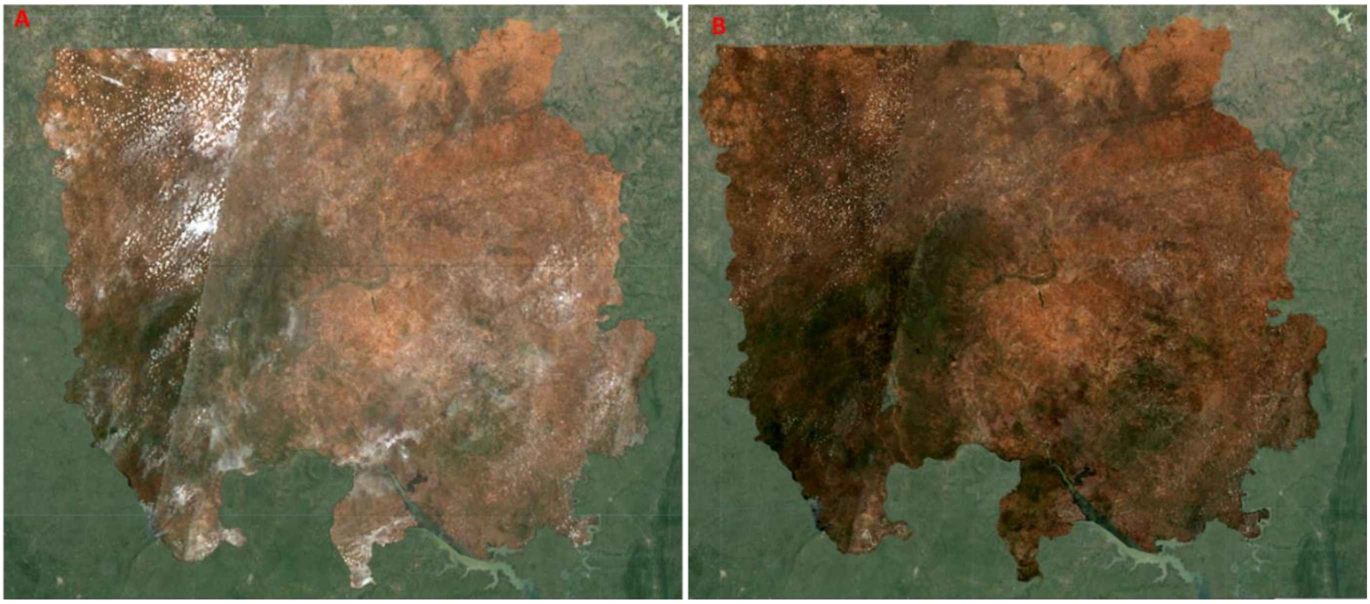

Figure 1. Image collection of April 2018 before (A) and after (B) applying the cloud mask from the QA60 band

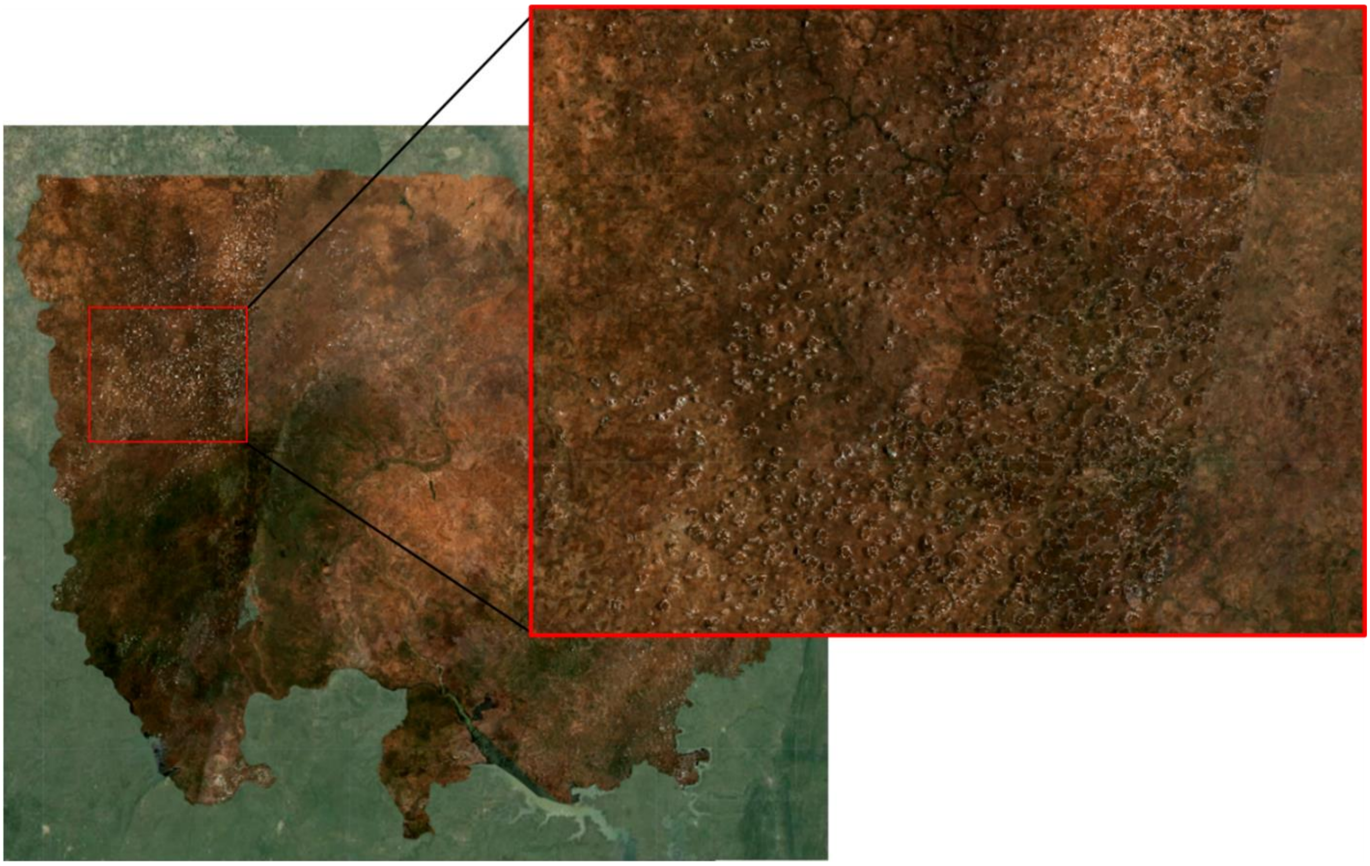

Figure 2. Enlarged area of clouds persistence after applying QA60 band

2.3 THRESHOLDING

In applying the algorithm, users can transform the cloud probability map into a cloud mask by applying a threshold to the cloud probability map. The developers suggest a recommended threshold value of 40% to reduce the occurrence of cloud omission errors [4]. However, our study found an optimal threshold value of 30% after applying different thresholds between 10% and 40%. We observed that some parts of the study area and small reservoirs, were detected as clouds and shadows, leading to their removal after applying 10% and 20% thresholds. For instance, parts of the Vea dam were removed after using the 10% and 20% thresholds (Figure 3).

Conversely, cloud footprints were persistent after using a threshold value of 40%. As such, a threshold value of 30% was applied to all the images used for the study. Figure 4 compares the April 2018 image at different cloud probability thresholds between 10% and 40%.

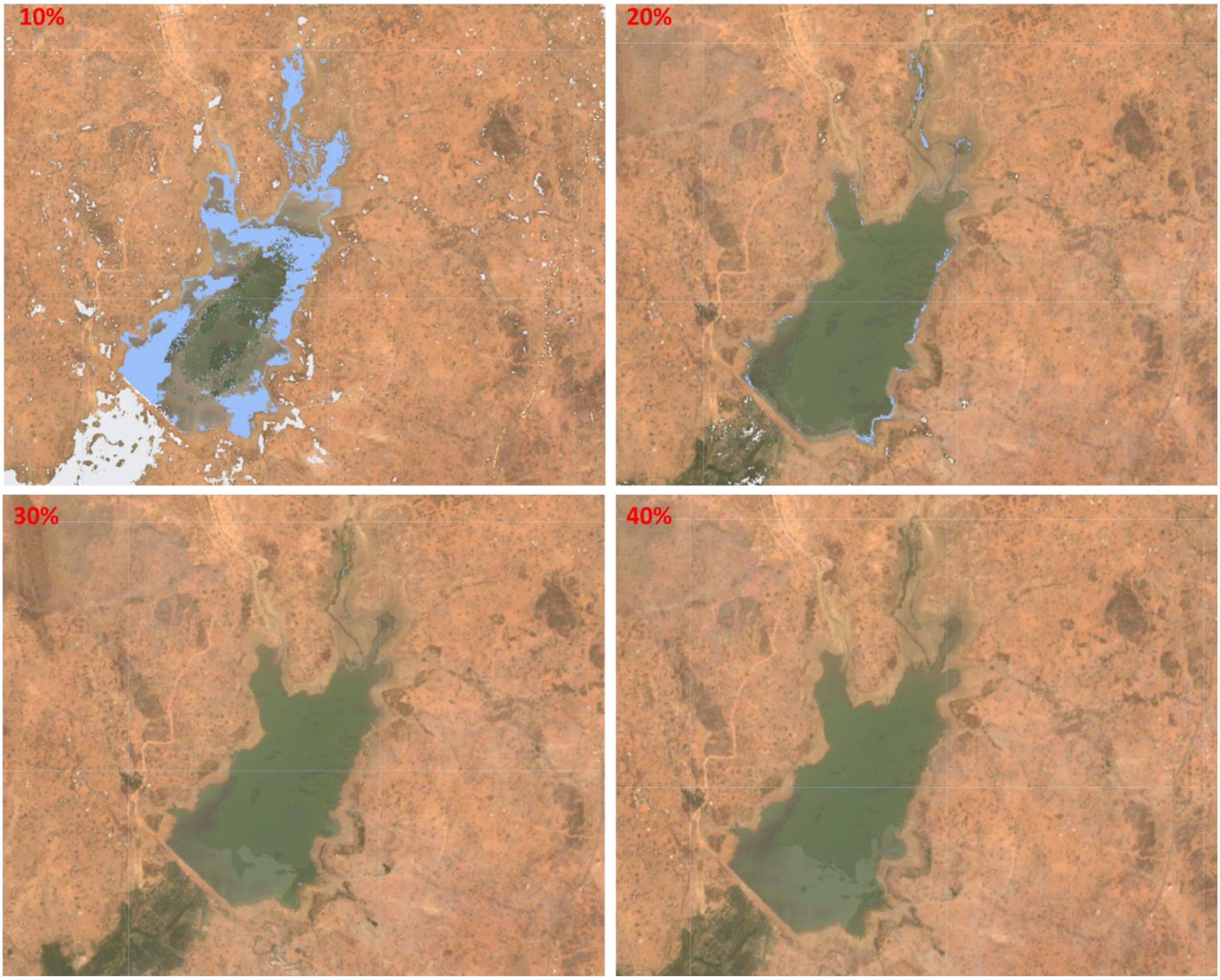

Figure 3. Vea dam after using the different cloud probability thresholds

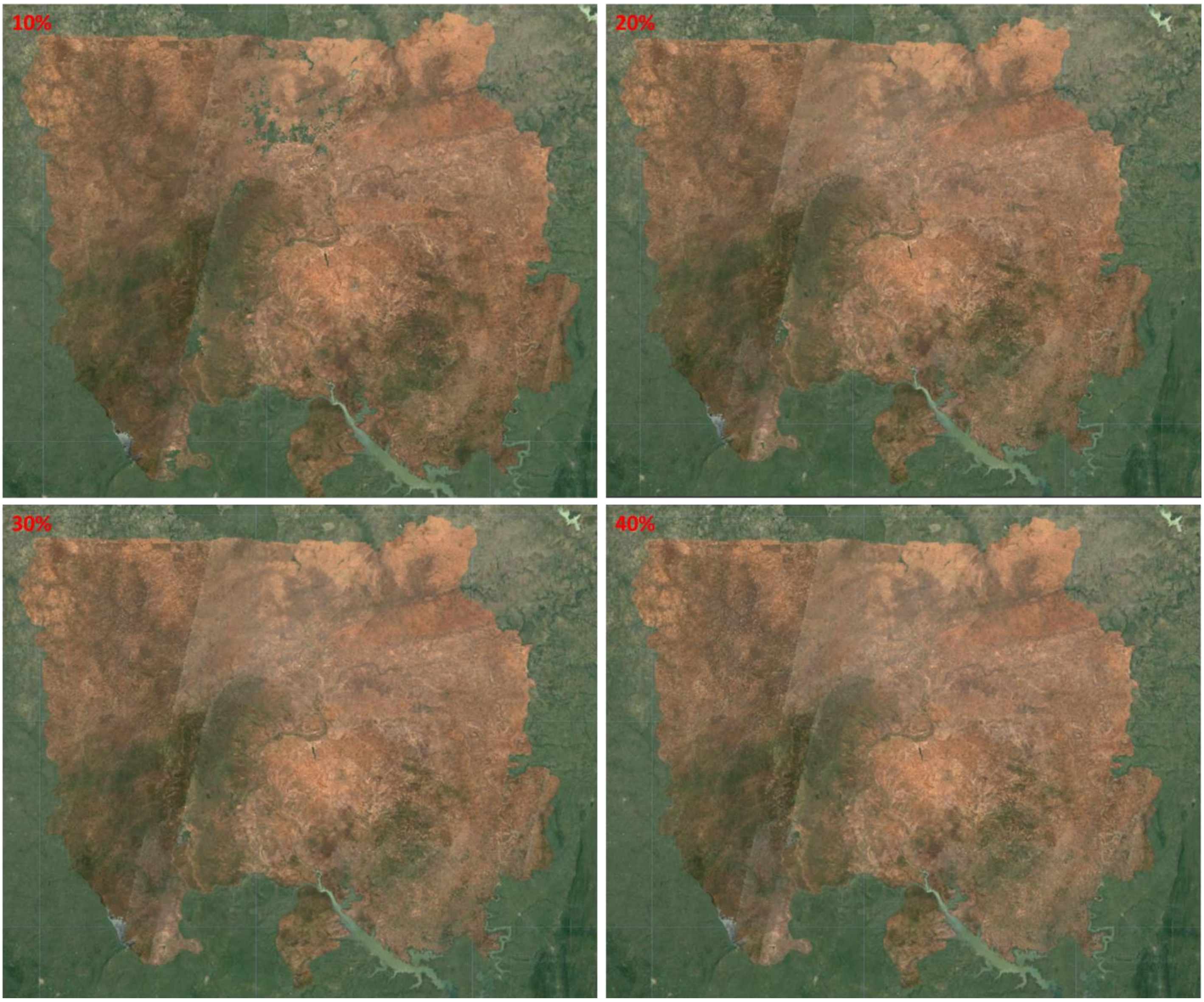

Figure 4. Comparison of April 2018 images at different thresholds between 10% and 40%

2.4 IMAGE ANALYSIS

A total of 5 bands were used for the study (Table 1). Two spectral indices were computed using the pre-process band collection. These spectral indices were mapped across the image collections for each month. This was done to enhance the classification accuracy of the model. The modified Normalized Water Index (MNDWI) and the Normalized Difference Vegetation Index (NDVI) were added as bands to all image collections and composited. The composited image collections formed the input data for the random forest algorithm. The MNDWI [5] and NDVI [6] are given as equations 1 and 2, respectively;

$$MNDWI = \frac{B3 - B11}{B3 + \mathrm{B11}} \quad (1)$$

$$NDVI = \frac{B8 - B4}{B8 + B4} \quad (2)$$

Table 1. Details of bands used for the study

Name	Scale	Pixel Size (meter)	Wavelength		Description
			Sentinel 2A	Sentinel 2B	
B2	0.0001	10	496.6nm	492.1nm	Blue
B3	0.0001	10	560nm	559nm	Green
B4	0.0001	10	664.5nm	665nm	Red
B8	0.0001	10	835.1nm	833nm	NIR
B11	0.0001	20	1613.7nm	1610.4nm	SWIR 1

2.5 MACHINE LEARNING CLASSIFICATION

Despite having several machine learning algorithms for mapping, we utilized the Random Forest algorithm (RFA) to classify small reservoirs in the study area. The RFA was utilized over other machine learning algorithms such as support vector classifiers, NaiveBayes, and Classification and Regression Trees (CART) because It's an ensemble learning method that builds multiple decision trees and merges their outputs. Combining these trees reduces overfitting and improves accuracy compared to single decision tree models. During the training phase, the RFA generates an extensive array of decision trees and subsequently determines a class based on the highest frequency, thereby demonstrating resilience against overfitting [7], [8]. The RFA adheres to the fundamental principles of bootstrap aggregating (bagging). This method, from a training set alongside respective responses, consistently chooses random samples with replacement and constructs trees based on these selected samples. However, the distinction of RFA from the typical bagging technique lies in its utilization of a modified tree learning algorithm. This algorithm opts explicitly for a random subset of the available features at each potential split during the learning process. Following the training phase, predictions can be generated by aggregating the majority vote derived from the classification trees.

2.6 GENERATING TRAINING SAMPLES AND TRAINING OF RFA

The training dataset comprises a FeatureCollection containing a property that stores the class label, while other properties hold predictor variables. Class labels need to be integers starting from 0 in consecutive order. Also, remap() ensures class values are in sequential integers. Additionally, the predictor variables should be numeric. To train the RFA for our study, a reference sample dataset was generated for both water and non-water. First, using the geometry drawing tools, a high-resolution satellite base map in the GEE environment was used to sample water and non-water points and polygons. Class labels were then assigned to the water and non-water samples, where 0 and 1 represent non-water and water, respectively. Each feature collection was given a property called 0 and 1, signifying the properties for storing predictor variables of the training sample. Next, we generated at least 80-100 feature collections (encompassing points and polygons) for each class using the drawing tool in the code editor. Finally, we merged the water and non-water samples and divided it into training (80%) and testing (20%) datasets.

Next, the RFA builds an individual decision tree for each sample. Utilizing the predictors (which comprise bands extracted from composited image collections), these trees collectively determine the classification of each pixel as water or non-water through voting. Subsequently, each pixel is assigned the value that receives the most support from these decision trees. We employed the .smileRandomForest in the GEE environment with 100 trees and 4 randomly selected predictors per split. The water property was extracted from the classes and used as the class property. Also, four bands (B4, B5, B3, MNDWI) were used as the input properties. Figure 5a shows the RFA input, training, and output process, while Figure 5b shows the overall development of the RFA. The validation error matrix and overall accuracy were used to evaluate the performance of the RFA.

The classified small reservoir extent for each month was cleaned by masking all unconnected pixels in the GEE environment. Also, a stream network was extracted from 1 arcsec Shuttle Radar Topographic Mission (SRTM) Digital Elevation Model (DEM) for the study area. In extracting the streams in the ArcGIS environment, all sinks in the DEM were filled using the fill tool. Flow direction and accumulation were then generated using the flow direction and flow accumulation tools, Finally, the map Algebra was used to extract the stream network of the study area. The extracted stream network was converted into a vector and imported into the GEE environment. Next, we masked all permanent waters using a 100m buffered stream network vector. The final classified reservoir extent was then converted into a vector and exported. Some features observed as unusual were manually removed. Also, all small reservoirs with a maximum extent of less than 0.09 hectors (ha) were removed [9]. This ensures that the final small reservoir extent does not include dugouts.

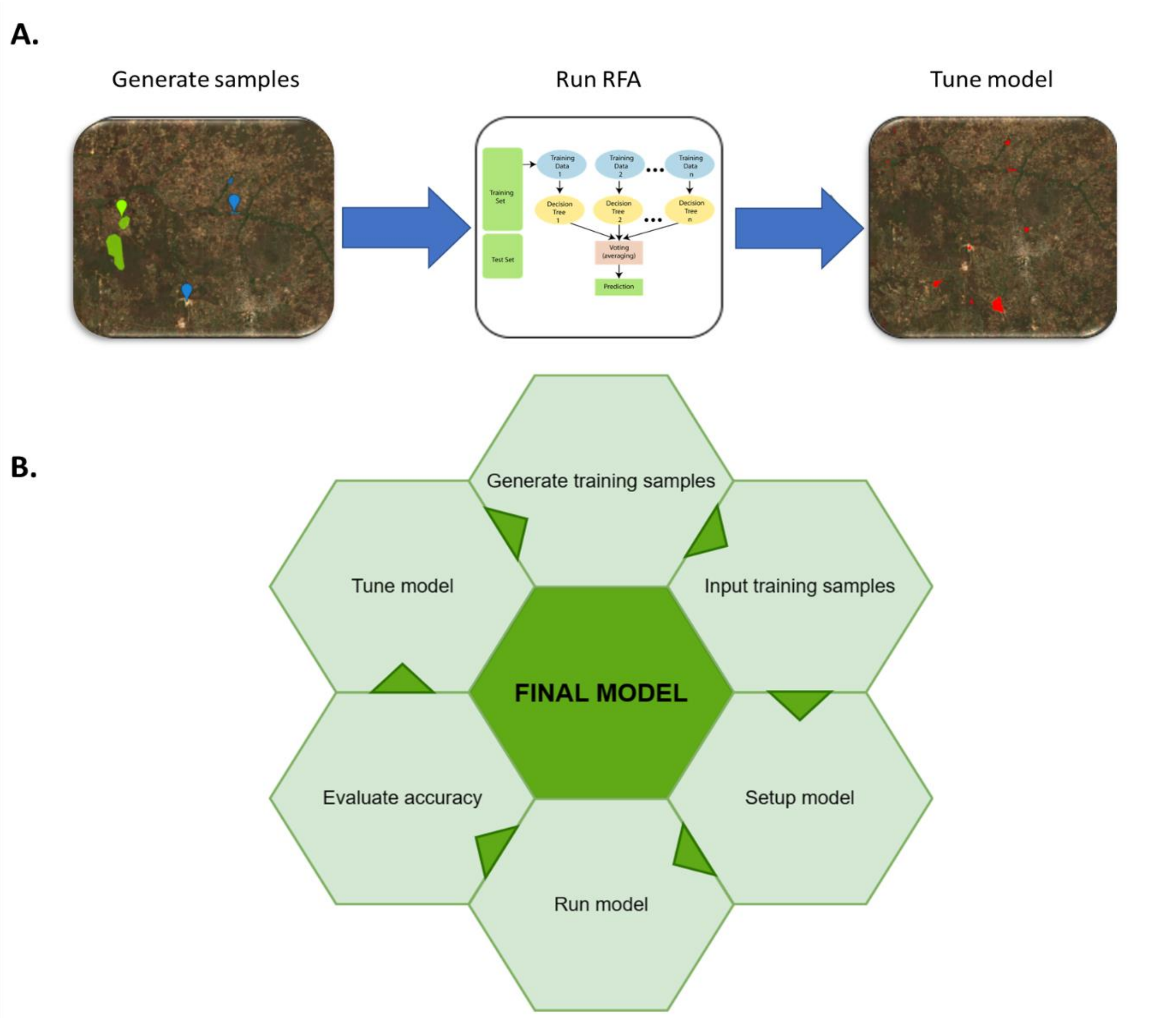

Figure 5. A. The input, training, and output process. B. Overall model development process of the RFA.

2.7 MODEL VALIDATION AND PERFORMANCE EVALUATION

We validated the classified small reservoir extent by comparing it to a high-resolution Google Earth image. This was done to assess the correlation between the actual and predicted small reservoir extents across the study area. We followed three major steps to validate the predicted small reservoir extents. First, 72 small reservoir extents were selected from the predicted dataset at varying sizes (ranging from 0.18 to 54.08). Next, the 72 small reservoir extents were digitized from Google Earth on different dates (November and December 2018, November and December 2019, and January and November 2020) based on their corresponding selected small reservoirs. This produced a validated small reservoir layer of the study area. Finally, the extent values of the observed and actual small reservoirs were compared to assess the lateral accuracy of the classified small reservoir extents.

In assessing the performance of the RFA, we employed the confusion matrix and the F1-score. The confusion matrix shows various relevant measures such as overall accuracy (OA), ' 'Producer's Accuracy (PA), ' 'User's

Accuracy (UA), and the Kappa coefficient (Foody, 2002). The F1-score, on the other hand, denotes the accuracy of binary classification and is calculated based on precision and recall using equations (3) and (4). These metrics collectively aid in assessing the reliability and performance of the classification process.

$$P = \frac{TP}{TP+FP} \tag{3}$$

$$R = \frac{TP}{TP+FN} \tag{4}$$

Where,

P is precision
R is Recall
TP is True positives
FP is False positives
FN is False negatives

2.8 CHARACTERIZATION OF SMALL RESERVOIRS

2.8.1 Size characterization

The predicted extents of small reservoirs were categorized into three size categories. Reservoirs with areas below 0.6 hectares were categorized as small, those ranging from 0.6 to 6 hectares were categorized as medium, and reservoirs exceeding 6 hectares were classified as large reservoirs.

2.8.2 Landscape characterization

Here, we categorized the extent of the predicted reservoirs by considering their location within the landscape. The locations of the small reservoirs contextualized, based on the hypothesis that there is an increasing trend in the storage capacity of the small reservoirs from upstream, midstream, and downstream of the landscape. Initially, the small reservoirs were categorized based on their flow accumulation capacity, considering their location in the landscape [10]. We generated a flow accumulation raster for the study area using the digital elevation model from the 1 arcsec Shuttle Radar Topographic Mission (SRTM) data. Subsequently, we extracted the cell value of the location of the maximum flow accumulation (approximated as the outlet position) for each small reservoir. These values were categorized into three intervals based on their standard deviations from the mean where values less than 2.8, between 2.8-5.6 as well as those greater than >5.6 were categorized as upstream, midstream, and downstream, respectively [9], [11].

2.8.3 Risk of falling dry

The risk of falling dry defines the Probability of the small reservoir drying within a specific dry year (i.e., November 2018 to April 2019) of the entire study period (i.e., November 2018 to April 2023). Here, the months of water occurrence (Ranging from 1-to 30, which represents the number of months for the entire study period) was used to categorize the small reservoirs. Small reservoirs with less than 5 months of water occurrence were categorized under very high risk of falling dry. This is followed by those with water occurrence between 5-10 months, categorized as high risk of falling dry. Those between 10-15, 15-20, 20-25, and above 25 months were categorized under medium, low, very low, and extremely low risk of falling dry, respectively. For the risk of falling dry for each dry year, the months of water occurrence ranged from 1-6 months. Therefore, water occurrence for 1-6 months represented Very high, High, Medium, Low, Very low, and Extremely low risk of falling dry, respectively.

We estimated the Probability of water occurrence (PWO) for each dry year as (equation 5)

$$PWO = \frac{n^i}{N} \times 100 \qquad (5)$$

Where;

n^i is number of months that water occurred in i^{th} small reservoir
N is total number of months per dry year or study period.

The probability of risk of falling dry (PRF) is estimated as (equation 6):

$$PRF = 100 - PWO \qquad (6)$$

2.9 DATABASE DEVELOPMENT

We developed a geodatabase for the small reservoir extent using unique IDs. This was achieved after a series of different activities. First, we merged all monthly small reservoir extents for each dry year. Next, we dissolved all geometries into a single geometry for each year. The outputs for each year were then overlayed and dissolved to generate an overall small reservoir extent from 2018-2023. Then the multi polygon output was fragmented into individual polygons, and unique IDs were generated for each polygon. We then converted the coordinate system to Geographic WGS 84 CRS (EPSG:4326) and estimated the centroids. This was followed by extracting the longitudes and latitudes for each polygon. The final output is a reference layer. This layer was used to assign unique IDs and transfer longitude and latitude to the individual months of small reservoir extents for each dry year and saved in their respective subfolders.

3. RESULTS

3.1 VALIDATION RESULTS

Figure 6 presents the validation results demonstrating the correlation between the extent classified by RFA and the digitized extent obtained from high-resolution Google Earth imagery. This comparison resulted in an R-squared value of 0.98, indicating a strong correlation between the extent of 72 observed and classified small reservoirs. The surfaces of the 72 observed and classified small reservoir extents between 0.18ha and 54ha were compared (Figure 6). Figure 7 further exhibits six specific instances of classified small reservoirs overlaid on Google Earth imagery at varying dates, offering a visual representation of the accuracy achieved in the classification process. Also, the overall accuracy ranged from 0.94 to 0.99, with November 2021 and April 2023 recording the highest and lowest scores, respectively (Table 2). Similarly, Kappa and F1 score ranged from 0.887 to 0.999 and 0.998 to 0.861 (Table 2).

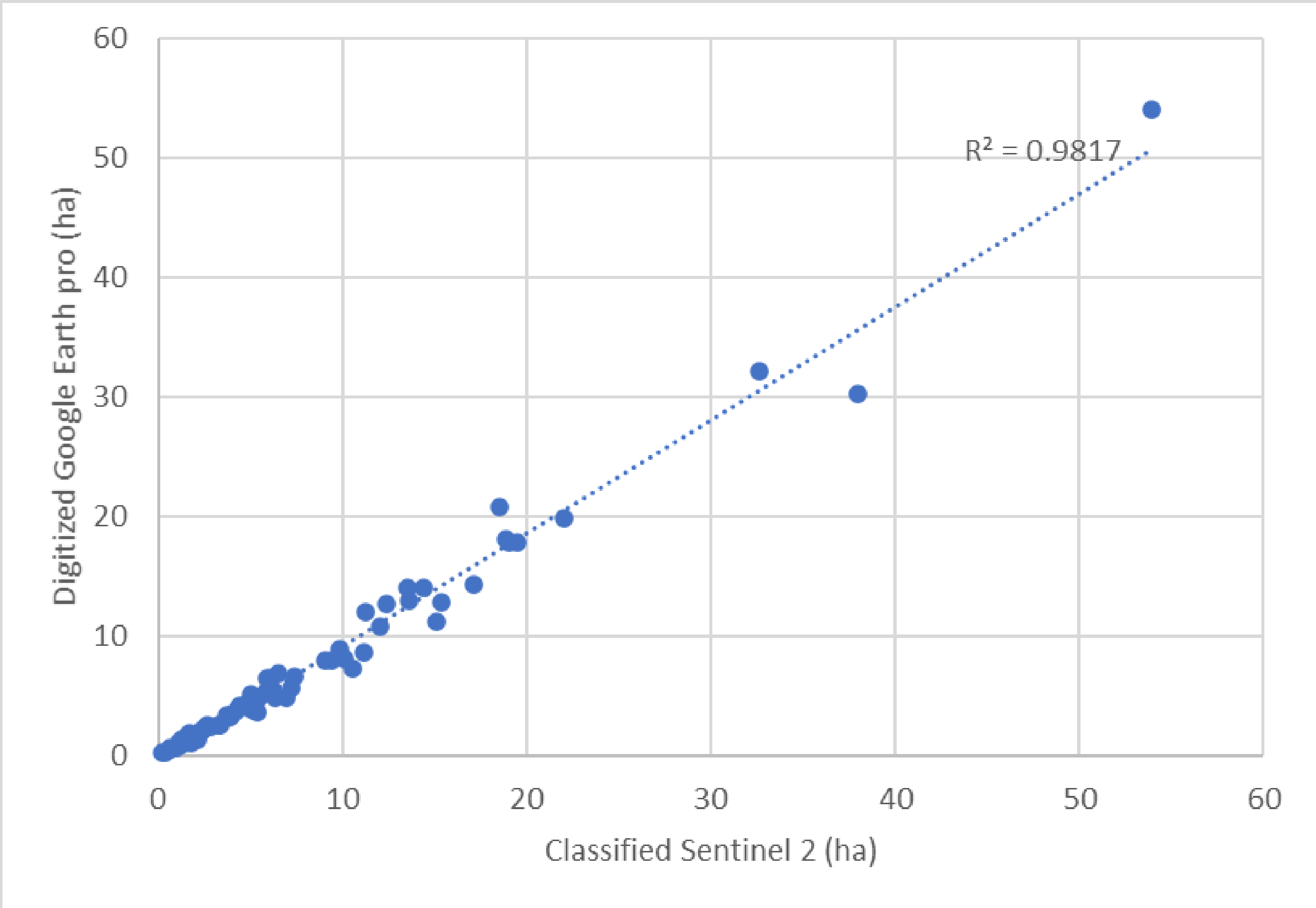

Figure 6. Scatter plot showing the correlation between the classified and digitized small reservoir extent from RFA and high-resolution Google Earth imagery, respectively

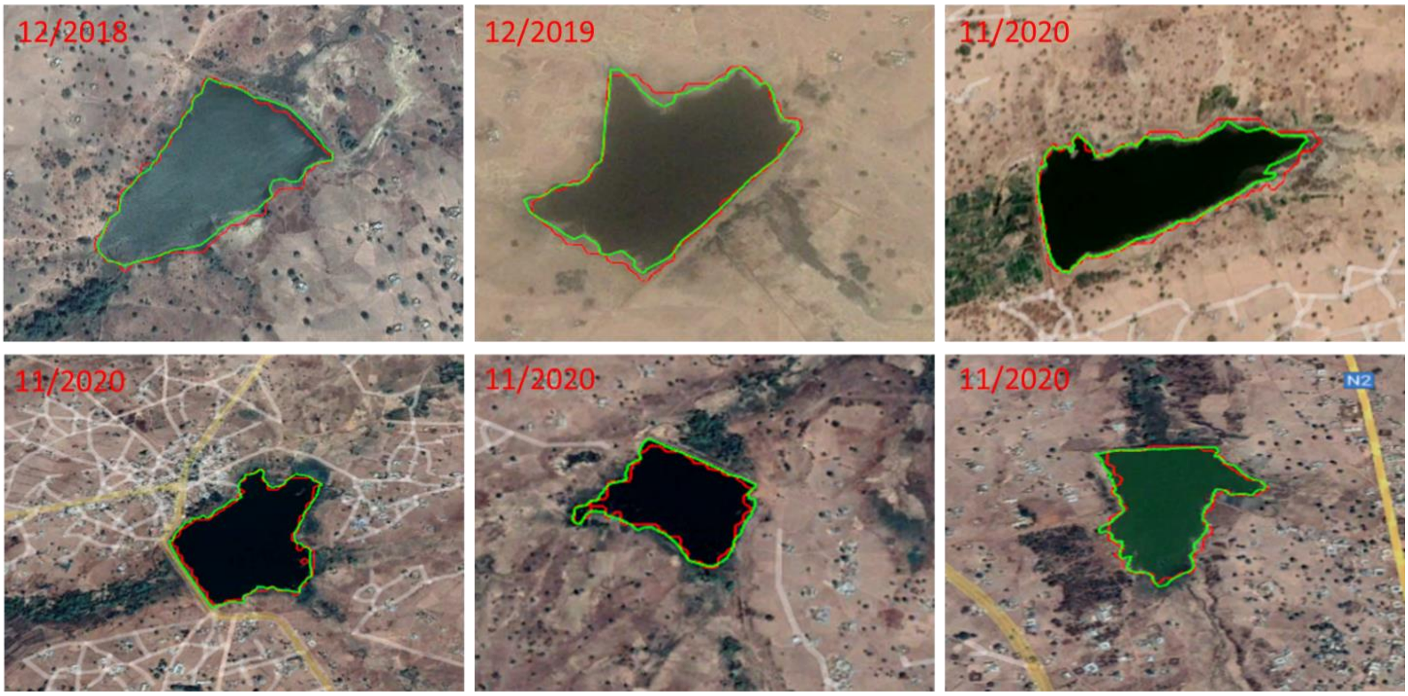

Figure 7. Comparison of the classified and digitized small reservoir extent from RFA and Google Earth Engine, respectively, at varying dates (NB: Green = Digitized, Red = Classified)

3.2 ANALYSIS OF SMALL RESERVOIR EXTENT AND NUMBER

The trend in the number of small reservoirs during each dry season indicates a consistent decline from November to April. For instance, in November 2018, there were 1,422 small reservoirs, which decreased to 446 by April 2019 (Figure 8). This phenomenon could be ascribed to the escalating intensity of the dry season, which tends to peak during the dry months, notably in April of each consecutive dry year. Similarly, the trend of the surfaces of the small reservoirs declined as the months progressed, with intensification in April (Figure 9) throughout all the dry years. As such, due to progressive drying, most small reservoirs tend to reduce in size (Figure 10a) or completely dry up (Figure 10b) by April. The overall reservoir extents varied between 6738ha and 7368ha in November and, similarly, between 3013 hectares and 3647 hectares in April, spanning across the dry years (Figure 9).

Conversely, there is a noticeable increase. The number of small reservoirs steadily rose as the dry years progressed, reaching its peak in November 2022 (Figure 8). The initial count of 1422 small reservoirs in November 2018 escalated to 2292 by November 2022.

Table 2. Model performance evaluation

Dry year	Month	Overall Accuracy	Kappa	Producer Accuracy	User Accuracy	F1-Score
2018–2019	Nov	99.10	0.991	95.30	99.71	0.991
	Dec	99.41	0.999	96.02	98.10	0.985
	Jan	99.37	0.990	96.24	99.95	0.987
	Feb	99.12	0.972	94.91	98.42	0.960
	Mar	98.65	0.977	93.99	98.05	0.952
	Apr	98.01	0.961	91.17	96.81	0.926
2019–2020	Nov	99.93	0.995	96.95	99.78	0.998
	Dec	99.78	0.995	95.42	99.60	0.993
	Jan	98.79	0.986	89.78	96.85	0.978
	Feb	98.81	0.979	87.07	97.50	0.974
	Mar	96.67	0.940	83.79	96.58	0.960
	Apr	96.88	0.931	81.49	97.74	0.969
2020–2021	Nov	99.92	0.980	97.59	99.76	0.990
	Dec	98.87	0.988	98.91	99.69	0.987
	Jan	98.40	0.966	96.28	99.87	0.979
	Feb	97.77	0.947	95.72	98.90	0.962
	Mar	97.98	0.939	93.91	98.71	0.943
	Apr	96.81	0.921	91.97	96.79	0.948
2021-2022	Nov	99.95	0.980	97.50	99.91	0.970
	Dec	98.99	0.986	95.64	99.94	0.962
	Jan	98.70	0.953	94.20	99.75	0.946
	Feb	97.78	0.925	92.09	99.54	0.927
	Mar	96.45	0.918	90.42	96.88	0.904
	Apr	95.51	0.902	89.30	94.09	0.899
2022-2023	Nov	99.94	0.983	97.14	99.78	0.982
	Dec	98.99	0.987	96.40	99.81	0.940
	Jan	99.90	0.966	94.75	97.90	0.930
	Feb	97.81	0.929	94.09	97.49	0.881
	Mar	96.87	0.905	92.70	95.93	0.874
	Apr	94.79	0.887	87.90	92.41	0.861

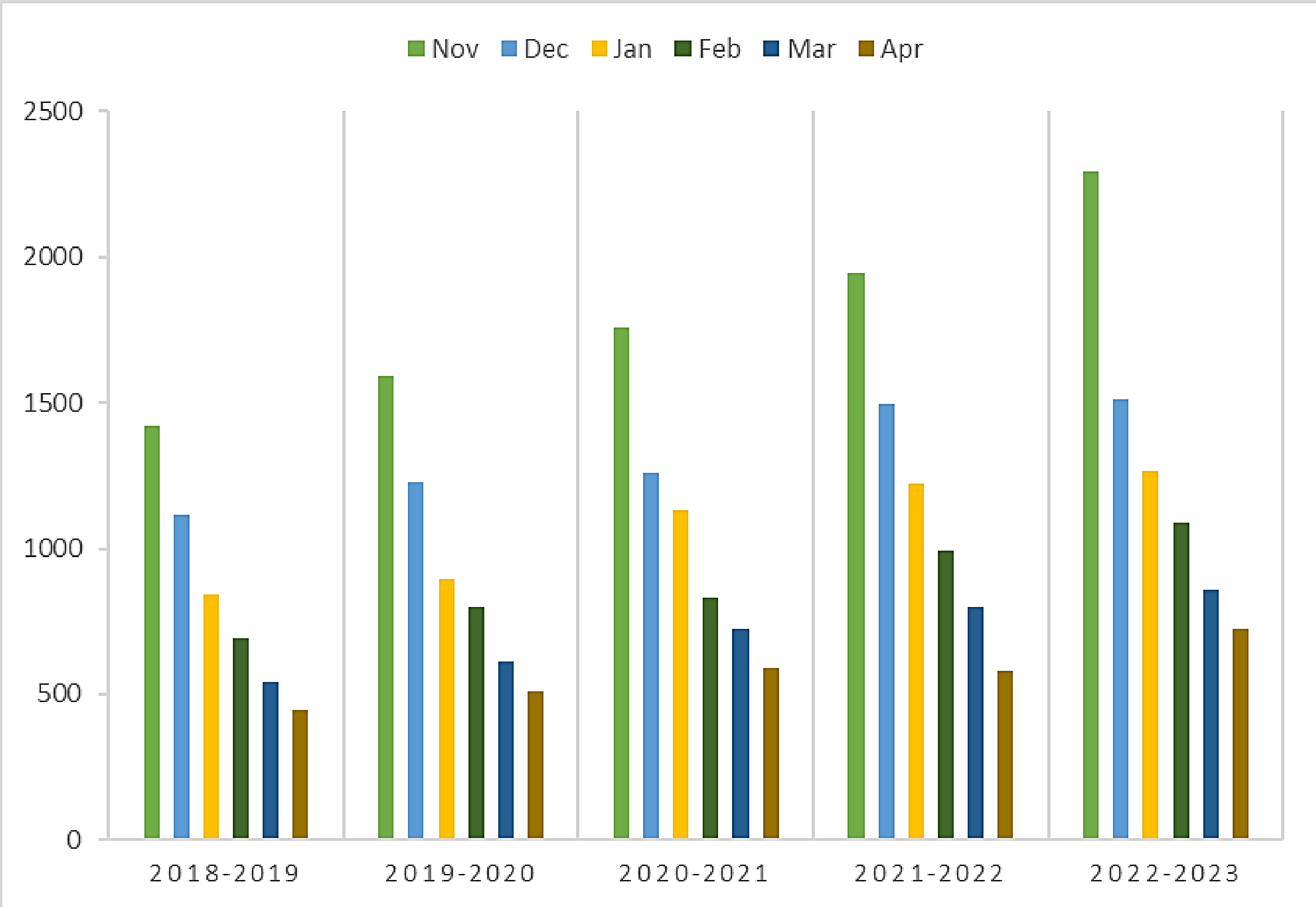

Figure 8. Variations in the number of small reservoirs both between seasons and within a single season

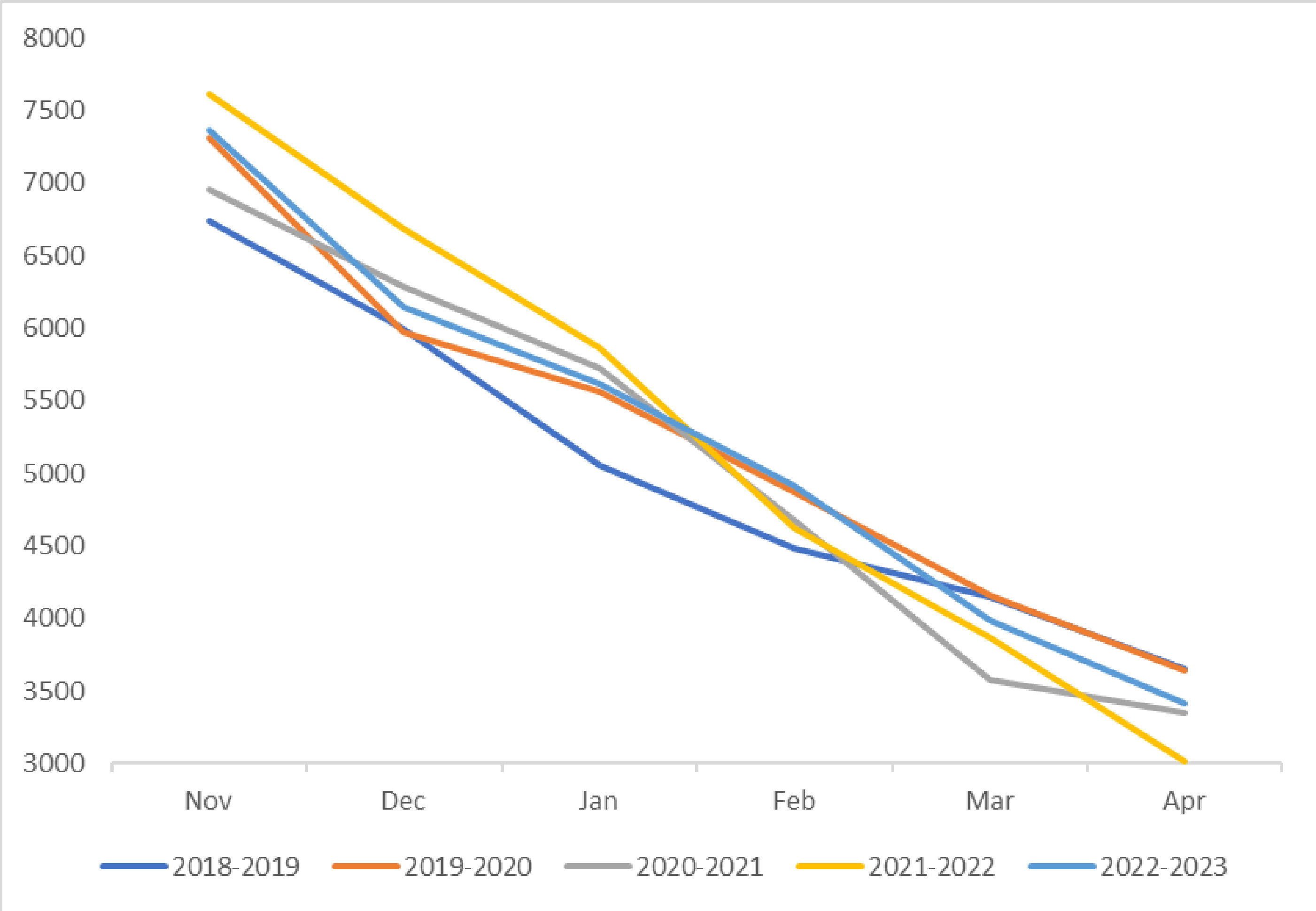

Figure 9. Variations in the surface extent of small reservoirs across five distinct dry seasons.

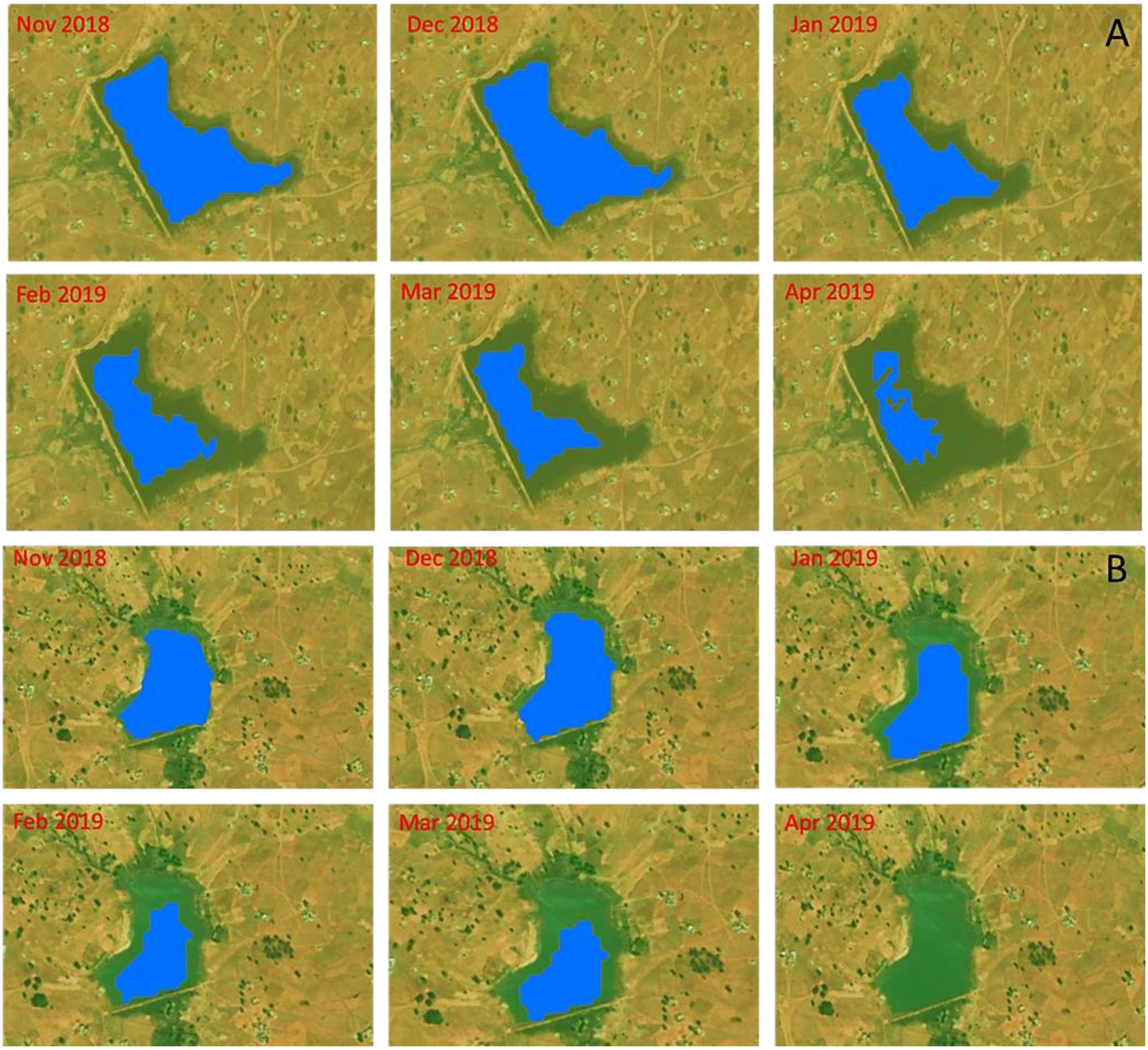

Figure 10. (A) An example of a small reservoir holding water throughout the 2018-2019 dry year and (B) drying by the end of the 2018-2019 dry year

3.3 CHARACTERIZATION OF RESERVOIRS

3.3.1 SIZE

Figure 11 illustrates the spatial distribution of small reservoirs categorized by size during the dry year of 2018-2019. The larger reservoirs (> 6.0 hectares) are notably clustered in the Upper East region, surpassing the concentrations observed in other regions, as depicted in Figure 11. Following this, the Upper West and Northern regions exhibit lesser but still noticeable concentrations of large reservoirs. November 2018 showcased a notable aggregation of small reservoirs (< 0.6 hectares). However, this concentration of small reservoirs gradually diminished from December 2018 to April 2019.

The November classifications served as the basis for assessing the numerical variations in the number of reservoirs categorized by size at the onset of the dry years. Across all the years, as depicted in Figure 12, small-sized reservoirs (< 0.6 hectares) dominated in number when compared to medium and large reservoirs. For instance, in November 2022, the count of small reservoirs ranged between 754 and 1349, surpassing the numbers of medium reservoirs (which increased from 478 to 746 in November 2021 with a slight decrease in November 2022) and large reservoirs (fluctuating between 174 and 235, with minor declines noted in November 2020 and 2022, respectively).

Table 3 shows the total number of small reservoirs within our analysis at the beginning (November) and end (April) of each dry year. As previously highlighted, there was a consistent, gradual decline in reservoirs from the beginning to the end of each dry year. For instance, in November 2018, the number of small reservoirs decreased from 771 (54%) to 158 (11%) by April 2019 (Table 3), resulting in a total of 613 (43%) small reservoirs that had dried up. Similarly, medium-sized reservoirs decreased from 478 to 212, and large reservoirs reduced from 174 to 76 by the end of the 2018 dry season. Consequently, this led to a total of 266 (19%) dried medium-sized reservoirs and 98 (7%) dried large-sized reservoirs during the dry year of 2018-2019.

They commenced the 2019-2020 dry season with a total count of 1592 reservoirs, which dwindled to 510 reservoirs by the season's end in April 2020. This decline accounted for a total reduction of 1082 reservoirs, equivalent to approximately 68% of the initial reservoir count. Among the different size categories, small reservoirs experienced the most significant drying compared to medium and large reservoirs (Table 3). Specifically, 544 small reservoirs, constituting about 34% of the total reservoirs, dried up by the season's end, while 401 medium and 137 large reservoirs exhibited drying.

Comparable patterns persist across the subsequent dry years (i.e., 2020-2021, 2021-2022, and 2022-2023), where a larger proportion of small reservoirs dried up by the end of the dry season in contrast to medium and large reservoirs. Notably, there is a consistent escalation in the number of dried-up reservoirs as each year advances. The 2022-2023 dry year recorded the highest cumulative count of dried-up reservoirs, reaching 1567 by the end of the dry season (Table 3).

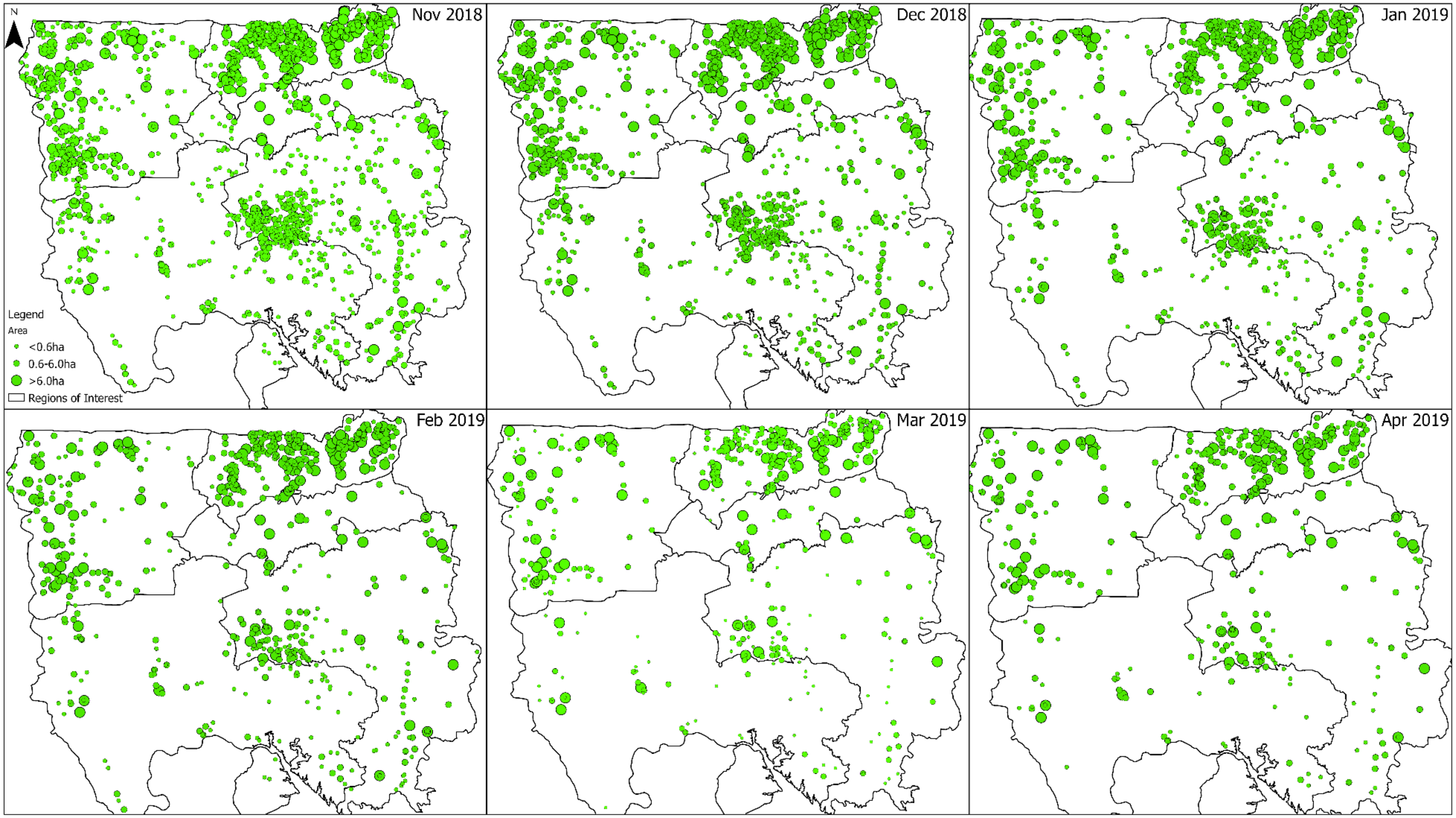

Figure 11. Spatial distribution of small reservoirs by size for the 2018-2019 dry year

Table 3. Number of reservoirs at the beginning and end of each dry year

Beginning of Dry season (November)

Size	2018	%	2019	%	2020	%	2021	%	2022	%
Small	771	54.18	754	47.36	919	52.28	969	49.69	1349	58.86
Medium	478	33.59	627	39.38	633	36.01	746	38.26	724	31.59
Large	174	12.23	211	13.25	206	11.72	235	12.05	219	9.55
Total	1423	100	1592	100	1758	100	1950	100	2292	100

End of Dry season (April)

Size	2019	%	2020	%	2021	%	2022	%	2023	%
Small	158	11.10	210	13.19	258	14.68	268	13.74	293	12.78
Medium	212	14.91	226	14.21	258	14.68	248	12.72	351	15.31
Large	76	5.34	74	4.65	76	4.32	65	3.33	81	3.53
Total	446	31.34	510	32.04	592	33.67	581	29.79	725	31.63

Dried Reservoirs (Change)

	2018-19	%	2019-20	%	2020-21	%	2021-22	%	2022-23	%
Small	613	43.08	544	34.17	661	37.61	701	35.95	1056	46.07
Medium	266	18.69	401	25.19	375	21.33	498	25.54	373	16.27
Large	98	6.89	137	8.61	130	7.39	170	8.72	138	6.02
Total	977	68.66	1082	67.96	1166	66.33	1369	70.21	1567	68.37

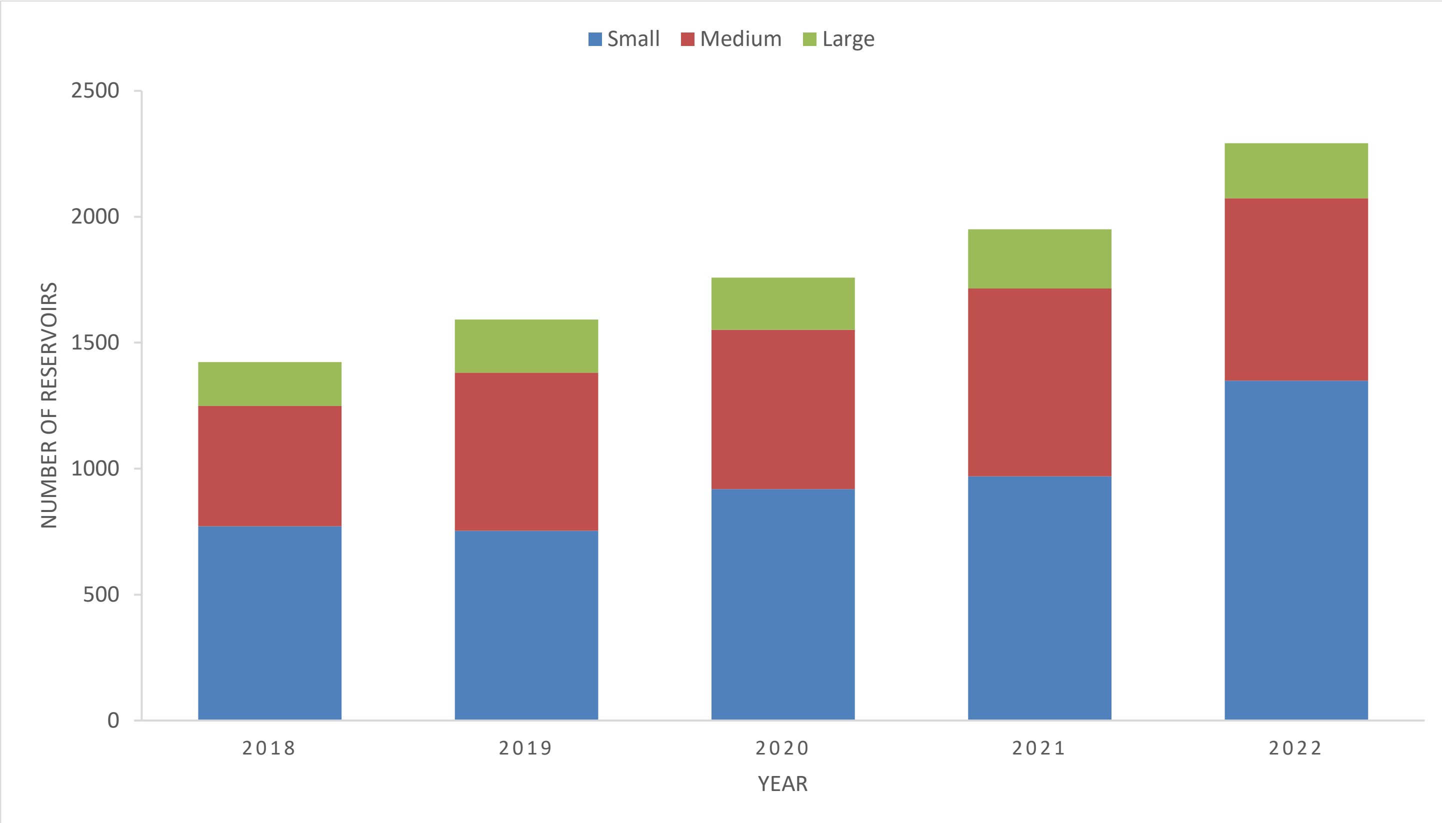

Figure 12. Number of reservoirs by size at the beginning of dry year

3.3.2 LANDSCAPE

The Reservoirs were categorized based on their positioning in the landscape and size. Among the categorized reservoirs, approximately 50% (1534) were situated in the midstream, with about 30% (934) found in the downstream (Table 4). Approximately 20% of the reservoirs were located in the upstream. Figure 13 illustrates the spatial distribution of small reservoirs based on their landscape positioning. Meanwhile, Figure 13 indicates the prevalence of small-sized reservoirs (<0.6ha) in the upstream (555) and midstream (873). Moreover, the downstream exhibited a higher presence of medium-sized reservoirs (464). Notably, while there were no large reservoirs upstream, the midstream and downstream regions recorded 81 and 246 large reservoirs, respectively, as depicted in Figure 14.

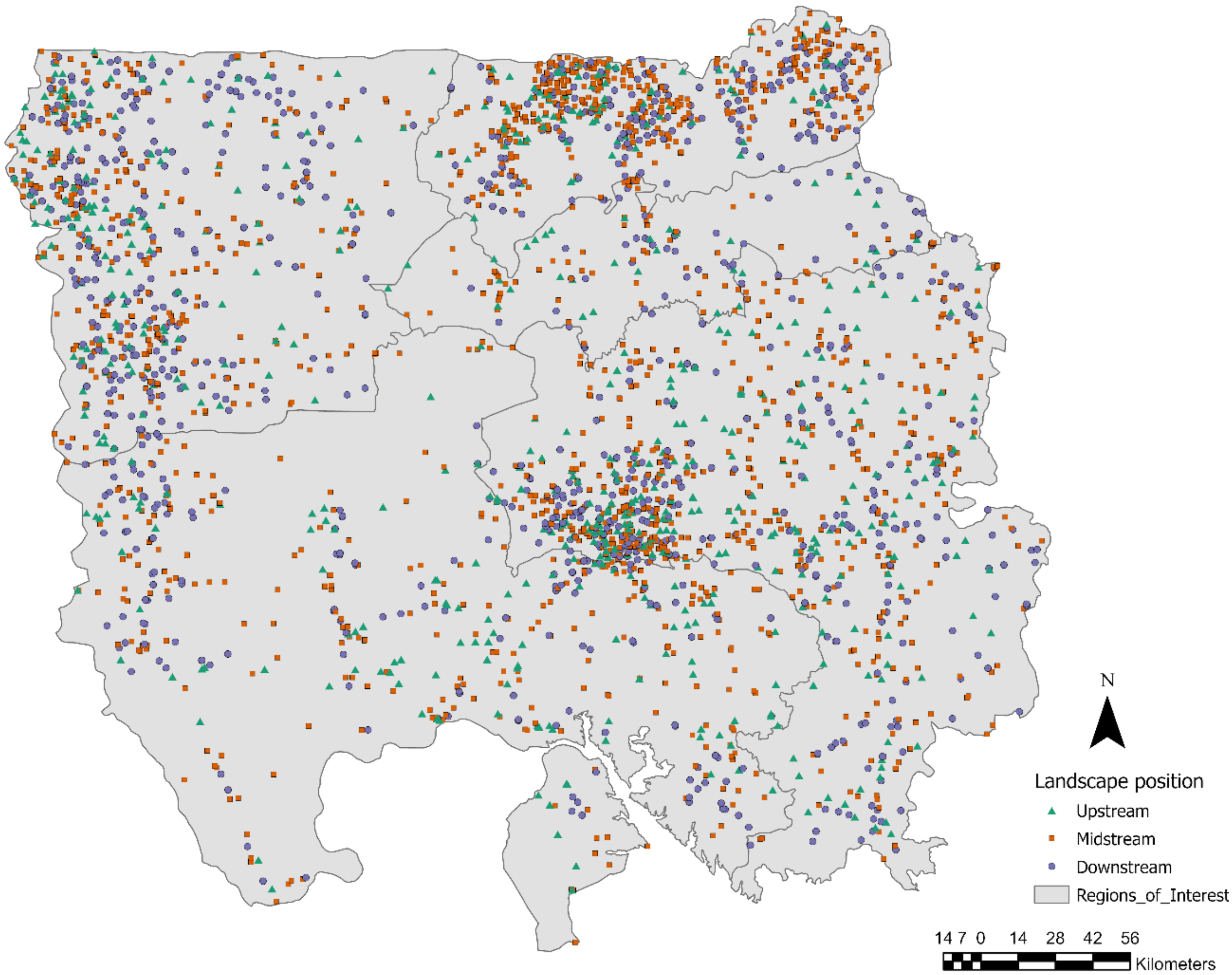

Figure 13. Spatial distribution of small reservoirs by position in landscape

Table 4 4. Position of small reservoirs in the landscape

Landscape Position	Count	Percent
Upstream	611	19.84
Midstream	1534	49.82
Downstream	934	30.33
Total	3079	100

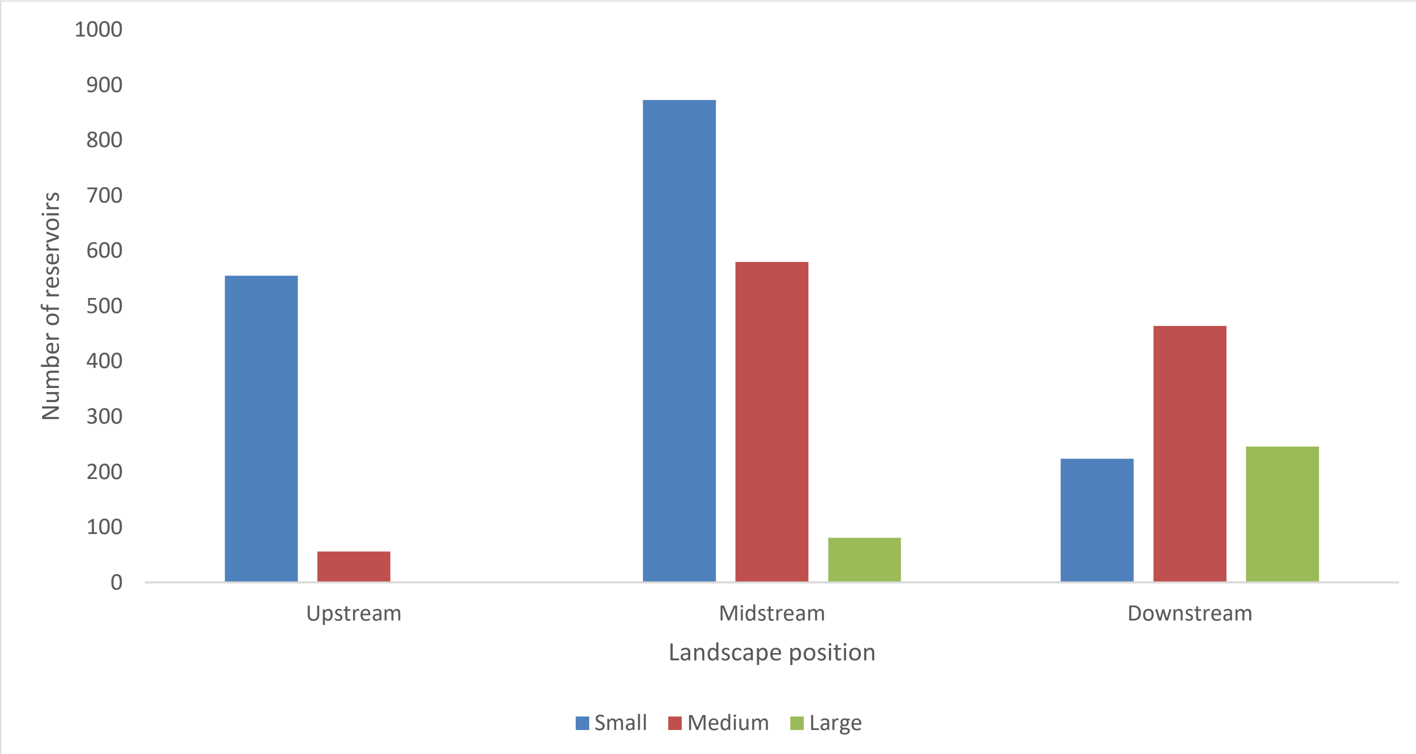

Figure 14. Small reservoir characteristics due to Size and landscape position

3.3.3 RISK OF FALLING DRY

The study assessed the probability of small reservoirs drying up during the study duration and individual dry years. Overall, findings suggest that approximately 1506 small reservoirs, constituting about half of the total, face a very high risk of drying up (Table 5). Additionally, approximately 454 (15%) and 312 (10%) small reservoirs are classified as having high and medium risk of drying up, respectively. However, there are 245 (8%), 196 (6%), and 366 (12%) small reservoirs identified as having low, very low, and extremely low risks, respectively, of drying up.

During each specific dry year, a significant portion of the reservoirs, totaling 391, exhibit an extremely low risk of drying up in the 2018-2019 dry period. Subsequently, roughly 284 reservoirs (21%) are identified as having a very high risk of drying up. Similar trends persist in subsequent dry years, such as 2019-2020, 2020-2021, and 2021-2022, where the majority of small reservoirs are classified as having an extremely low risk of drying up. However, in the 2022-2023 dry year, a substantial portion (33%) of small reservoirs is observed to be at a very high risk of drying up, while around 25% are noted to have an extremely low risk of drying up.

Reservoirs were assessed for their likelihood of drying based on their size. Surprisingly, most reservoirs, regardless of their size, fell into the category of very high risk for drying up. For example, roughly 50% (164) of the large reservoirs face a very high risk of drying up. In contrast, comparatively smaller percentages—around 9%, 5%, and 17%—of large reservoirs are classified as having low, very low, and extremely low risks of drying up, respectively (as detailed in Table 6). However, when comparing sizes, small reservoirs exhibited the highest proportion, approximately 53%, falling within the very high-risk of drying up.

Concerning the risk of drying based on their landscape position, it's notable that most reservoirs, regardless of their landscape placement, tend to face a very high risk of drying up. Specifically, approximately 56%, 50%, and 52% of small reservoirs were identified as having a very high risk of drying up, respectively (as outlined in Table 7). The remaining small reservoirs were distributed across various risk levels within the landscape, ranging from high to extremely low.

Table 5. Risk of falling dry for each dry year

Risk of falling dry	Overall		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023	
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
Very high	1506	48.91	284	21.35	393	25.37	424	25.68	438	23.46	748	33.11
High	454	14.75	257	19.32	294	18.98	180	10.9	273	14.62	188	8.32
Medium	312	10.13	147	11.05	114	7.36	281	17.02	266	14.25	391	17.31
Low	245	7.96	135	10.15	173	11.17	170	10.31	161	8.62	210	9.31
Very low	196	6.37	116	8.72	122	7.88	82	4.97	231	12.37	165	7.3
Extremely low	366	11.89	391	29.41	453	29.24	514	31.13	498	26.67	557	24.66

NB: Overall risk of failing is calculated using all 30 months for the 5 dry years compared to 6 months for each dry year

Table 6. Risk of falling dry due to Size

Risk of falling dry	Large	%	Medium	%	Small	%
Very high	164	50.15	543	49.36	873	52.85
High	37	11.31	137	12.45	206	12.47
Medium	26	7.95	114	10.36	172	10.41
Low	30	9.17	75	6.82	140	8.47
Very low	15	4.59	76	6.91	105	6.36
Extremely low	55	16.82	155	14.09	156	9.44
Total	327		1100		1652	

Table 7. Risk of falling dry due to position in Landscape

Risk of falling dry	Upstream	%	Midstream	%	Downstream	%
Very high	340	55.65	756	49.28	484	51.82
High	66	10.8	183	11.93	131	14.03
Medium	59	9.66	169	11.02	84	8.9
Low	44	7.2	115	7.51	86	9.21
Very low	37	6.06	101	6.58	58	6.21
Extremely low	65	10.64	210	13.69	91	9.74
Total	611		1534		934	

4. CONCLUSIONS

The comprehensive analysis of small reservoir dynamics and their vulnerability to drying revealed substantial insights into their characteristics and trends across various parameters. The validation results demonstrated a robust correlation between the extent of observed and classified small reservoirs, with an R-squared value of 0.98. Moreover, the temporal patterns observed during dry seasons unveiled a consistent decline in the number and size of small reservoirs from November to April, attributing this phenomenon to the escalating intensity of the dry season, particularly in April. This progressive drying trend was evident across multiple dry years, resulting in a reduction in reservoir extents between November and April, between 6738ha and 7368ha in November and between 3013 ha and 3647 ha in April. Conversely, examining trends across dry years, there was a noticeable increase in small reservoirs, peaking in November 2022 from an initial count of 1422 in November 2018. In terms of size, small-sized reservoirs (< 0.6 hectares) dominated across all years, outnumbering medium and large reservoirs.

The evaluation of reservoirs based on landscape positioning highlighted distinct spatial distributions, with midstream locations accommodating around 50% of the categorized reservoirs, followed by downstream areas (30%) and upstream (20%). Moreover, different sizes of reservoirs exhibited varying distributions across these landscape positions, with small-sized reservoirs prevailing in upstream and midstream areas. At the same time, downstream regions had a higher presence of medium and large reservoirs.

Assessing the risk of drying across different parameters revealed that approximately half of the small reservoirs faced a very high risk, with smaller proportions at high, medium, low, very low, and extremely low risks of drying. Notably, the consistency in the decline of small reservoirs by the end of each dry year was evident, with a greater proportion of small reservoirs drying than medium and large reservoirs. Furthermore, specific dry years exhibited varying proportions of dried-up reservoirs, with the 2022-2023 dry year recording the highest cumulative count of dried-up reservoirs. Remarkably, despite variations in size and landscape positioning, most reservoirs faced a very high risk of drying, indicating a pervasive vulnerability among these small reservoirs irrespective of their characteristics.

In conclusion, this comprehensive analysis provides valuable insights into the dynamics of small reservoirs, highlighting their vulnerability to drying, consistent patterns across dry seasons and years, and the influence of landscape positioning and size on their susceptibility to dryness. These findings can serve as a foundational understanding for effective water resource management strategies, particularly in mitigating the impact of drying among small reservoirs in the studied region.

REFERENCES

[1] Hollstein A., Segl K., Guanter L., Brell M., and Enesco M., Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images, *Remote Sens.*, vol. 8, no. 8, p. 666, Aug. 2016, doi: 10.3390/rs8080666.

[2] Schmitt M., Hughes L. H., Qiu C., andZhu X. X., """AGGREGATING CLOUD-FREE SENTINEL-2 IMAGES WITH GOOGLE EARTH ENGINE""," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. IV-2/W7, pp. 145–152, Sep. 2019, doi: 10.5194/isprs-annals-IV-2-W7-145-2019.

[3] Zupanc A., "Improving cloud detection with machine learning," *Accessed Oct*, vol. 10, p. 2019, 2017.

[4] Skakun S. *et al.*, "Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2," *Remote Sens. Environ.*, vol. 274, p. 112990, Jun. 2022, doi: 10.1016/j.rse.2022.112990.

[5] Xu H., "Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery," *Int. J. Remote Sens.*, vol. 27, no. 14, pp. 3025–3033, Jul. 2006, doi: 10.1080/01431160600589179.

[6] Pettorelli N., *The normalized difference vegetation index*. Oxford University Press, USA, 2013.

[7] Liu X. *et al.*, "Large-Scale High-Resolution Coastal Mangrove Forests Mapping Across West Africa With Machine Learning Ensemble and Satellite Big Data," *Front. Earth Sci.*, vol. 8, p. 560933, Jan. 2021, doi: 10.3389/feart.2020.560933.

[8] Shelestov A., Lavreniuk M., Kussul N., Novikov A., and Skakun S., "Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping," *Front. Earth Sci.*, vol. 5, Feb. 2017, doi: 10.3389/feart.2017.00017.

[9] Ghansah B., Foster T., Higginbottom T. P., Adhikari R., and Zwart S. J., "Monitoring spatial-temporal variations of surface areas of small reservoirs in Ghana's Upper East Region using Sentinel-2 satellite imagery and machine learning," *Phys. Chem. Earth Parts ABC*, vol. 125, p. 103082, Feb. 2022, doi: 10.1016/j.pce.2021.103082.

[10] Zhou G., Wei H., and Fu S., "A fast and simple algorithm for calculating flow accumulation matrices from raster digital elevation," *Front. Earth Sci.*, vol. 13, no. 2, pp. 317–326, Jun. 2019, doi: 10.1007/s11707-018-0725-9.

[11] Chao S. R., Ghansah B., and Grant R. J., "An exploratory model to characterize the vulnerability of coastal buildings to storm surge flooding in Miami-Dade County, Florida," *Appl. Geogr.*, vol. 128, p. 102413, Mar. 2021, doi: 10.1016/j.apgeog.2021.102413.

5. APPENDIX 1

Table A 1. Summary of various methods tested in the study

Method	Date	Source	Pros	Cons
Fitoprinciple	January 30, 2020	[fitoprincipe/gee-composite: Make cloud-free Landsat and Sentinel 2 composites using Google Earth Engine Python API (github.com)]	Able to filter clouds using cloud cover percentage to filter the collection Choose whether to mask the clouds out or not based on the maximum score for a given "day of year.", Maximum score for the best satellite in the given period, Minimum score for the pixels next to clouds (mask in general), Maximum score to the image with fewer masked pixels Index, and the maximum score for pixels with a given vegetation index (momentarily set to 0.8)	'Couldn't separate some shadows from water pixels. Transparent clouds are not removed.
Hollstein et al's 2016	2016	https://doi.org/10.3390/rs8080666	Categorizes cirrus, opaque, and shadows Categorizes bad and suitable pixels into a binary option (0,1) Uses a decision tree to select suitable pixels from bad pixels. Returns suitable pixels per scene	Detects some shadows as water pixels and removes them Footprints of clouds are not removed.
Aggregating cloud-free sentinel-2	September 20, 2019	ISPRS-Annals - AGGREGATING CLOUD-FREE SENTINEL-2 IMAGES WITH GOOGLE EARTH ENGINE (copernicus.org)	It does not infer pixels based on statistical or machine learning models but makes use of posterior information, which	Not able to remove transparent clouds

images (Schmitt et al, 2019)			was measured by Sentinel-2 While being able to generate mostly cloud-free images even for severely cloud-affected regions of interest (ROIs), the method always strives to create as clean images and artifact-free as possible. Using ' 'GEE's cloud computation infrastructure, it can efficiently produce cloud-free images for large numbers of ROIs and time frames in a parallel manner.	Images with more than 75% clouds are not corrected. User Memory demanding Orbit limits on the composites are visible.
Sentinel-2 Cloud Masking with s2cloudless	June 1, 2022	Sentinel-2 Cloud Masking with s2cloudless | Google Earth Engine | Google for Developers	Clouds are identified from the S2 cloud probability dataset (s2cloudless), and shadows are defined by cloud projection intersection with low-reflectance near-infrared (NIR) pixels.	In some cases, it is still possible to spot orbit limits on the composites.

fitoprincipe method

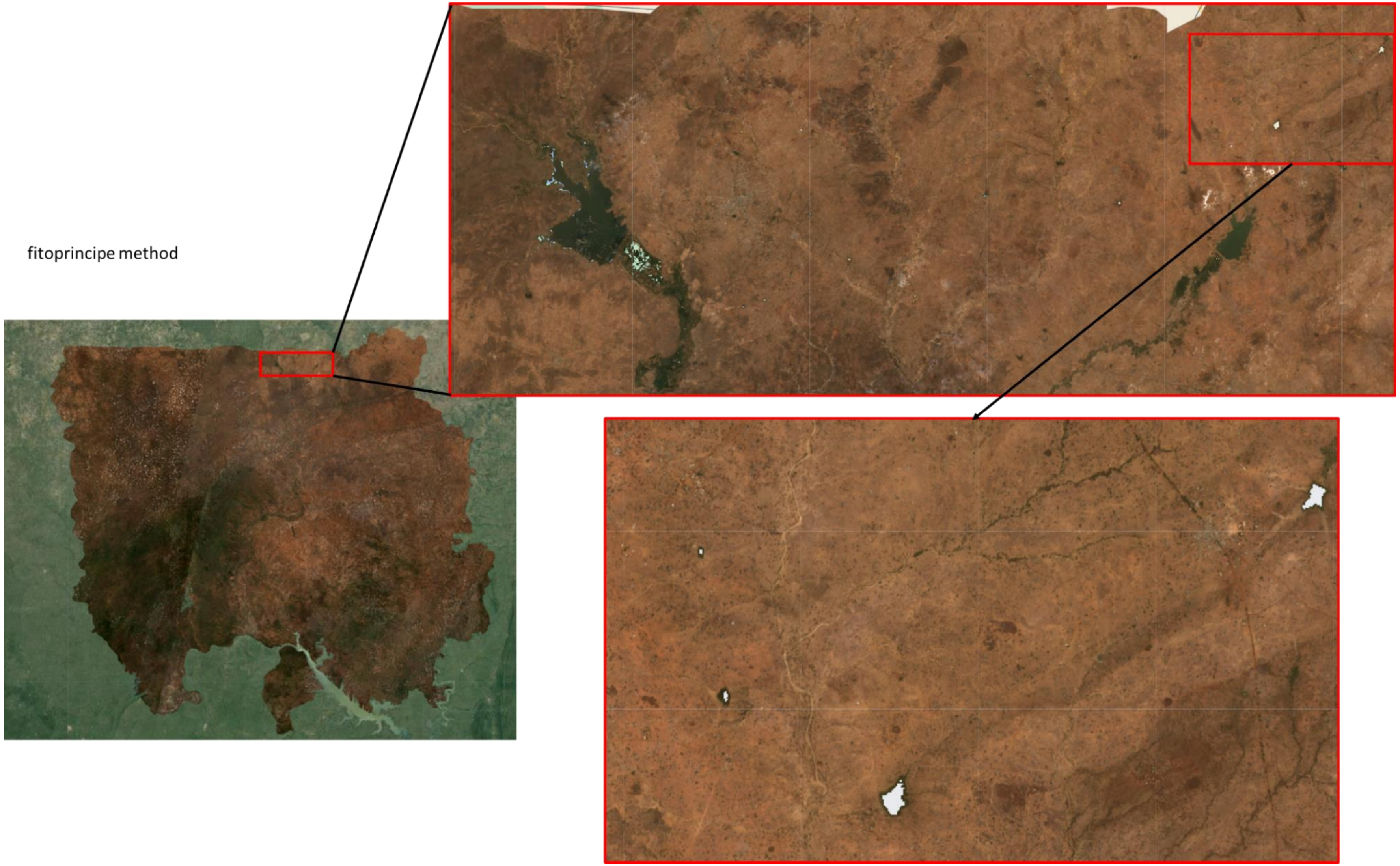

Figure A 1. Output of the fitoprincipe cloud removal technique. Enlarged areas showing part of some reservoir extents masked out as clouds with some reservoirs completely removed.

Figure A 2. The output of the Hollstein et al. (2016) cloud removal technique. The enlarged area shows part of some reservoir extents masked out as clouds with some reservoirs completely removed, as well as persistent cloud footprints in some areas.

Figure A 3. The output of the Schmitt et al. (2019) cloud removal technique. Enlarged areas showing part of the study area contaminated by cloud