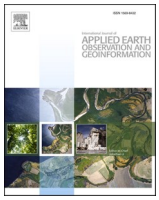
International Journal of Applied Earth Observation and Geoinformation 122 (2023) 103418

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Mapping irrigated agriculture in fragmented landscapes of sub-Saharan Africa: An examination of algorithm and composite length effectiveness

Timon Weitkamp [a,b,*], Gert Jan Veldwisch [b], Poolad Karimi [c], Charlotte de Fraiture [c]

[a] Resilience BV, Bevrijdingsstraat 38, 6703 AA Wageningen, the Netherlands
[b] Water Resources Management Group, Wageningen University, Droevendaalsesteeg 4, 6708 PB Wageningen, the Netherlands
[c] IHE Delft Institute for Water Education, PO Box 3015, 2601 DA Delft, the Netherlands

ARTICLE INFO

Keywords:
Random forest
Support vector machine
Neural network
k-nearest neighbor
Hotspot map
Remote sensing

ABSTRACT

Accurately identifying irrigated areas is crucial for sustainable development, food security, and effective land and water resource management. However, incomplete or outdated national estimates of irrigated areas underestimate the extent of it, particularly among smallholders. This study aimed to address this issue by investigating the impact of different algorithms and composite lengths on predicting irrigated agriculture in four study areas in Mozambique. The study found that the choice of algorithm and composite length notably impacted the accuracy of identifying irrigation. Shorter composite lengths, such as 2-monthly or 3-monthly composites, were more effective in identifying irrigation in fragmented and dynamic landscapes, while longer composite lengths were better suited to stable classes and homogeneous landscapes. Artificial neural networks, support vector machines, and random forests were all effective algorithms for classifying irrigation. However, the study emphasised the importance of considering hotspots and agreement maps when identifying irrigation. Agreement maps combine the classification results of multiple models, providing better insights into the core areas of irrigated agriculture and allowing for a better understanding of irrigation dynamics and policy decision-making, particularly among smallholder systems. This research provides valuable insights for those working on remote sensing-based irrigation mapping and monitoring in sub-Saharan Africa, focusing on identifying smallholder irrigation with greater certainty.

1. Introduction

Obtaining accurate information about irrigation is vital for making informed decisions about land and water resource management for food security and sustainable development (Bofana et al., 2020; Wellington and Renzullo, 2021). Unfortunately, national estimates of irrigated areas are often based on limited on-ground surveys or low-resolution remote sensing data for large-scale applications (Wellington and Renzullo, 2021). The available information is often outdated or incomplete (Beekman et al., 2014; Espey, 2019; Venot et al., 2021). Furthermore, limited budgets prevent officials from conducting regular in-person agriculture monitoring (Ajaz et al., 2019; de Bont, 2019; Ramezan et al., 2019).

African smallholder agriculture is a complex system that often takes place on small, irregular-shaped fields with in-class variance such as inter- and mix-cropping systems and variability in the timing of agronomic activities such as planting, harvesting and irrigation (Bégué et al., 2018; Izzi et al., 2021; Veldwisch et al., 2019). It is often found in mosaic landscapes where agriculture and natural vegetation alternate over short distances, resulting in frequent changes in land cover/use over short distances.

Distinguishing irrigated from rainfed agriculture or natural vegetation can be challenging, particularly in areas where soil moisture does not quickly deplete, such as near streams or in wetlands, which may have similar soil moisture patterns as irrigated croplands.

Despite the challenges of accurate mapping, quantifying and monitoring irrigation practices, remote sensing (RS) imagery has become popular for land use classification. Evaluating how different machine learning algorithms perform in classifications is one of the most studied aspects of land use classifications (Marín Del Valle and Jiang, 2022), of which the random forest (RF), support vector machine (SVM), artificial neural networks (ANN) and k-nearest neighbours (k-NN) are among the

* Corresponding author.
E-mail addresses: timon@resiliencebv.com (T. Weitkamp), gertjan.veldwisch@wur.nl (G. Jan Veldwisch), p.karimi@un-ihe.org (P. Karimi), c.defraiture@un-ihe.org (C. de Fraiture).

https://doi.org/10.1016/j.jag.2023.103418
Received 18 April 2023; Received in revised form 23 June 2023; Accepted 7 July 2023
Available online 22 July 2023

most mature and widely used (Maxwell et al., 2018; Sheykhmousa et al., 2020; Thanh Noi and Kappas, 2017). RF is popular for its ease of use and high accuracy (Belgiu and Drăguţ, 2016), while SVM is often chosen due to its ability to perform well with few training samples (Mountrakis et al., 2011). ANN is frequently used when detecting trends or patterns is difficult, and with the increase in computation power, it is being utilised more frequently (Abdolrasol et al., 2021). The k-NN classifier, although simple, has been found to compete with more complex classifiers in terms of performance (Abu Alfeilat et al., 2019). However, few studies compare two or more algorithms in the field of (smallholder) irrigation mapping.

Simple methods to use satellite data for classification are through single images or composites (Gella et al., 2021). Composites are widely used to generate cloud-free spatially consistent images from satellite time series and can be created based on summary measures extracted from the time series (Khatami et al., 2020), such as mean, minimum, or maximum pixel values. Vegetation phenology can be characterised by creating shorter composites, such as monthly or seasonal composites (Bey et al., 2020; Khatami et al., 2020; Kumar et al., 2022). However, using them could reduce the classification accuracies because they contain less information than, for example, time series data (Marín Del Valle and Jiang, 2022), although contrasting findings suggest that the opposite effect is also possible (Hasenbein et al., 2022). Alternatively, the temporal variation can be captured by calculating the geometric median, which preserves high-dimensional relationships between spectral bands, and three median absolute deviation statistics of temporal variation (Roberts et al., 2018, 2017). These composites and statistics have successfully been used in classifying irrigated croplands in Zimbabwe (Wellington and Renzullo, 2021) and seem promising for our study.

Enough studies have already investigated the effect of different machine learning algorithms or composites in land use classification. Bey et al. (2020) found high accuracy using the median composite with RF for mapping smallholder croplands in Mozambique, whereas Abubakar et al. (2020) achieved high accuracy in mapping maise fields in Nigeria with RF and SVM but used single images instead of composites. Furthermore, Bofana et al. (2020) compared four algorithms using combined seasonal input data but did not explore other composite lengths. However, to our knowledge, no study exists in which different algorithms and composite lengths are compared over the same study area. This study examines how different algorithms and composite lengths affect the accuracy of predicting irrigated agriculture in Mozambique. The research evaluates four classifiers (RF, SVM, ANN, and k-NN) and four composite lengths (12-monthly, 6-monthly, 3-monthly, and 2-monthly) and introduces "agreement maps" to show core areas of irrigated agriculture surrounded by an uncertainty zone. These maps can combine the strengths of multiple models and reduce the possibility of false positives. This unique method focuses on specific class distribution and classification certainty.

2. Materials and methods

We analyse the impact of different algorithms and composite lengths on the accuracy of irrigated agriculture in two stages (Fig. 1). Firstly, we test four algorithms and select the one with the highest accuracy. Secondly, we test this algorithm with different composite lengths, limiting each phase to one study area per province. We present maps of classifications and measures of accuracy for each combination of algorithm, composite length, and study area. A new method for consolidating the results by identifying hotspots is introduced. Table 1 summarises the different classifications, with each combination of algorithm and composite referred to as a distinct model (16 models in total).

2.1. Study area

This study was conducted in four areas in Mozambique: Chokwe and Xai-Xai in Gaza province and Manica and Catandica in Manica province (Fig. 2). These areas were chosen for their diverse agroecological characteristics and the presence of irrigated agriculture, including small-scale and large-scale systems. The case studies covered approximately 40x40 km in size.

Mozambique's rainy season occurs from November to April, with peak rainfall between December and February (Fig. 3). Chokwe receives 650 mm/year (Kajisa and Payongayong, 2011), Xai-Xai receives 950 mm/year (Brandt et al., 2009), and both Manica and Catandica receive 1100 mm/year (Gumbo et al., 2021; Weemstra et al., 2014). Irrigation occurs during the dry season, with two cycles occurring roughly from April to July and August to November.

In Manica province, the landscape is mountainous, with small streams serving as irrigation sources. Farmers redirect the water into earthen canals called "furrows" and use sprinkler irrigation, small pumps, and bucket irrigation. These systems are smaller than those in Gaza province and vary based on water availability. Horticultural crops are irrigated during the dry season, while maise is grown during the rainy season.

In Gaza province, there are both large- and small-scale irrigation systems along the banks of the Limpopo River. Flooding is a common practice, and pumps are used to access higher areas. Near Xai-Xai, there are irrigated areas with shallow groundwater tables that require drainage after the rainy season. Horticulture and maise are common crops in the irrigation season, while rice and maise dominate the rainy season.

2.2. Sampling design, labelling protocol, and field data sampling

Field reference data were collected from August to November 2020 using three different sampling strategies: random clustered, opportunistic, and regular clustered designs. A random clustered sampling design was initially used to minimise travel time but resulted in overlapping polygons and limited samples of irrigated agriculture. An opportunistic sampling design was used to gather more irrigated agriculture samples specifically, while a regular clustered design was used to prevent overlap and ensure sufficient polygon size for Sentinel pixels. The collected data was cleaned and analysed, resulting in 823 unevenly distributed polygons among different classes and areas. Polygons were digitally drawn from high-resolution images for hard-to-reach areas. Table 2 describes the classes following the ESA WorldCover definition (Zanaga et al., 2022), while Table 3 provides the number of polygons and total hectares per class and area. The classes cropland, grassland, shrubland, and tree cover were labelled in the field using Open Data Kit (ODK) Collect, a smartphone application that allows rapid and scalable field data collection (ODK collect, 2022).

The cropland classes (irrigated and rainfed) were distinguished by the period in which crops were actively grown, specifically during the rainy season (water is primarily supplied through rainfall) or the dry season (water is actively managed on the fields, either by applying or by draining water).

2.3. Input variables: Data collection and preprocessing – Digital Earth Africa

Satellite data for the four areas were collected within the Digital Earth Africa (DEA) 'sandbox', which provides access to Open Data Cube products in a Jupyter Notebook environment.[1] Sentinel-1 and 2 geomedian products (a robust high-dimensional statistic like the normal median that maintains relationships between spectral bands, DEA, 2021; Roberts et al., 2018) were generated at 10-meter resolution for four different composite lengths (one 12-monthly, two 6- monthly, four 3-

[1] Sandbox link and explanation can be found on https://docs.digitalearthafrica.org/en/latest/sandbox/index.html.

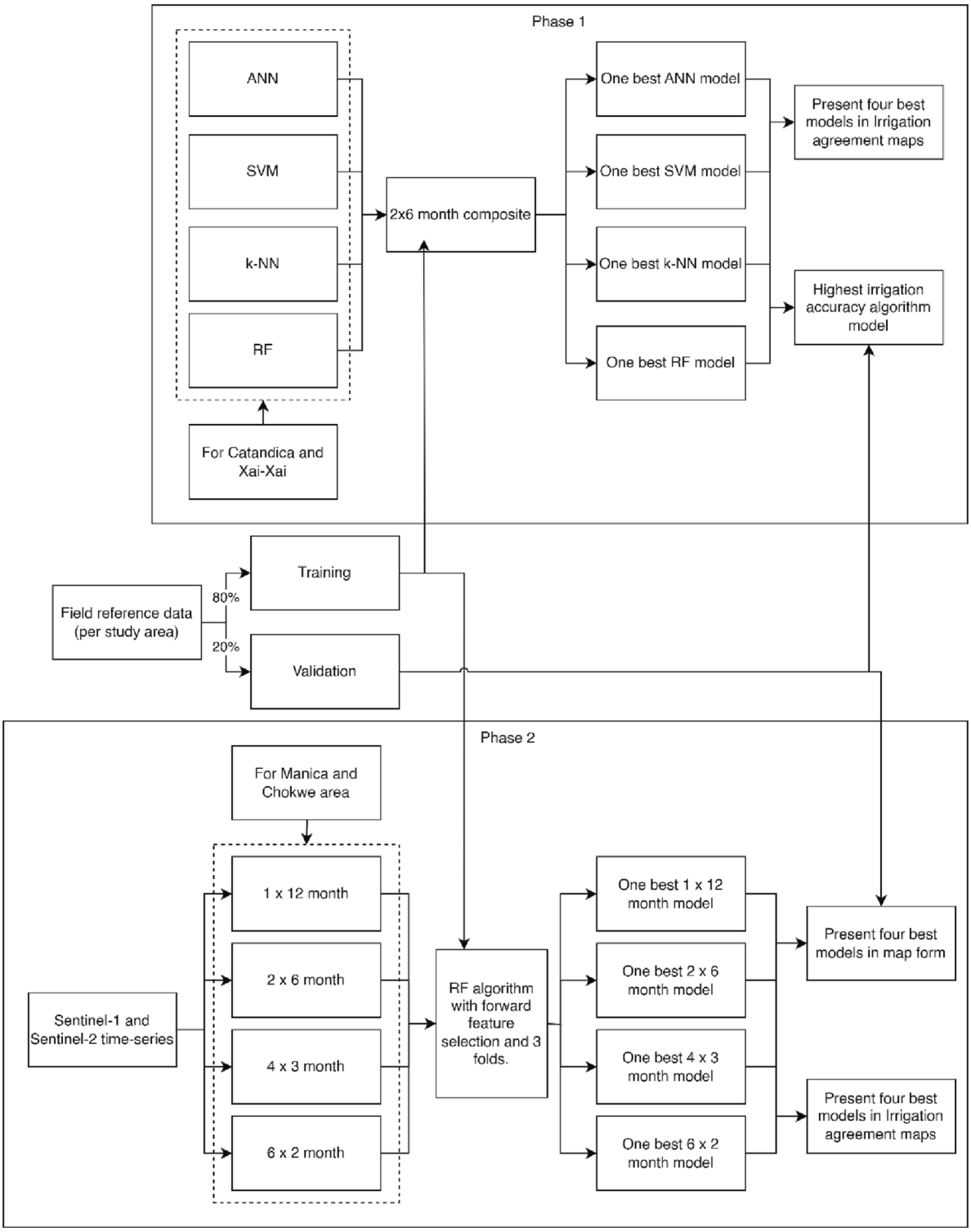


Fig. 1. Flow chart illustrating the two phases and methods used per phase.

monthly, and six 2-monthly), covering October 2019 – September 2020, corresponding to the hydrological year (wet and dry season). Images with more than 30% cloud cover (Sentinel 2) were filtered out. The specific Python scripts used for preparing the satellite data on DEA can be found at https://github.com/TimonWeitkamp/Mapping-irrigated-agriculture.

From Sentinel-2, we calculated the Normalised Difference Vegetation Index (NDVI), Bare Soil Index (BSI), and Normalised Difference Water Index (NDWI), using the DEA indices package for the Sentinel-2 composites (Wellington and Renzullo, 2021), while the Chlorophyll Index Red-Edge (CIRE) (Gitelson et al., 2005; Segarra et al., 2020) was calculated in R. Three second-order statistics (Median Absolute Deviations (MADs)) were also calculated, which are change statistics based on the geomedian: the Euclidean (EMAD, based on Euclidean distance), Spectral (SMAD, based on cosine distance), and Bray-Curtis (BCMAD, based on Bray-Curtis dissimilarity) MADs (Roberts et al., 2018). Wellington & Renzullo (2021) used these change statistics, as well as a few of the indices in their classification of irrigated areas, with success.

Table 1
Overview of the different models/classifications.

	Algorithm	Composite	Study areas
Phase 1	RF	2 × 6	Catandica & Xai-Xai
	SVM	2 × 6	Catandica & Xai-Xai
	k-NN	2 × 6	Catandica & Xai-Xai
	ANN	2 × 6	Catandica & Xai-Xai
Phase 2	RF	1 × 12	Manica & Chokwe
	RF	2 × 6	Manica & Chokwe
	RF	4 × 3	Manica & Chokwe
	RF	6 × 2	Manica & Chokwe

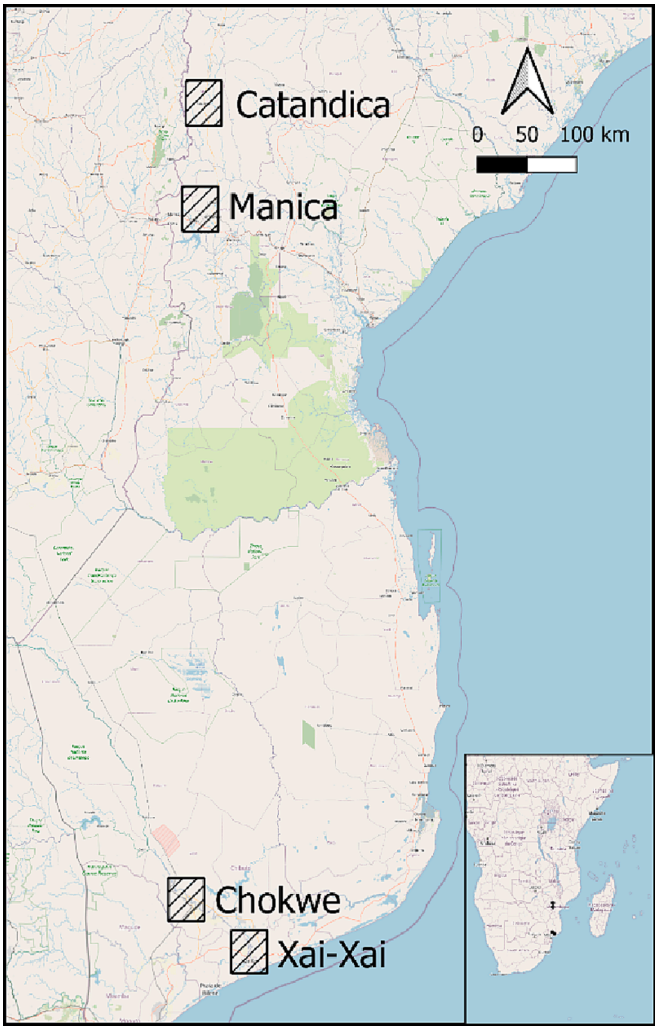


Fig. 2. The four study areas in Mozambique, from top to bottom: Catandica, Manica (Manica province), Chokwe, and Xai-Xai (Gaza province). See Annex 2 for detailed classifications per study area.

We used these indices and statistics to cover the different phases of croplands, from browning (BSI) to greening (NDVI, CIRE), the NDWI for water detection, while the MADs are suitable for change detection, particularly for irrigation (Wellington and Renzullo, 2021).

We also used Sentinel-1, specifically the VV and VH bands, and calculated the Radar Vegetation Index (RVI). These have also been used in recent agriculture mapping studies (Abubakar et al., 2020; Gella et al., 2021; Venot et al., 2021). The VV polarisation data is sensitive to soil moisture, whereas the VH polarisation data is more sensitive to volume scattering, which depends strongly on the geometrical alignment and vegetation characteristics. Therefore, VH data has a limited potential for estimating soil moisture compared to VV data but higher sensitivity to vegetation (Gao et al., 2018). The RVI can be used to separate soil from vegetation (Jennewein et al., 2022; Mandal et al., 2020). Additionally, the study area experiences frequent cloud cover for parts of the year, and the synthetic-aperture radar (SAR) data is less affected by cloud cover. As a result, the SAR composites of the cloudy seasons contain fewer missing observations and improve classification results, as radio-frequency radiation from SAR can penetrate through clouds.

All bands and indices were merged into one dataset, forming an 18-variable dataset (Table 4). This was done per composite length (4 lengths) and per area (4 areas).

2.4. Classification

2.4.1. Conceptual description of the machine-learning algorithms

We used four different algorithms, namely a radial support vector machine (SVM), random forest (RF), artificial neural networks (ANN), and k-nearest neighbours (k-NN). We used the *caret* package (Kuhn, 2008) to systematically compare different algorithms (svmRadial, rf, nnet, and knn, respectively) and composites in a standardised method in the free statistical software R Studio. The scripts can be found on GitHub https://github.com/TimonWeitkamp/Mapping-irrigated-agriculture.

Since our focus is on the application of the algorithms rather than the theoretical aspects of their design, we provide only a short description of each algorithm.

- SVM split the classes by fitting an optimal separating hyperplane (OSH) between classes using the training samples within feature space (i.e., all the pixel band values within the training sample) and to maximise the margins between OSH and the closest training samples (the support vectors) (Mountrakis et al., 2011).
- RF is an ensemble learning technique that generates many random decision trees that are then aggregated to compute a classification (Belgiu and Drăguţ, 2016).
- ANN design is based on the biological nervous systems, which is where their name comes from. An ANN is made up of neurons, which are organised in layers. The key characteristic of an ANN is that all neurons in one layer are connected to all neurons in all adjacent layers, and these connections have weights (Abdolrasol et al., 2021).
- The k-NN classifier is different from the other classifiers. Instead of producing a model, each unknown sample is directly compared against the original training data and is assigned to the most common class of the k training samples that are nearest in the feature space to the unknown sample (Maxwell et al., 2018).

2.4.2. Spatial folds and parameter settings

The polygon shapefiles and images were read into R, after which all pixel values for all variables were extracted. After extracting the pixel values, the field data was split into 80% training and 20% validation data using a fixed seed number (i.e., the same data used in each model's training and validation), stratified per landcover class. The *CreateSpacetimeFolds* from the *caret* package (Kuhn, 2019) was used to create three spatial folds, meaning all pixels within a polygon remain together in either the training or testing phase, instead of some pixels within the same polygon being used for training, and their neighbouring pixels being used for testing. This reduces spatial overfitting, i.e., it avoids over-optimistic models (Meyer et al., 2018). Five cross-validation folds were used during the training phase (*caret::ffs()*). These scripts can be found on GitHub.

The *caret::ffs()* function, or forward feature selection, first trains a model with two predictors using all possible pairs of predictor variables, after which the best initial model is kept. Iteratively, a new predictor is added to the model, and again the best combination is kept. This process stops when there is no increase in model performance. This function reduces the complexity of the model; however, combining all predictors takes time. Doubling the number of variables results in roughly four times as many sub-models to process.

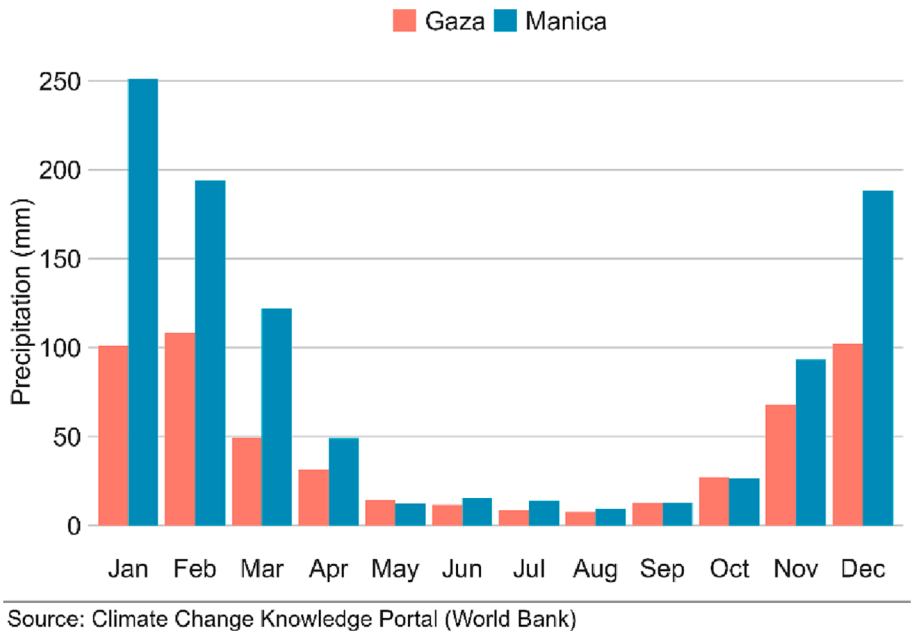


Fig. 3. Mean monthly precipitation (1991–2020) per province. Irrigation occurs during the dry season, with a first cycle roughly from April to July and a second from roughly August to November.

Table 2
Class descriptions.

Cropland irrigated	Croplands under management mainly during the dry season. Any active form of water management is considered, from drainage to application through buckets.
Cropland rainfed	Croplands under management mainly during the wet season
Tree cover	Natural vegetation comprises mainly trees and dense undergrowth.
Shrubland	Natural vegetation comprising of mainly low shrubs, grasses, and some scattered trees.
Grassland	Natural vegetation of primarily grass.
Wetland	Natural vegetation that is submerged part of the year (mainly during the rainy season and first part of the dry season).
Water	Water bodies and rivers.
Built-up area	Man-made surfaces and built-up areas, including bare areas such as sand (no vegetation).

All hyperparameters were tuned through the *tuneLength* (in *caret::ffs ()*) option, which generated five random tuning parameter combinations. A manual hyperparameter setting was considered but not used. The classification model with the highest overall accuracy was used to predict the entire extent of each site.

2.5. Accuracy/error assessment

We evaluated the performance of the models using a range of metrics, including overall map accuracy, user accuracy, and producer accuracy. These metrics were calculated using the unbiased accuracy assessment method described by Olofsson et al. (2014) and the *mapac* package in R (Pflugmacher, 2022).

To assess the models, we used a cross-validation approach, in which the training data was split into folds, and the model with the highest result was compared to 20% of the validation data (the same 20% in each run). The results for each model were then reported in a confusion matrix. It is important to note that the overall accuracy can be biased towards the most abundant class in the training data. Therefore, it is useful also to consider the user's and producer's accuracies, which provide more detailed information about the model's performance for a specific class.

2.6. Presentation of results

Using multiple models to assess irrigated agriculture is crucial, but defining boundaries can vary. To address uncertainty, "irrigation hotspots" can be identified as areas where irrigation is known to exist but cannot be accurately measured. "Agreement maps" combine model classifications, showing consensus on irrigation locations. A value of 4 out of 4 models signifies unanimous identification, while 1 out of 4

Table 3
Polygon distribution and size (hectares) per area and class of the collected field data.

	Catandica		**Manica**		**Xai-Xai**		**Chokwe**	
	# polygons	hectares	# polygons	hectares	# polygons	hectares	# polygons	hectares
Cropland irrigated	45	16,4	58	10,2	157	38,3	68	166
Cropland rainfed	34	10,9	32	7	19	5,8	48	40,4
Tree cover	9	148	19	104	9	37,2	15	12,5
Shrubland	25	89,5	20	11,3	28	26	104	187
Grassland	0	0	0	0	52	111	0	0
Wetland	0	0	0	0	6	27	12	144
Water	0	0	9	113	9	42,6	5	17,2
Built-up area	10	3,4	10	5,6	10	18,1	10	11,5
Total	123	268,2	148	251,1	290	306	262	578,6

Table 4
Overview of variables per composite time-length.

Group	Variable	Equation
Sentinel-2	Blue	
	Green	
	Red	
	Near Infrared (NIR)	
	Red-edge 1 (RE1)	
	Red-edge 2 (RE2)	
	Shortwave Infrared 1 (SWIR1)	
	Shortwave Infrared 2 (SWIR2)	
Indices S2	Normalised Difference Vegetation Index (NDVI)	(NIR - Red)/(NIR + Red)
	Normalised Difference Water Index (NDWI)	(NIR – SWIR1)/(NIR + SWIR1)
	Bare Soil Index (BSI)	((Red + SWIR1) - (NIR + Blue))/ ((Red + SWIR1) + (NIR + Blue))
	Chlorophyll index (CI)	(NIR / Red Edge 1) − 1
Temporal variation	3 MADS S2	See Roberts et al. (2018) and Wellington and Renzullo (2021) for more details on equations
Sentinel-1	VV	
	VH	
Indices S1	RVI	4 × VH / (VV + VH)

models means only one identified irrigation.

3. Results

In the first section, we explore the influence of the different algorithms (using only the 2 × 6-month composite). In the second section, we explore the influence of composite length on the visibility of irrigated agriculture (using only the rf algorithm).

3.1. Comparison: Algorithms

We use the 2 × 6-monthly composites to compare how well irrigated agriculture is classified using different algorithms for Catandica and Xai-Xai regions. This composite length is used because of the balance between a low number of parameters (i.e., computation time) and acceptable accuracies.

3.1.1. Accuracies and classifications for different algorithms

The results in Table 5 show the accuracies of various models that use different algorithms for classifying irrigated agriculture in two study areas. The knn algorithm had low user and producer accuracy (7–26%) in both areas but had higher overall accuracy due to its good performance in identifying tree cover in Catandica and grassland in Xai-Xai. The nnet and svmRadial algorithms had very high user and producer accuracy (95–99%) in Catandica, while the rf algorithm had reasonable user and producer accuracy (80–85%) in both areas, and the svmRadial algorithm had reasonable user and producer accuracy (75–85%) in Xai-Xai. The overall accuracies were higher than the class-specific accuracies, indicating that certain classes, such as dense and shrubland in Catandica and grassland in Xai-Xai, were classified better. The confusion matrices in Annex 1 shows that in Xai-Xai, irrigated agriculture was mainly confused with grassland and shrubland, while in Catandica, it was mostly confused with both light and tree cover (for the rf classification only).

Table 5
User, producer and overall accuracy for irrigated agriculture for the algorithm models.

		Accuracy		
		Producer's	User's	Overall
Catandica	knn	7.4 ± 0.6	25.8 ± 2.0	61.1 ± 0.5
	nnet	99.6 ± 0.3	97.7 ± 0.7	98.4 ± 0.2
	rf	80.3 ± 1.7	79.3 ± 1.9	93.7 ± 0.3
	svmRadial	93.5 ± 1.1	94.9 ± 1.0	98.1 ± 0.2
Xai-Xai	knn	10.6 ± 0.6	18.5 ± 1.1	41.9 ± 0.4
	nnet	85.8 ± 0.9	91.0 ± 0.8	91.6 ± 0.3
	rf	85.9 ± 0.9	86.2 ± 1.0	91.8 ± 0.3
	svmRadial	74.3 ± 1.0	84.8 ± 1.0	85.3 ± 0.3

Fig. 4 demonstrates that while the nnet and svmRadial algorithms have similar levels of accuracy, they produce different maps of irrigated agriculture. The nnet algorithm shows more irrigated agriculture spread over the whole area, whilst the svmRadial algorithm only shows irritation in clusters. While the rf and svmRadial algorithms have different levels of accuracy, the maps they produce are similar. The knn algorithm greatly overestimates the extent of irrigated agriculture, with almost the entire map showing this class except for areas of tree cover in the bottom left corner. All four algorithms also incorrectly classify trees in Catandica town (located in the centre of the map, see Annex 2 for more details) and some rock outcroppings (not present in the training data) as irrigated agriculture.

In Xai-Xai, the rf and nnet algorithms have similar levels of accuracy, and their classified maps are also similar (Fig. 5). However, both of these algorithms, as well as the svmRadial algorithm, incorrectly classify many individual trees in towns (located in the bottom right quadrant of the map, see Annex 2 for more details) and groups of trees in predominantly rainfed agriculture areas (on the east and west sides of the map, outside of the Limpopo river valley) as irrigated areas. The maps produced by these three algorithms show many areas of irrigated agriculture along the edges of the valley and the river. In contrast, the map produced by the knn algorithm shows no clear structures that follow the landscape.

3.1.2. Irrigation agreement maps

The knn algorithm tends to overestimate the area of irrigated agriculture, making it unsuitable to use in agreement maps for visualising hotspots accurately. Consequently, we will exclude its results and only consider the outcomes from the remaining three algorithms. By overlaying the estimated maps from these algorithms, which identify the irrigation class, in an 'agreement map,' we can identify hotspots (Fig. 6).

In the top inset map (A), smallholder irrigation is near Catandica's urban region, correctly classified as irrigation by all three algorithms. However, the algorithms wrongly classify most of the urban trees as irrigation, and their boundaries differ slightly, leading to some areas with uncertainty. We call the pixels where all models agree (3/3 in this case) the *core areas* and the pixels surrounding these core areas the *uncertainty zone*. In the bottom inset map (B), the three algorithms accurately identify most of a tea plantation as irrigated agriculture, but some minor patches are classified differently by one or two algorithms. The knn algorithm, not included in this figure, classified all the surrounding grasslands as irrigated areas (Fig. 4), overestimating the extent and location of irrigated agriculture.

In Xai-Xai (Fig. 7), we excluded knn results from the agreement map. In area A (top inset map), smallholder irrigated fields have clusters of '3 models', indicating agreement between the results, but with fewer certain areas in between. The bottom right part of area A, an urban area (Xai-Xai), has most of the trees misclassified as irrigated areas. Area B (bottom inset map) shows a large, irrigated rice scheme (Hubei-Gaza Rice project). There is a major cluster of irrigated agriculture recognised by all models in the centre of this map, but the remaining fields are only identified by one or two of the algorithms.

The main overview map also shows that there are a lot of irrigated areas in the northeast quadrant, which are mostly misclassified pixels (1 model); this area has a higher elevation (+20 m) than most of the irrigated fields (which are in the Limpopo valley), where we primarily find rainfed agriculture, small patches of tree cover, shrubland, and urban areas.

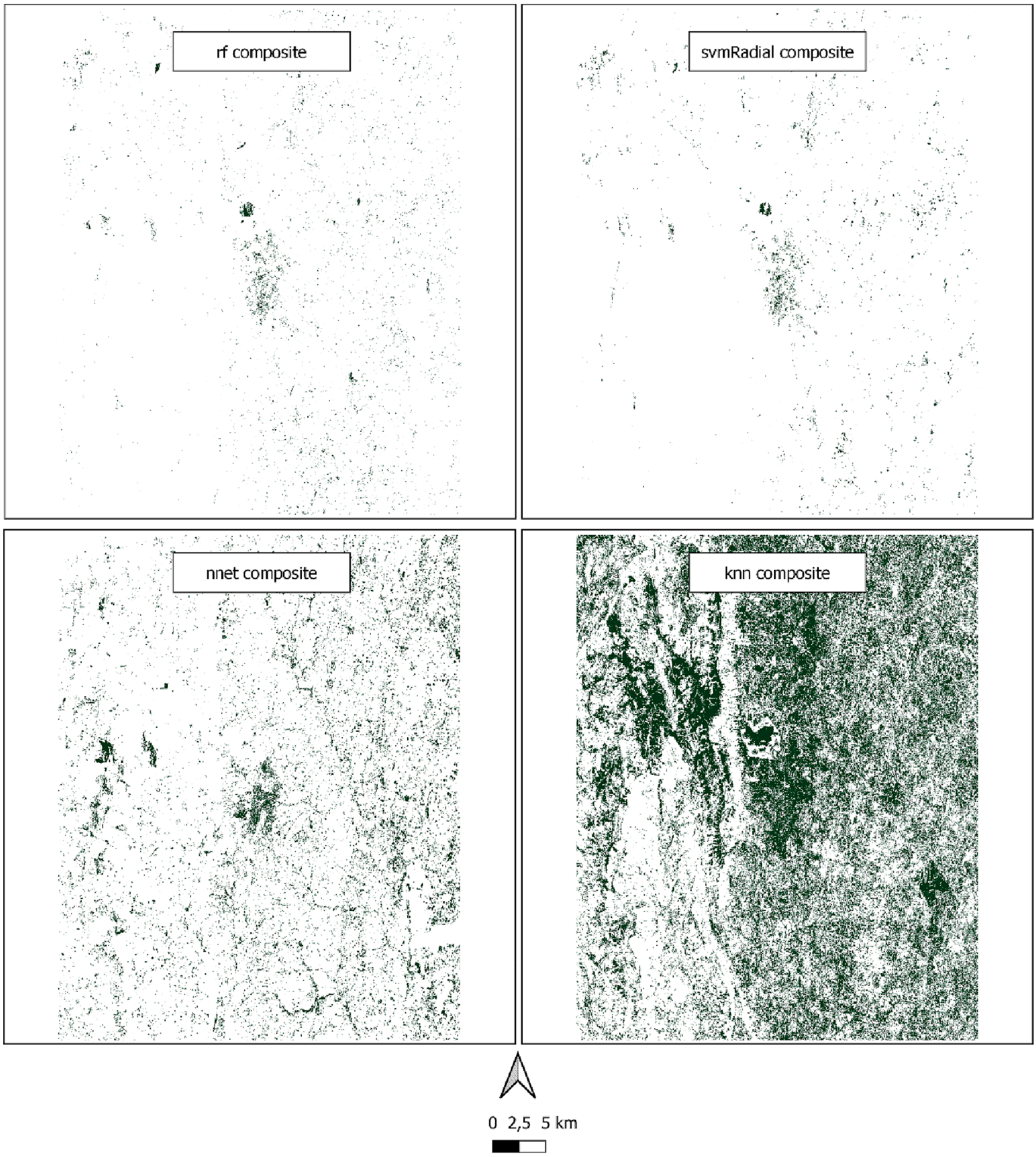


Fig. 4. Extent of irrigated agriculture per algorithm (tlbr: rf, svmRadial, ann, knn) for Catandica.

Table 6 shows hotspot accuracy and classified hectares for Catandica and Xai-Xai, with three categories based on the agreement between models: 3 models refer to three models classifying the same pixel as irrigated agriculture. The table shows that 3 models pixels are almost 100% correctly classified as irrigated agriculture, indicating high confidence in the core hotspots. However, there is an uncertainty zone surrounding the core areas. In Catandica, the 2 models ring is still accurate, while in Xai-Xai, only two-thirds of the pixels were accurately classified. Pixels identified by only one model are usually incorrect and can be excluded from final assessments.

3.2. Comparison: Composite lengths

Here we present the results of the different composite lengths using the rf. We used this algorithm because of its high computation speed, ease of use, and widespread use within the community.

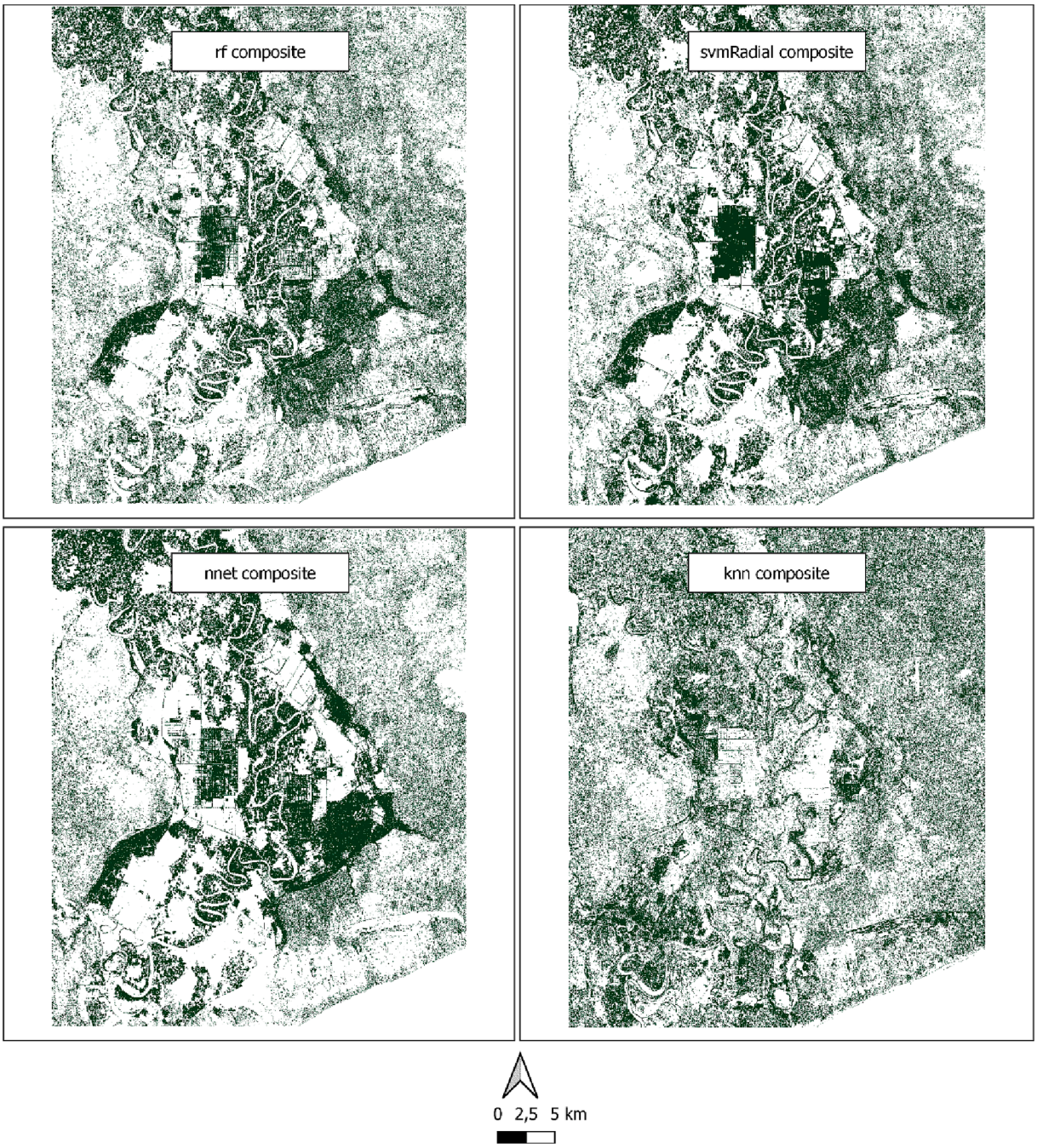


Fig. 5. Extent of irrigated agriculture per algorithm (tlbr: rf, svmRadial, nnet, knn) for Xai-Xai.

3.2.1. Accuracies and classifications for different algorithms

Table 7 displays the accuracies of various models that used different composite lengths to classify irrigated agriculture in two study areas. All models had high overall accuracies (above 95%). A single 12-month composite may not be sufficient to capture the differences between irrigated agriculture, rainfed agriculture, and shrubland in a complex landscape, such as the one found in Manica. This composite performs better in the slightly less complex landscape of Chokwe. Based solely on overall accuracy, Chokwe should be classified using the 2 × 6-month composites, while Manica should be classified using the 6 × 2-month composites. However, doubling the number of variables results in roughly four times as many sub-models to process, with only a limited increase in accuracy. Additionally, accuracy alone is insufficient to base conclusions on, as discussed in Section 3.1.

Confusion of irrigated agriculture in Chokwe was mostly with shrubland in all models (Annex 1). In Manica, irrigated agriculture was confused with several classes, primarily rainfed agriculture, followed by shrubland.

Fig. 8 shows the extents of irrigated agriculture for the four composites for Manica. At first glance, the four results seem similar, with irrigated agriculture following the rivers and slopes of the mountains. However, the urban area of Messica (located at the bottom centre of the map, see Annex 2 for more details) contains trees that have been misclassified as irrigated agriculture. Compared to Chokwe, Manica shows more small-scale irrigation spread out over the landscape.

Fig. 9 shows the similarities in the extent of irrigated agriculture in Chokwe and reveals that most of the fields at the head end of the Chokwe irrigation scheme are classified as irrigated agriculture and are actively

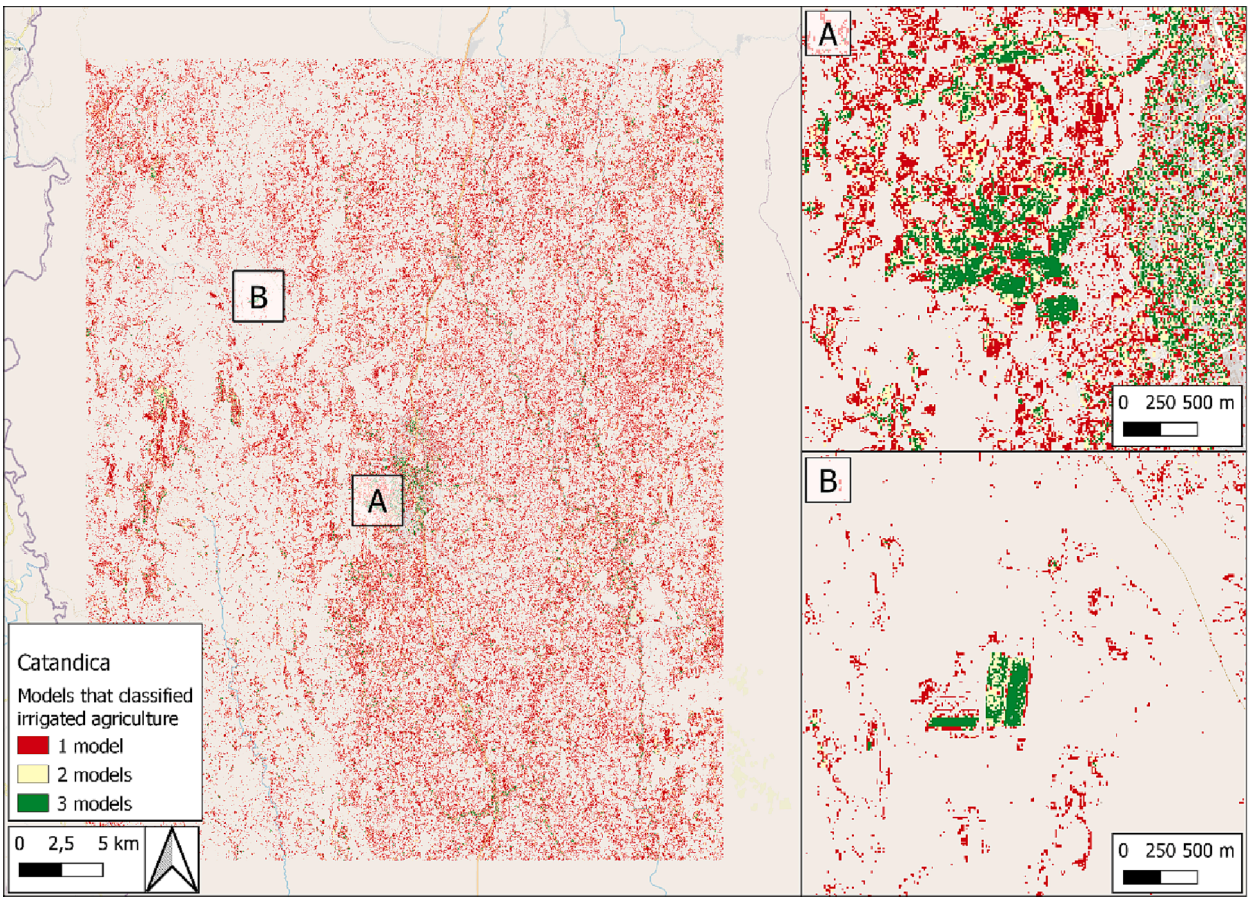


Fig. 6. Map of Catandica shows how many models classified a pixel as irrigated agriculture and two zoom-ins of a smallholder irrigation scheme (A) and part of the larger tea plantation (B). The values in the legend show how many models classified a pixel as irrigated agriculture: 3 means agreement in 3 models.

cultivated, while the tail end shows less irrigated agriculture – this reflects the actual situation well. The 3-month and 2-month composites follow the same trends but show a smaller overall area of irrigated agriculture. The 6-month composite stands out from the other three in its lower misclassification of shrubland in the map's top right and bottom left parts. The other composites show small clusters of irrigated agriculture in these areas, which are not present in the 6-month composite. The urban area of Chokwe (located at the centre of the map, see Annex 2 for more details) hardly shows any irrigated agriculture, similar to the other three study areas.

3.2.2. *Irrigation agreement maps*

In the main map of Manica (Fig. 10), we can see irrigation occurring in riverbeds and near mountains, with some large clusters of fields as well as many small patches. Area A (top inset map) shows an area with two known, clearly delineated smallholder irrigation schemes. Some core areas (4 models) are surrounded by areas that gradually change from 3 models to 1 model in a short distance, the uncertainty zone. Area B (bottom inset map) focuses on a few centre pivots (circular shapes). It shows that only parts of these fields are labelled with 4 models – an agreement by all four models –but as all pixels covering the field are irrigated by the centre pivot, we would expect all pixels of those fields to be labelled irrigation by all four models. If we had used only one classification, these areas would not have been recognisable as centre pivots.

Chokwe (Fig. 11) shows a similar pattern of core area and uncertainty zone. The map clearly shows the large-scale Chokwe irrigation scheme along the Limpopo River's south bank and some smaller schemes on the north bank, such as area A (top inset map). It also shows that some of the models have identified irrigated agriculture on islands in the river (1 model), which is certainly possible but may be natural vegetation that has been misclassified. This area also contains clusters of trees in predominantly rainfed areas that have been misclassified as irrigated agriculture (1 model). Area B (bottom inset map) box highlights part of the Chokwe irrigation scheme, of which we know only part is still actively used.

Table 8 summarises the accuracy of the classification of irrigated agriculture in Manica and Chokwe using different composite models. In Manica, the 3 and 4 models agreement achieved 100% accuracy, while the 1 model and 2 models (uncertainty zone) had lower accuracy rates of 1.40% and 64.20%, respectively. In Chokwe, the 4 models achieved 100% accuracy, while the 1 model and 2 models had accuracy rates of 1.40% and 84.60%, respectively.

4. Discussion and recommendations

We examined how different composite lengths and algorithms affect the accuracy of remote sensing-based models in identifying irrigated agriculture in four distinct study areas. Our analysis of 16 models revealed that the composite length and algorithm choice can considerably impact the results. Therefore, it is necessary to integrate the results of various models to account for model-specific biases. The following sections discuss our key findings.

4.1. *Algoruthim*

Our study found that the choice of algorithm can greatly impact the accuracy of remote sensing-based models in identifying irrigated agriculture. Our experiments showed that ANNs, SVMs, and RFs effectively classified irrigated areas. However, there was no straightforward "best" algorithm, as all achieved user, producer, and overall accuracies ranging

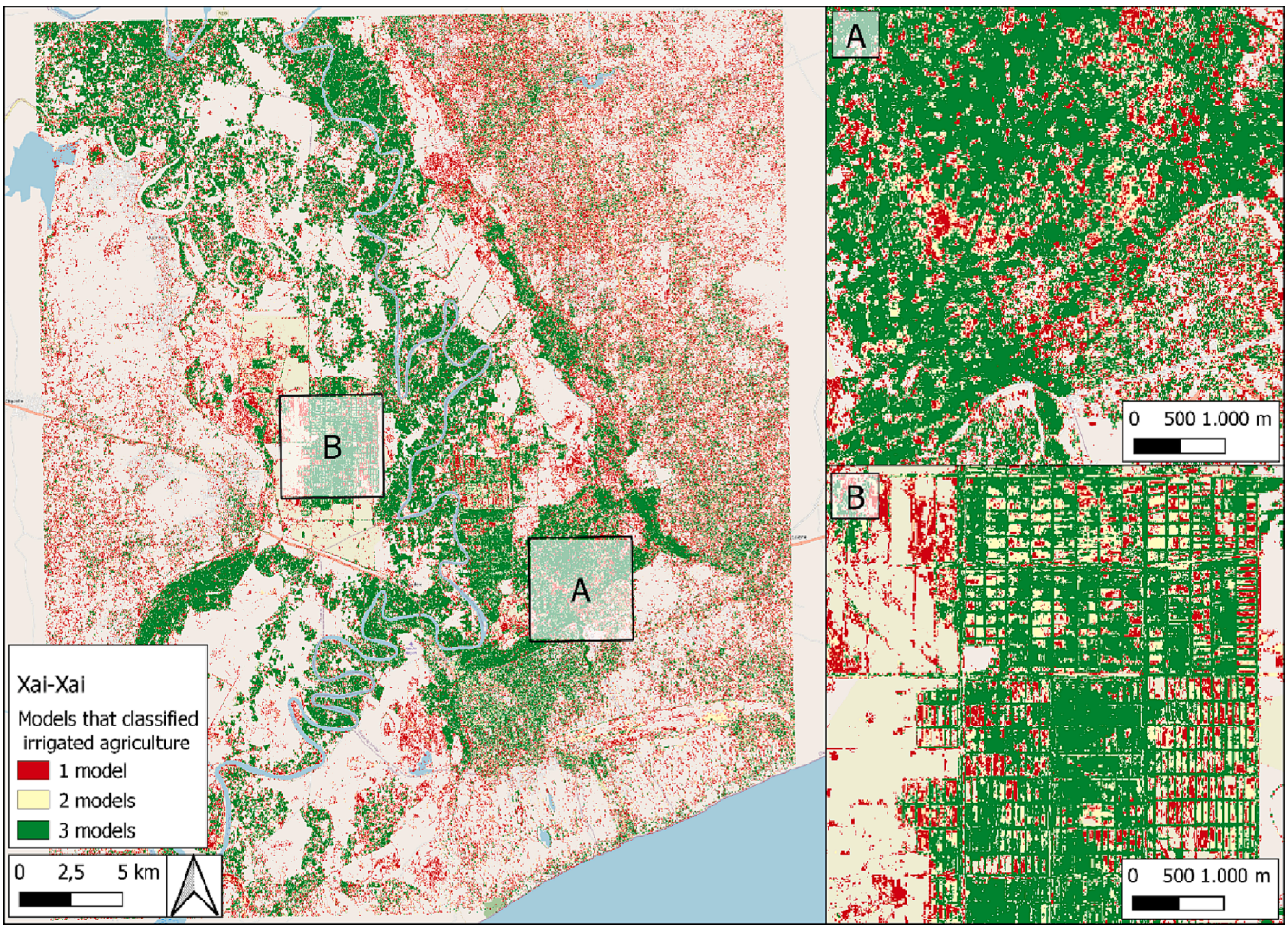


Fig. 7. Map of Xai-Xai showing how many of the models classified a pixel as irrigated agriculture, and two zoom-ins of a smallholder irrigation scheme (A) and part of the large rice irrigation scheme (B). The values in the legend show how many models classified a pixel as irrigated agriculture: 3 means agreement in 3 models.

Table 6
Accuracy of hotspot values and total number of hectares classified.

			Training data			
		Agreement	# pixels correctly classified	Total # pixels classified	% correctly classified	Total hectares
Algorithms	Catandica	1 model	26	242	11%	23,124
		2 models	189	200	95%	6660
		3 models	2145	2145	100,0%	1205
	Xai-Xai	1 model	355	2255	16%	20,872
		2 models	839	1315	64%	17,744
		3 models	4656	4704	99%	26,537

Table 7
User, producer, and overall accuracy for irrigated agriculture for the algorithm models.

		Accuracy		
		Producer's	User's	Overall
Catandica	knn	98.9 ± 0.2	98.0 ± 0.2	96.1 ± 0.2
	nnet	98.0 ± 0.2	96.2 ± 0.3	95.9 ± 0.2
	rf	97.5 ± 0.3	95.7 ± 0.3	94.9 ± 0.2
	svmRadial	99.7 ± 0.1	99.3 ± 0.1	98.0 ± 0.1
Xai-Xai	knn	73.9 ± 2.0	74.8 ± 2.3	94.5 ± 0.3
	nnet	94.8 ± 1.2	91.8 ± 1.5	98.8 ± 0.1
	rf	92.8 ± 1.3	90.2 ± 1.6	98.2 ± 0.2
	svmRadial	90.4 ± 1.5	91.2 ± 1.6	98.3 ± 0.2

from 80% to 95%.

Based on the agreements and differences observed between the different algorithm maps, we recommend using at least three algorithms and focusing on hotspots to consider both the heterogeneous and homogeneous parts of the landscape in the model. Additional research could assess the algorithmic sensitivity to the diverse methods employed in farmer-led irrigation. This could be accomplished by analysing the performance of the models in scenarios where training data from these farmers are either excluded or included, allowing for a comparison between the two.

4.2. Composite length

The study found that composite length is crucial in accurately identifying irrigated agriculture in diverse landscapes. Shorter

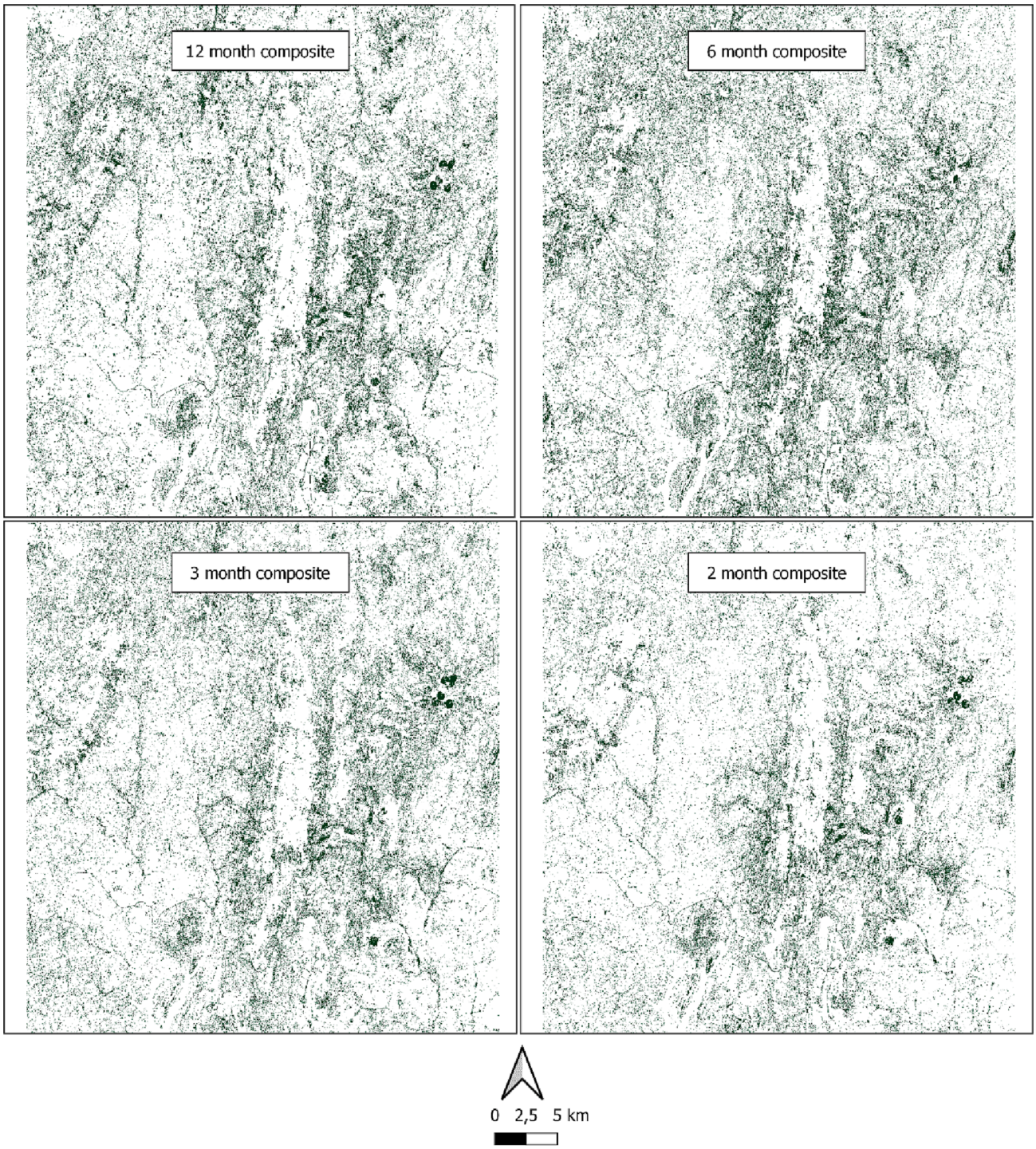


Fig. 8. Extent of irrigated agriculture per composite length (tlbr: 12, 6, 3 & 2 months) for Manica.

composites are better for complex landscapes, while longer composites are sufficient for homogeneous ones. It is important to consider composite length when creating remote sensing-based models and to focus on hotspots. The 6-month and 3-month composites are promising options due to their lower computation time and data size. Using agreement maps incorporating multiple composites enhances the visibility of features like centre pivots.

Further investigation could centre on determining the optimal selection of months to include or exclude in the composite. In the current research, a 12-month dataset was used, distributed across various composite lengths. However, it is worth exploring the possibility of achieving comparable results by solely utilising the dry season months. This approach may offer the advantage of requiring less data and reducing model complexity.

4.3. Model agreement method: Hotspot maps

Our analysis revealed that combining models with different composite lengths and algorithms can improve the accuracy of identifying irrigated agriculture. Hotspot maps provide valuable information for decision-making and prioritising targeted field surveys or management decisions. For complex landscapes with dynamic and heterogeneous

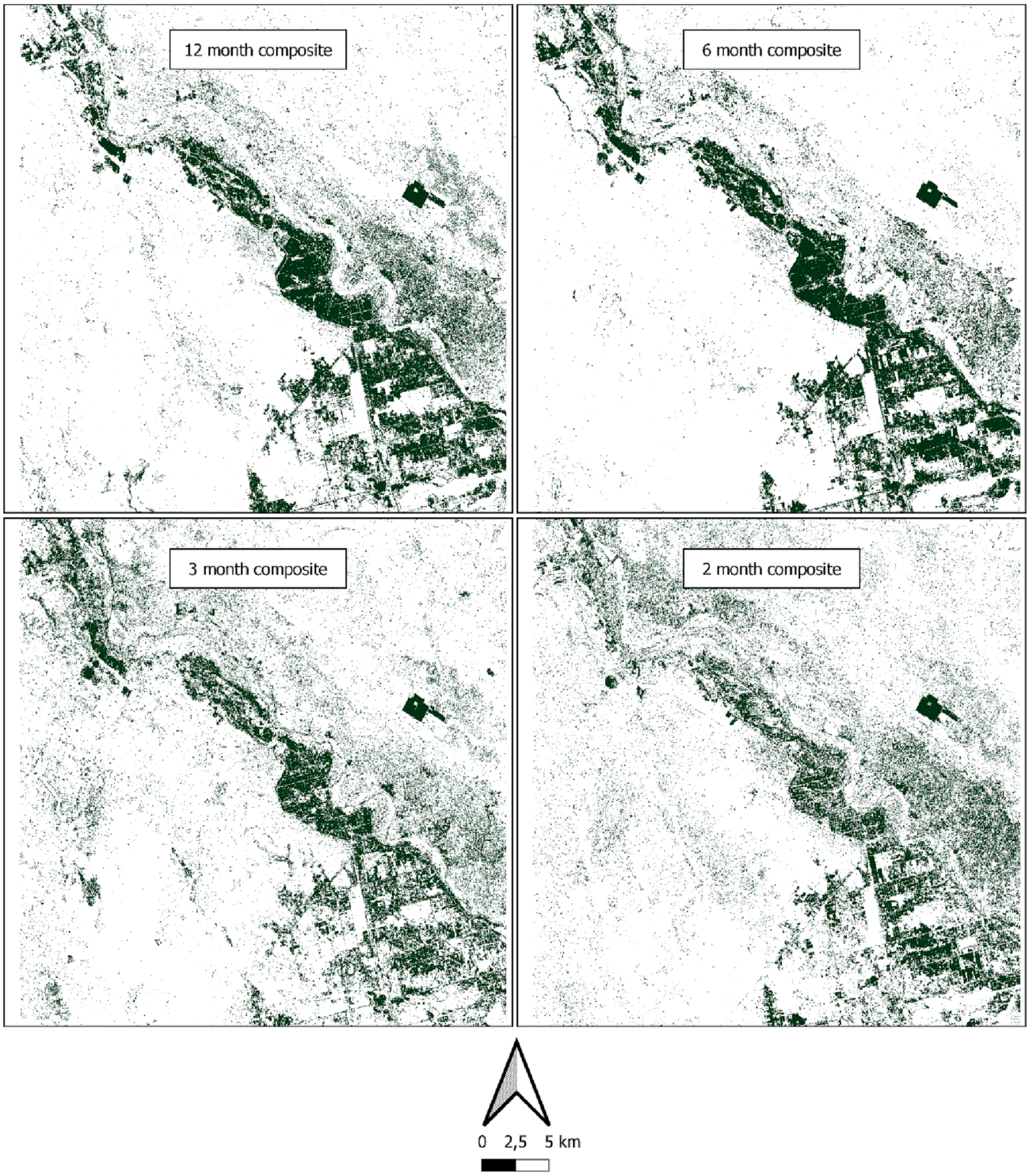


Fig. 9. Extent of irrigated agriculture in green per composite length (tlbr: 12, 6, 3, 2 months) for Chokwe.

classes, combining models can provide better insights into the core areas of hotspots. The example of the Chokwe irrigation scheme shows that although a vast area is equipped for irrigation, considerable areas are not irrigated. That irrigated area can be smaller or larger depending on the model used; however, the hotspot maps show the minimum irrigated area with greater certainty. We recommend including at least three models to improve the accuracy of the core areas.

4.4. Reflection: Other aspects that likely influenced our results

Our findings suggest using multiple composite lengths to capture the dynamic nature of irrigated agriculture. Shorter composites (quarterly or bi-monthly) are necessary to identify highly dynamic classes like irrigated agriculture accurately, while longer composites (annual or seasonal) may be more effective for stable classes like tree cover and urban areas. Focusing on specific periods, such as the end of the rainy season and the start of the dry season, can also help capture changes in irrigation and vegetation patterns.

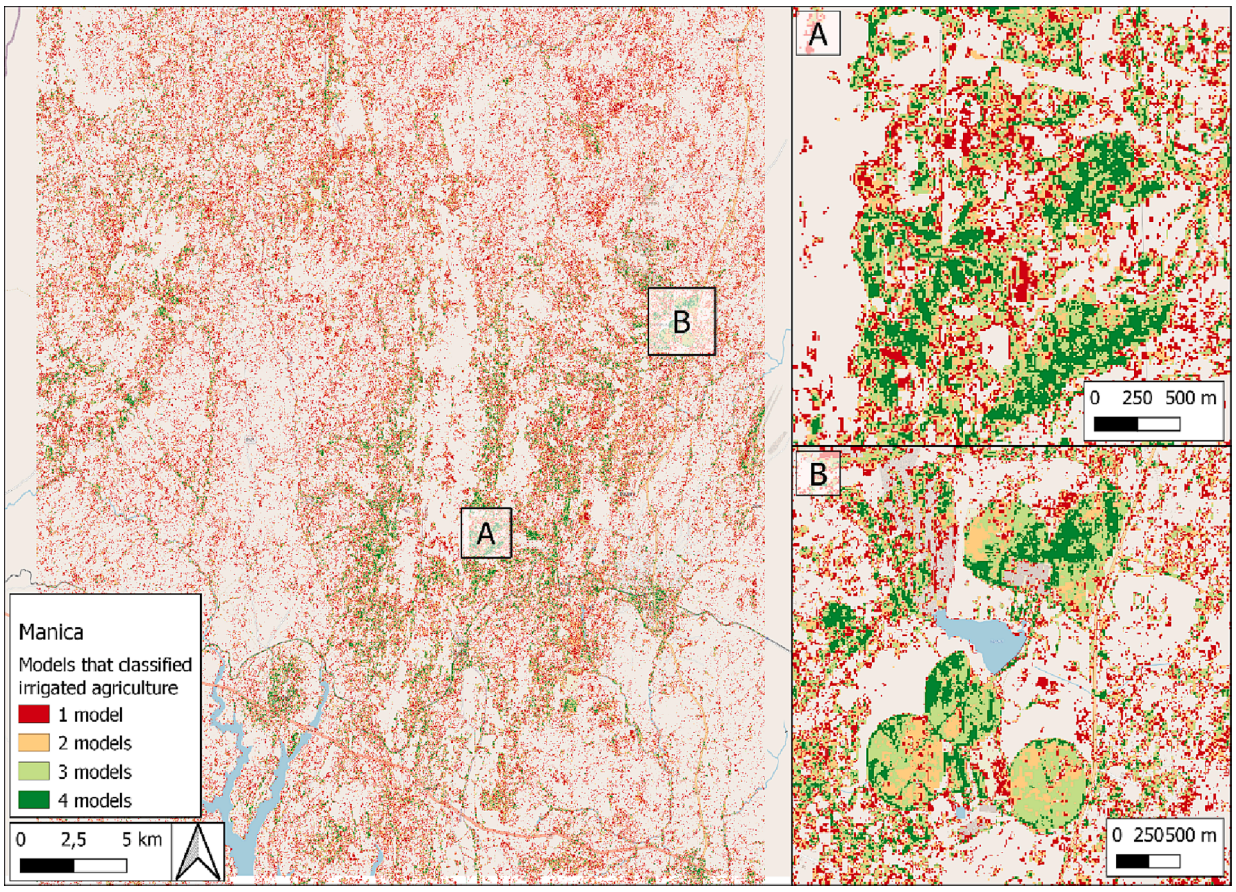


Fig. 10. Map of Manica shows how many models classified a pixel as irrigated agriculture, and two zoom-ins of a smallholder irrigation scheme (A) and part of an irrigated estate by means of centre pivots (circular shapes within area B). The values in the legend show how many models classified a pixel as irrigated agriculture: 4 means agreement in 4 models.

We chose variables and composite statistics based on previous studies on mapping irrigated agriculture (Elwan et al., 2022; Lebourgeois et al., 2017; Wellington and Renzullo, 2021; Xie et al., 2019). Our aim was not to determine the "best" variables or statistics, as this is context-dependent. Different combinations of variables were important for different runs, and the geomad statistic was sufficient to show the influence of composite length and algorithm choice. Although these methods have the potential to improve accuracies further, our results were already high, which raises the question of whether more effort should be focused on field data collection or improving models at optimal performance.

The training data was collected during the dry season, and the labels for rainfed agriculture were based on leftover maise stalks and shrubland. The training data may have been imbalanced, with fewer samples for less prevalent classes. The regular clustered sampling design was used due to the tradeoff between complete random data collection and travel time. The data collected through abandoned strategies were still used, but the overall size was small. Some areas, such as bare rocks and sand, were included in the built-up class, which resulted in inaccurate classification by the algorithms.

The study was conducted over four areas chosen because of their differences in weather, topography, and agricultural uses. We hoped to capture various irrigation circumstances but undoubtedly missed some practices. Hence the findings on composite and algorithm use may be helpful for some areas of Mozambique but less so in areas further away. For example, Wellington & Renzullo (2021) found that the annual composite was optimal for classifying irrigated agriculture in Zimbabwe.

5. Conclusion

We investigated the impact of different composite lengths and algorithms on the accuracy of remote sensing-based models for identifying irrigated agriculture in four sub-Saharan African study areas. Our findings showed that the choice of algorithm and composite length can considerably affect model outcomes. We found that SVMs, RFs, and ANNs effectively classified irrigated areas, while the k-nearest neighbour algorithm was ineffective in this task. Shorter composite lengths, such as 2-monthly or 3-monthly composites, were more effective for identifying irrigated agriculture in complex and dynamic landscapes, while longer composite lengths were more appropriate for stable classes.

Our study also highlighted the importance of considering hotspots and agreement maps when identifying irrigated agriculture. Combining the outputs of various models into agreement maps can provide better insights into the core areas and uncertainty zones of hotspots. These findings can help decision-makers remotely situated to understand irrigation dynamics better.

5.1. Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in order to improve readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

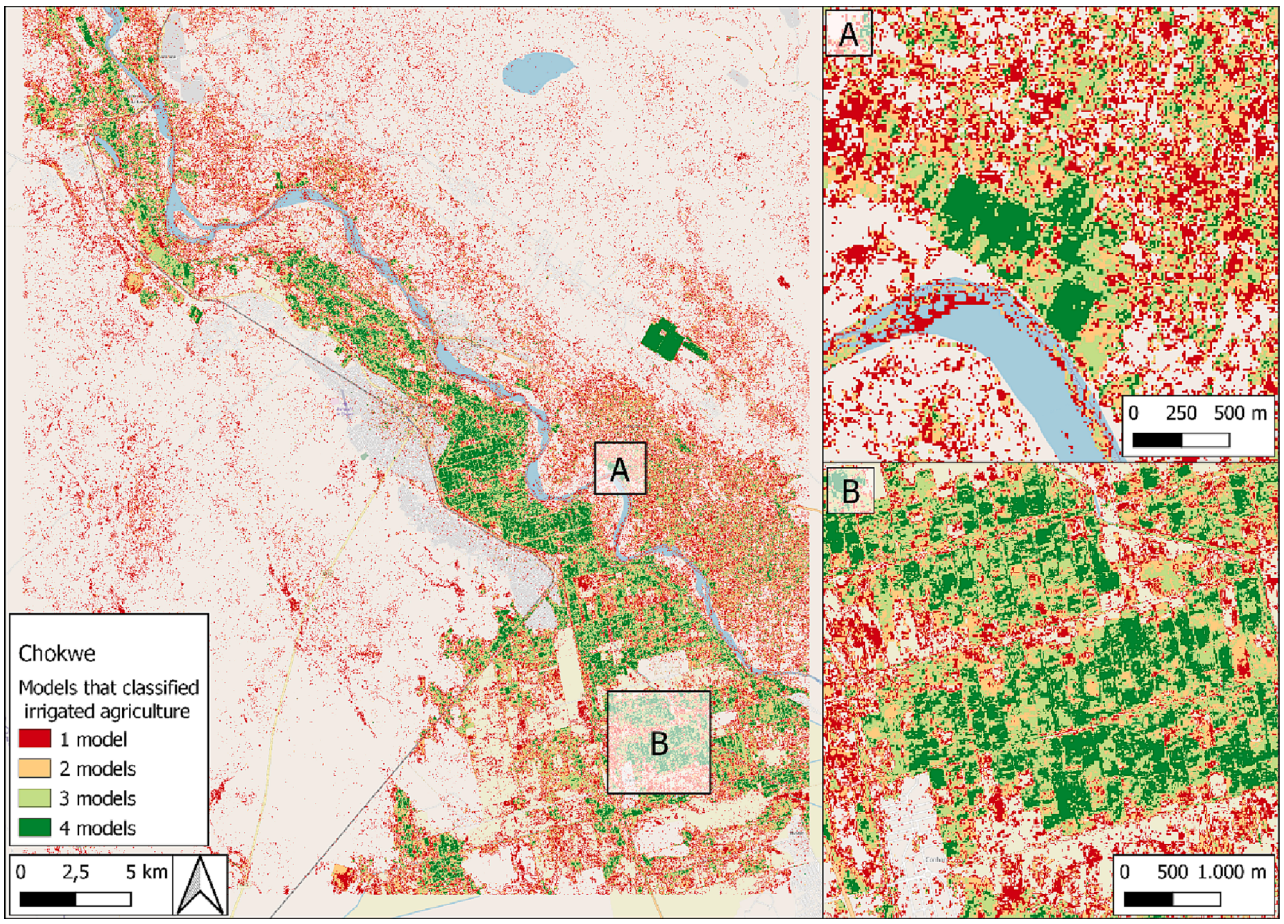


Fig. 11. Map of Chokwe showing how many of the models classified a pixel as irrigated agriculture, and two zoom-ins of a smallholder irrigation scheme (blue area) and part of the large-scale Chokwe irrigation scheme (red area). The values in the legend show how many models classified a pixel as irrigated agriculture: 4 means agreement in 4 models. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 8
Accuracy of hotspot values and total number of hectares classified.

			Training data			
		Agreement between models	Irrigated agriculture	Total pixels classified	Irrigation correctly classified	Total hectares
Composites	Manica	1 model	6	440	1,40%	20,444
		2 models	61	95	64,20%	9565
		3 models	396	396	100,00%	5795
		4 models	1259	1259	100,00%	3289
	Chokwe	1 model	3	212	1,40%	16,866
		2 models	22	26	84,60%	8212
		3 models	370	370	100,00%	6736
		4 models	18,199	18,199	100,00%	5389

CRediT authorship contribution statement

Timon Weitkamp: Conceptualization, Methodology, Formal analysis, Writing – original draft, Visualization. **Gert Jan Veldwisch:** Conceptualization, Writing – review & editing. **Poolad Karimi:** Conceptualization, Writing – review & editing. **Charlotte de Fraiture:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

Annex 1: Confusion matrices

Chokwe

Comparison composite length - Chokwe 12m rf composite

	Reference class (sample counts)							Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	262	-	1	2	-	0	0	265	265	98.87 ± 0.65	98.87 ± 0.65	98.87 ± 0.46
Dv	-	319	6	123	-	0	2	450	344	92.73 ± 1.36	70.89 ± 2.14	80.35 ± 1.47
Ia	-	7	3634	45	21	0	1	3708	3675	98.88 ± 0.17	98.00 ± 0.23	98.44 ± 0.14
Lv	3	17	13	4324	61	0	26	4444	4643	93.13 ± 0.33	97.30 ± 0.24	95.17 ± 0.21
Ra	-	1	17	87	935	0	12	1052	1029	90.86 ± 0.86	88.88 ± 0.97	89.86 ± 0.65
Wa	0	0	0	0	0	381	0	381	381	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
We	0	0	4	62	12	0	3050	3128	3091	98.67 ± 0.20	97.51 ± 0.28	98.09 ± 0.17

Overall accuracy = 96.11 ± 0.16

Comparison composite length - Chokwe 6m rf composite

	Reference class (sample counts)							Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	268	-	-	1	1	0	0	270	268	100.00 ± 0.00	99.26 ± 0.52	99.63 ± 0.26
Dv	-	337	1	108	-	0	0	446	347	97.12 ± 0.89	75.56 ± 2.03	84.99 ± 1.33
Ia	-	3	3680	18	4	0	0	3705	3692	99.67 ± 0.09	99.33 ± 0.13	99.50 ± 0.08
Lv	-	7	6	4375	57	0	6	4451	4550	96.15 ± 0.26	98.29 ± 0.19	97.21 ± 0.16
Ra	-	-	5	43	988	0	1	1037	1050	94.10 ± 0.70	95.27 ± 0.66	94.68 ± 0.48
Wa	0	0	0	0	0	381	0	381	381	100.00 ± 0.00	100.00	100.00
We	0	0	0	5	0	0	3122	3127	3129	99.78 ± 0.08	99.84 ± 0.07	99.81 ± 0.06

Overall accuracy = 98.02 ± 0.11

Comparison composite length - Chokwe 3m rf composite

	Reference class (sample counts)							Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	272	-	-	7	1	0	0	280	274	99.27 ± 0.51	97.14 ± 1.00	98.19 ± 0.57
Dv	-	316	12	120	1	0	4	453	348	90.80 ± 1.50	69.76 ± 2.16	78.90 ± 1.49
Ia	-	7	3533	84	56	0	10	3690	3624	97.49 ± 0.25	95.75 ± 0.33	96.61 ± 0.21
Lv	2	22	52	4261	85	0	20	4442	4635	91.93 ± 0.36	95.93 ± 0.30	93.89 ± 0.23
Ra	-	1	21	123	942	0	2	1089	1088	86.58 ± 0.96	86.50 ± 1.04	86.54 ± 0.71
Wa	0	0	0	0	0	381	0	381	381	100.00 ± 0.00	100.00	100.00
We	0	2	6	40	3	0	3018	3069	3054	98.82 ± 0.19	98.34 ± 0.23	98.58 ± 0.15

Overall accuracy = 94.92 ± 0.18

Comparison composite length - Chokwe 2m rf composite

	Reference class (sample counts)							Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	266	-	-	3	4	0	0	273	267	99.63 ± 0.37	97.44 ± 0.96	98.52 ± 0.52
Dv	-	337	8	129	-	0	0	474	362	93.09 ± 1.30	71.10 ± 2.08	80.62 ± 1.42
Ia	-	1	3546	124	13	0	1	3685	3619	97.98 ± 0.23	96.23 ± 0.31	97.10 ± 0.20
Lv	1	24	42	4281	61	0	34	4443	4618	92.70 ± 0.35	96.35 ± 0.28	94.49 ± 0.22
Ra	-	-	23	47	977	0	1	1048	1055	92.61 ± 0.77	93.23 ± 0.78	92.91 ± 0.55
Wa	0	0	0	0	0	382	0	382	382	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
We	0	0	0	34	0	0	3092	3126	3128	98.85 ± 0.19	98.91 ± 0.19	98.88 ± 0.13

Overall accuracy = 95.9 ± 0.17

Manica

Comparison composite length - Manica 12m rf composite

	Reference class (sample counts)						Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	Map	Ref.	Producer's	User's	F-score
Bu	150	-	1	2	1	0	154	153	98.04 ± 1.10	97.40 ± 1.28	97.72 ± 0.85
Dv	-	2286	13	11	1	0	2311	2333	97.99 ± 0.28	98.92 ± 0.22	98.45 ± 0.18
Ia	-	20	264	20	49	0	353	357	73.95 ± 1.96	74.79 ± 2.31	74.37 ± 1.51
Lv	-	27	40	204	57	0	328	272	75.00 ± 2.31	62.20 ± 2.68	68.00 ± 1.86
Ra	3	0	39	35	155	0	232	263	58.94 ± 2.42	66.81 ± 3.09	62.63 ± 1.93
Wa	0	0	0	0	0	2412	2412	2412	100.00 ± 0.00	100.00	100.00

Overall accuracy = 94.49 ± 0.26

Comparison composite length - Manica 6m rf composite

	Reference class (sample counts)						Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	Map	Ref.	Producer's	User's	F-score
Bu	153	-	-	2	2	0	157	159	96.23 ± 1.47	97.45 ± 1.26	96.84 ± 0.97
Dv	-	2300	8	2	1	0	2311	2322	99.05 ± 0.20	99.52 ± 0.14	99.29 ± 0.12
Ia	1	9	291	9	9	0	319	322	90.37 ± 1.54	91.22 ± 1.58	90.80 ± 1.11
Lv	-	13	11	290	5	0	319	310	93.55 ± 1.34	90.91 ± 1.61	92.21 ± 1.05
Ra	5	0	12	7	211	0	235	228	92.54 ± 1.66	89.79 ± 1.98	91.14 ± 1.30
Wa	0	0	0	0	0	2454	2454	2454	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00

Overall accuracy = 98.34 ± 0.16

Comparison composite length - Manica 3m rf composite

	Reference class (sample counts)						Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	Map	Ref.	Producer's	User's	F-score
Bu	155	-	-	1	5	0	161	159	97.48 ± 1.22	96.27 ± 1.49	96.87 ± 0.97
Dv	-	2307	3	6	1	0	2317	2321	99.40 ± 0.16	99.57 ± 0.14	99.48 ± 0.10
Ia	1	8	323	9	17	0	358	348	92.82 ± 1.31	90.22 ± 1.57	91.50 ± 1.03
Lv	-	5	9	266	7	0	287	297	89.56 ± 1.65	92.68 ± 1.54	91.10 ± 1.13
Ra	3	1	13	15	206	0	238	236	87.29 ± 2.01	86.55 ± 2.21	86.92 ± 1.49
Wa	0	0	0	0	0	2421	2421	2421	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00

Overall accuracy = 98.2 ± 0.17

Comparison composite length - Manica 2m rf composite

	Reference class (sample counts)						Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Wa	Map	Ref.	Producer's	User's	F-score
Bu	137	-	1	1	3	0	142	137	100.00 ± 0.00	96.48 ± 1.55	98.21 ± 0.80
Dv	-	2316	2	3	0	0	2321	2330	99.40 ± 0.16	99.78 ± 0.10	99.59 ± 0.09
Ia	-	13	312	9	6	0	340	329	94.83 ± 1.18	91.76 ± 1.49	93.27 ± 0.96
Lv	-	1	7	280	8	0	296	301	93.02 ± 1.40	94.59 ± 1.31	93.80 ± 0.96
Ra	0	0	7	8	221	0	236	238	92.86 ± 1.59	93.64 ± 1.59	93.25 ± 1.13
Wa	0	0	0	0	0	2446	2446	2446	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00

Overall accuracy = 98.81 ± 0.14

Catandica

Comparison algorithm - Catandica 6m rf composite

	Reference class (sample counts)					Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Map	Ref.	Producer's	User's	F-score
Bu	84	-	1	1	25	111	92	91.30 ± 2.81	75.68 ± 4.07	82.76 ± 2.70
Dv	-	3165	38	38	0	3241	3293	96.11 ± 0.32	97.66 ± 0.27	96.88 ± 0.21
Ia	-	55	372	38	4	469	463	80.35 ± 1.68	79.32 ± 1.87	79.83 ± 1.26
Lv	0	73	39	1970	35	2117	2081	94.67 ± 0.47	93.06 ± 0.55	93.85 ± 0.36
Ra	8	0	13	34	351	406	415	84.58 ± 1.57	86.45 ± 1.70	85.51 ± 1.15

Overall accuracy = 93.66 ± 0.3

Comparison algorithm - Catandica 6m svmRadial composite

	Reference class (sample counts)					Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Map	Ref.	Producer's	User's	F-score
Bu	109	-	-	0	2	111	109	100.00 ± 0.00	98.20 ± 1.26	99.09 ± 0.64
Dv	-	3220	2	19	0	3241	3251	99.05 ± 0.17	99.35 ± 0.14	99.20 ± 0.11
Ia	-	8	445	13	3	469	476	93.49 ± 1.08	94.88 ± 1.02	94.18 ± 0.74
Lv	0	23	13	2064	17	2117	2101	98.24 ± 0.28	97.50 ± 0.34	97.87 ± 0.22
Ra	0	0	16	5	385	406	407	94.59 ± 1.09	94.83 ± 1.10	94.71 ± 0.77

Overall accuracy = 98.09 ± 0.17

Comparison algorithm - Catandica 6m knn composite

	Reference class (sample counts)					Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Map	Ref.	Producer's	User's	F-score
Bu	5	3	88	9	6	111	23	21.74 ± 8.44	4.50 ± 1.97	7.46 ± 2.75
Dv	18	2977	209	35	2	3241	3371	88.31 ± 0.48	91.85 ± 0.48	90.05 ± 0.34
Ia	-	115	121	232	1	469	1633	7.41 ± 0.55	25.80 ± 2.02	11.51 ± 0.70
Lv	0	276	1042	773	26	2117	1279	60.44 ± 1.02	36.51 ± 1.05	45.52 ± 0.86
Ra	0	0	173	230	3	406	38	7.89 ± 4.36	0.74 ± 0.43	1.35 ± 0.71

Overall accuracy = 61.14 ± 0.45

Comparison algorithm - Catandica 6m nnet composite

	Reference class (sample counts)					Σ		Accuracy		
Map class	Bu	Dv	Ia	Lv	Ra	Map	Ref.	Producer's	User's	F-score
Bu	110	-	-	0	1	111	111	99.10 ± 0.89	99.10 ± 0.90	99.10 ± 0.63
Dv	-	3207	-	34	0	3241	3237	99.07 ± 0.17	98.95 ± 0.18	99.01 ± 0.12
Ia	1	5	458	2	3	469	460	99.57 ± 0.31	97.65 ± 0.70	98.60 ± 0.39
Lv	0	25	1	2074	17	2117	2123	97.69 ± 0.32	97.97 ± 0.31	97.83 ± 0.22
Ra	0	0	1	13	392	406	413	94.92 ± 1.05	96.55 ± 0.91	95.73 ± 0.70

Overall accuracy = 98.38 ± 0.16

Xai-Xai

Comparison algorithm - Xai-Xai 6m rf composite

	Reference class (sample counts)								Σ		Accuracy		
Map class	Bu	Dv	Gr	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	429	-	1	7	1	5	0	0	443	436	98.39 ± 0.59	96.84 ± 0.83	97.61 ± 0.51
Dv	-	945	17	25	19	1	0	31	1038	1009	93.66 ± 0.74	91.04 ± 0.89	92.33 ± 0.58
Gr	-	22	2464	48	28	-	0	14	2576	2656	92.77 ± 0.46	95.65 ± 0.40	94.19 ± 0.31
Ia	1	9	91	1039	44	8	0	13	1205	1209	85.94 ± 0.91	86.22 ± 0.99	86.08 ± 0.67
Lv	2	6	76	65	448	31	0	7	635	568	78.87 ± 1.54	70.55 ± 1.81	74.48 ± 1.22
Ra	4	-	3	12	25	93	0	0	137	138	67.39 ± 3.47	67.88 ± 3.99	67.64 ± 2.64
Wa	0	0	0	0	0	0	990	0	990	990	100.00 ± 0.00	100.00	100.00
We	0	27	4	13	3	0	0	1012	1059	1077	93.96 ± 0.70	95.56 ± 0.63	94.76 ± 0.47

Overall accuracy = 91.8 ± 0.29

Comparison algorithm - Xai-Xai 6m svmRadial composite

	Reference class (sample counts)								Σ		Accuracy		
Map class	Bu	Dv	Gr	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	410	-	2	10	9	12	0	0	443	432	94.91 ± 1.01	92.55 ± 1.25	93.71 ± 0.81
Dv	-	873	34	44	-	-	0	87	1038	957	91.22 ± 0.87	84.10 ± 1.13	87.52 ± 0.73
Gr	-	28	2361	150	14	-	0	23	2576	2835	83.28 ± 0.55	91.65 ± 0.54	87.27 ± 0.39
Ia	-	23	141	1022	5	3	0	11	1205	1376	74.27 ± 1.00	84.81 ± 1.03	79.19 ± 0.73
Lv	13	1	266	97	202	44	0	12	635	257	78.60 ± 2.36	31.81 ± 1.85	45.29 ± 1.91
Ra	9	-	24	9	27	68	0	0	137	127	53.54 ± 3.80	49.64 ± 4.27	51.52 ± 2.90
Wa	0	0	0	0	0	0	990	0	990	1001	98.90 ± 0.33	100.00	99.45
We	0	32	7	44	0	0	11	965	1059	1098	87.89 ± 0.90	91.12 ± 0.87	89.48 ± 0.63

Overall accuracy = 85.25 ± 0.35

Comparison algorithm - Xai-Xai 6m knn composite

	Reference class (sample counts)								Σ		Accuracy		
Map class	Bu	Dv	Gr	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	440	-	-	3	-	-	0	0	443	1477	29.79 ± 0.51	99.32 ± 0.39	45.83 ± 0.60
Dv	-	827	18	58	-	-	0	135	1038	1138	72.67 ± 1.10	79.67 ± 1.25	76.01 ± 0.83
Gr	158	70	1290	1025	2	-	0	31	2576	2143	60.20 ± 0.80	50.08 ± 0.99	54.67 ± 0.67
Ia	48	34	516	223	-	-	0	384	1205	2109	10.57 ± 0.60	18.51 ± 1.12	13.46 ± 0.57
Lv	200	10	250	168	-	-	0	7	635	9	0.00 ± 0.00	0.00 ± 0.00	-
Ra	101	-	17	19	-	-	0	0	137	41	0.00 ± 0.00	0.00 ± 0.00	-
Wa	524	0	0	419	6	41	0	0	990	0	-	0.00 ± 0.00	-
We	6	197	52	194	1	0	0	609	1059	1166	52.23 ± 1.13	57.51 ± 1.52	54.74 ± 0.93

Overall accuracy = 41.93 ± 0.44

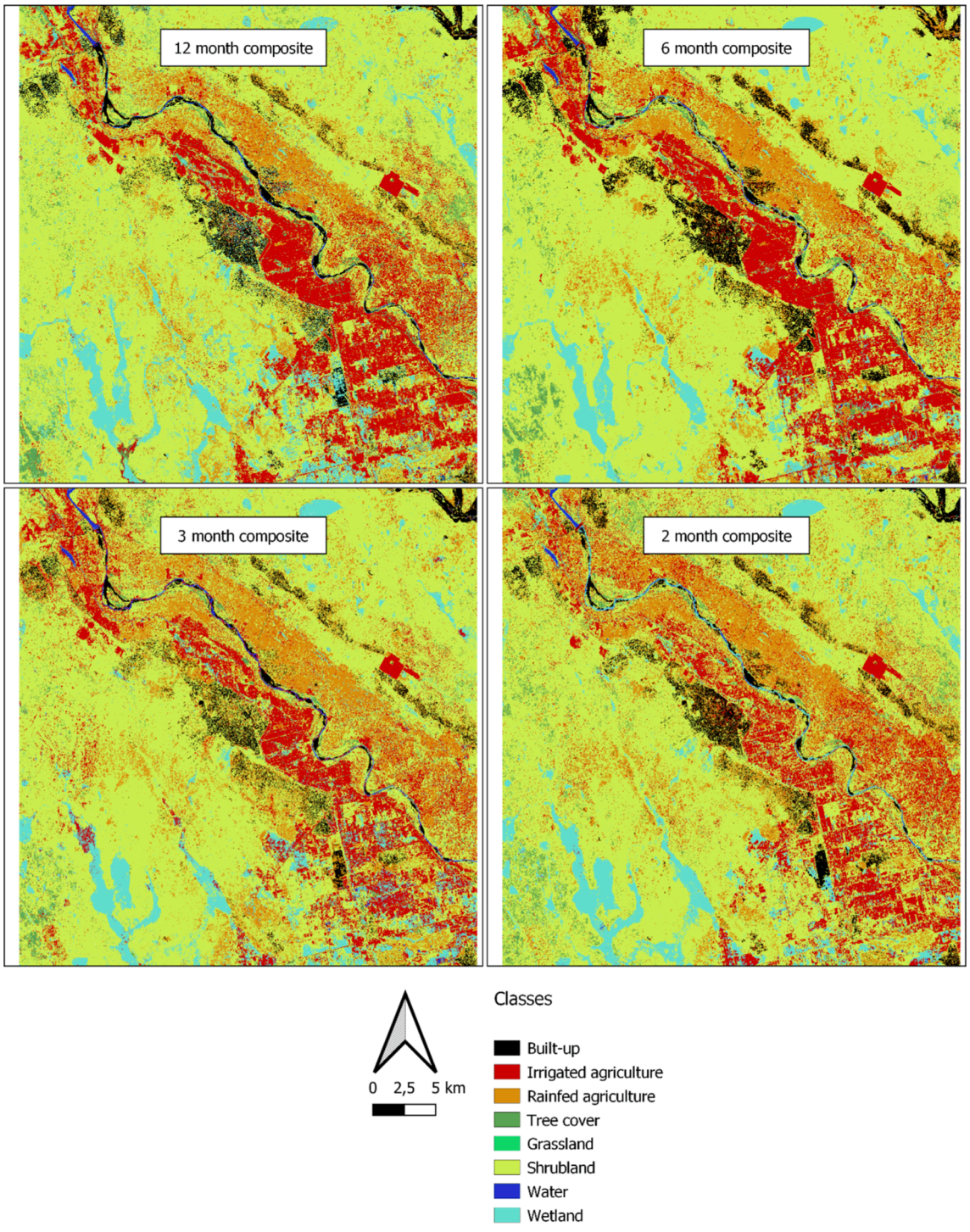
Comparison algorithm - Xai-Xai 6m nnet composite

	Reference class (sample counts)								Σ		Accuracy		
Map class	Bu	Dv	Gr	Ia	Lv	Ra	Wa	We	Map	Ref.	Producer's	User's	F-score
Bu	437	-	-	3	1	2	0	0	443	439	99.54 ± 0.32	98.65 ± 0.55	99.09 ± 0.32
Dv	-	963	22	29	19	-	1	4	1038	1010	95.35 ± 0.64	92.77 ± 0.80	94.04 ± 0.52
Gr	-	6	2439	77	32	-	0	22	2576	2662	91.62 ± 0.49	94.68 ± 0.44	93.13 ± 0.33
Ia	1	8	59	1097	26	4	0	10	1205	1278	85.84 ± 0.89	91.04 ± 0.82	88.36 ± 0.61
Lv	1	28	115	46	397	47	0	1	635	529	75.05 ± 1.60	62.52 ± 1.92	68.21 ± 1.32
Ra	-	-	6	18	53	60	0	0	137	113	53.10 ± 4.09	43.80 ± 4.24	48.00 ± 3.05
Wa	0	0	0	0	0	0	990	0	990	991	99.90 ± 0.10	100.00	99.95
We	0	5	21	8	1	0	0	1024	1059	1061	96.51 ± 0.55	96.69 ± 0.55	96.60 ± 0.39

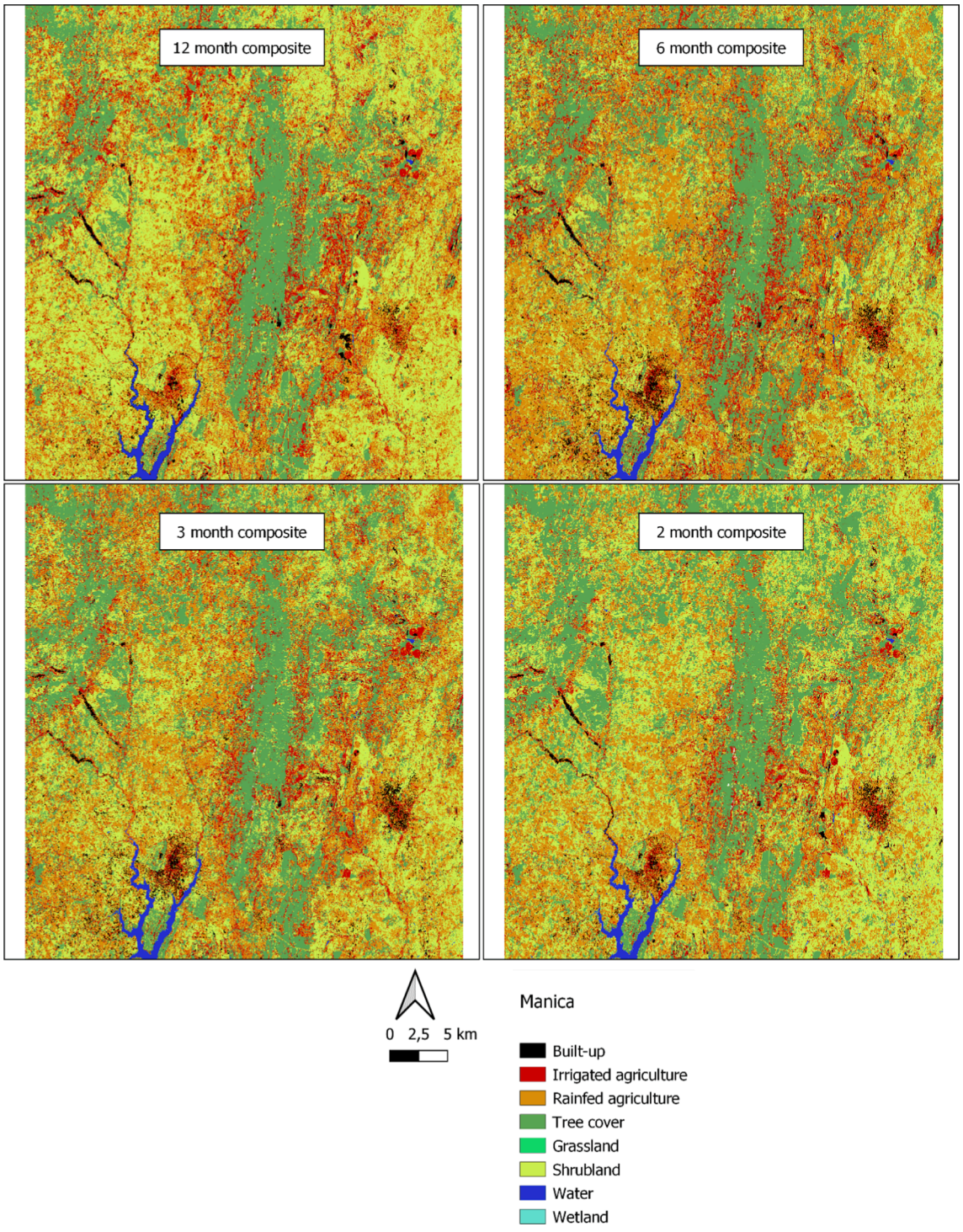
Overall accuracy = 91.64 ± 0.28

Annex 2: Classification maps

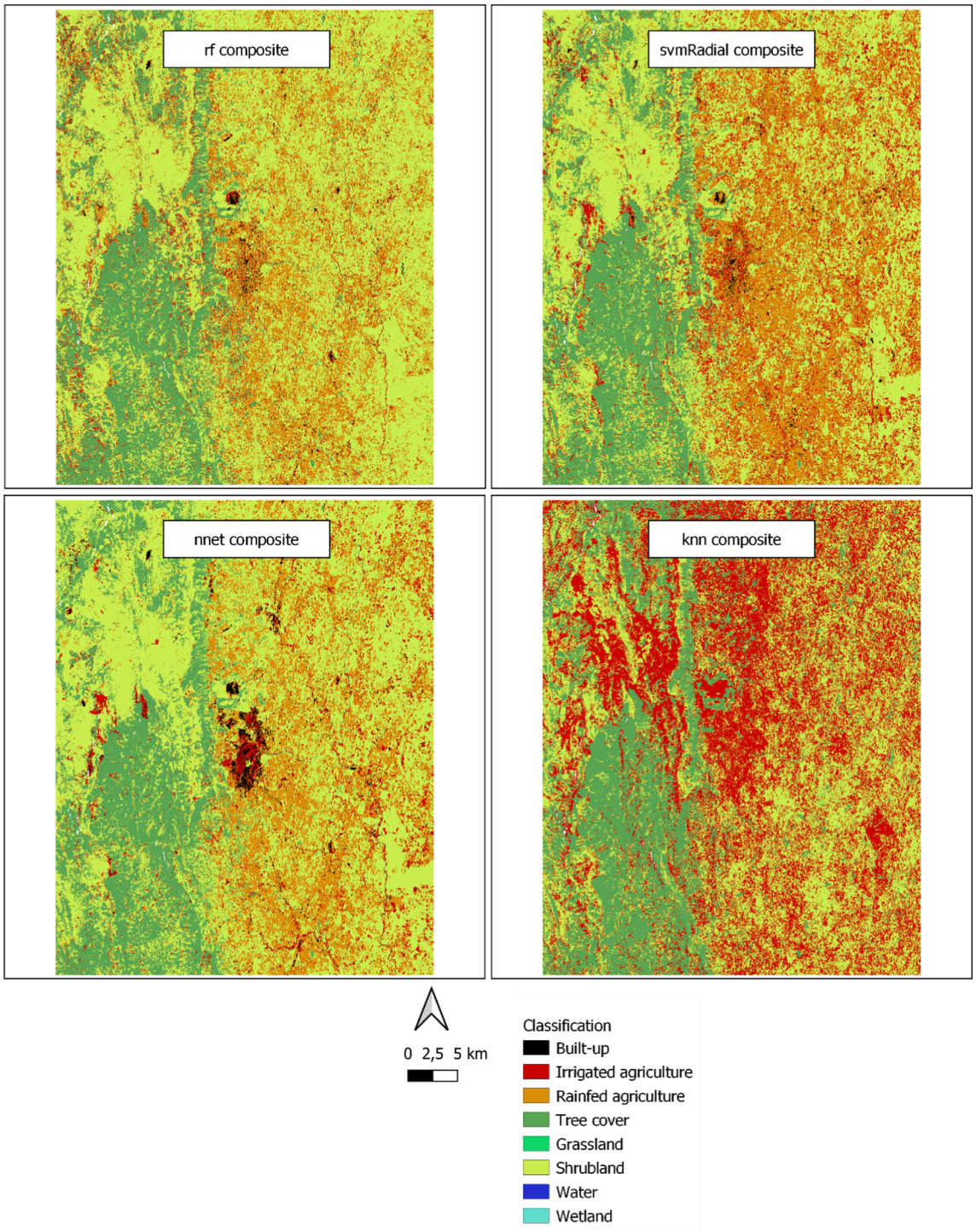
Chokwe


12 month composite
6 month composite
3 month composite
2 month composite
0 2,5 5 km
Classes
Built-up
Irrigated agriculture
Rainfed agriculture
Tree cover
Grassland
Shrubland
Water
Wetland

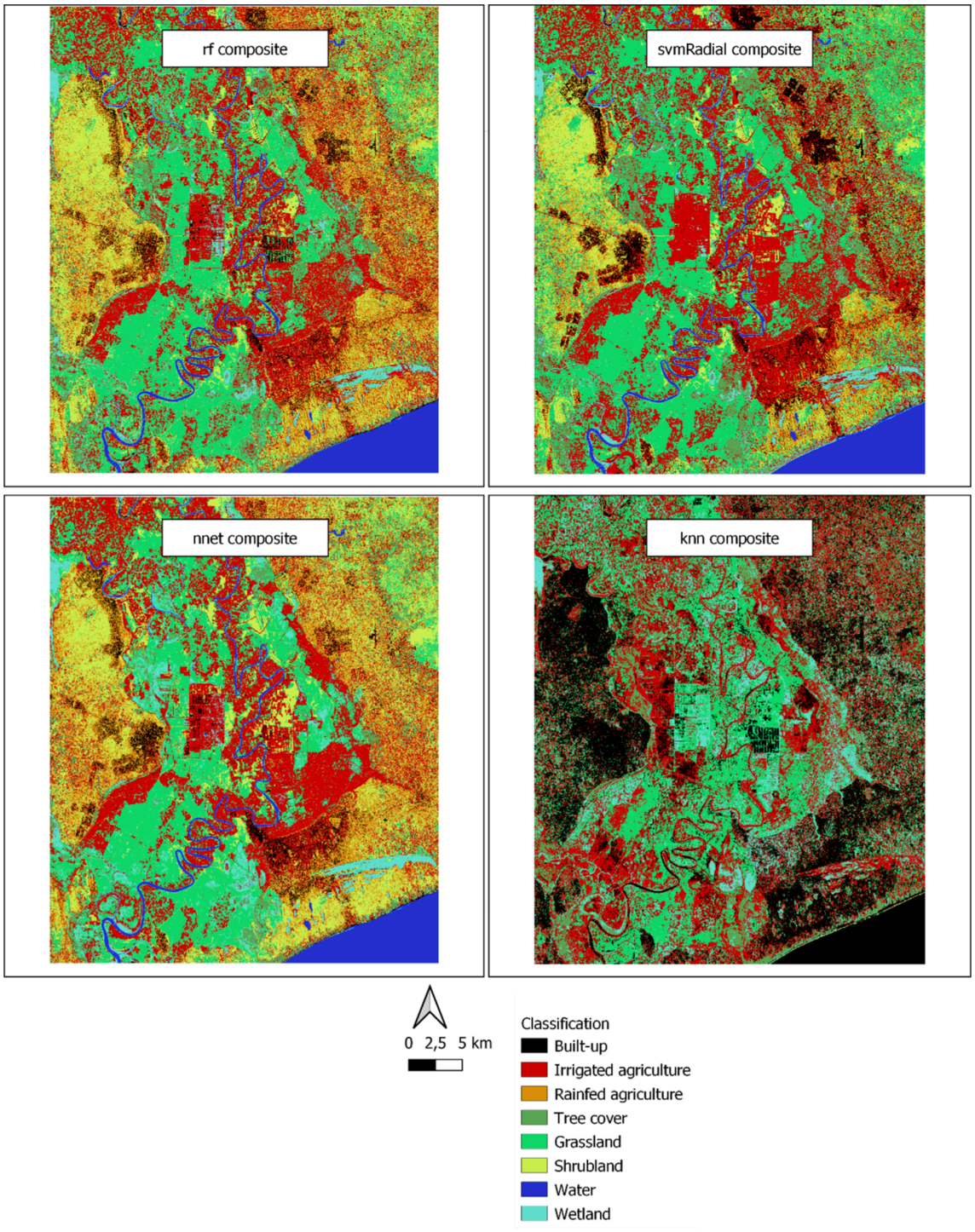
Manica


12 month composite
6 month composite
3 month composite
2 month composite
0 2,5 5 km
Manica
Built-up
Irrigated agriculture
Rainfed agriculture
Tree cover
Grassland
Shrubland
Water
Wetland

Catandica


rf composite
svmRadial composite
nnet composite
knn composite
0 2,5 5 km
Classification
Built-up
Irrigated agriculture
Rainfed agriculture
Tree cover
Grassland
Shrubland
Water
Wetland

Xai-Xai



References

Abdolrasol, M.G.M., Hussain, S.M.S., Ustun, T.S., Sarker, M.R., Hannan, M.A., Mohamed, R., Ali, J.A., Mekhilef, S., Milad, A., 2021. Artificial Neural Networks Based Optimization Techniques: A Review. Electronics 10, 2689. https://doi.org/10.3390/electronics10212689.

Abu Alfeilat, H.A., Hassanat, A.B.A., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B., Eyal Salman, H.S., Prasath, V.B.S., 2019. Effects of Distance Measure Choice on K-Nearest Neighbor Classifier Performance: A Review. Big Data 7, 221–248. https://doi.org/10.1089/big.2018.0175.

Abubakar, G.A., Wang, K., Shahtahamssebi, A., Xue, X., Belete, M., Gudo, A.J.A., Mohamed Shuka, K.A., Gan, M., 2020. Mapping Maize Fields by Using Multi-Temporal Sentinel-1A and Sentinel-2A Images in Makarfi, Northern Nigeria, Africa. Sustainability 12, 2539 https://doi.org/10/gkzh5s.

Ajaz, A., Karimi, P., Cai, X., De Fraiture, C., Akhter, M.S., 2019. Statistical Data Collection Methodologies of Irrigated Areas and Their Limitations: A Review. Irrig. Drain. 68, 702–713 https://doi.org/10/gk464q.

Beekman, W., Veldwisch, G.J., Bolding, A., 2014. Identifying the potential for irrigation development in Mozambique: Capitalizing on the drivers behind farmer-led irrigation expansion. Phys. Chem. Earth Parts ABC 76–78, 54–63 https://doi.org/10/gk8kfc.

Bégué, A., Arvor, D., Bellon, B., Betbeder, J., de Abelleyra, D., P.D. Ferraz, R., Lebourgeois, V., Lelong, C., Simões, M., Verón, R., S.,, 2018. Remote Sensing and Cropping Practices: A Review. Remote Sens. 10, 99 https://doi.org/10/gczn8x.
Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31 https://doi.org/10/f8ndk8.
Bey, A., Jetimane, J., Lisboa, S.N., Ribeiro, N., Sitoe, A., Meyfroidt, P., 2020. Mapping smallholder and large-scale cropland dynamics with a flexible classification system and pixel-based composites in an emerging frontier of Mozambique. Remote Sens. Environ. 239, 111611 https://doi.org/10/gk8kff.
Bofana, J., Zhang, M., Nabil, M., Wu, B., Tian, F., Liu, W., Zeng, H., Zhang, N., Nangombe, S.S., Cipriano, S.A., Phiri, E., Mushore, T.D., Kaluba, P., Mashonjowa, E., Moyo, C., 2020. Comparison of Different Cropland Classification Methods under Diversified Agroecological Conditions in the Zambezi River Basin. Remote Sens. 12, 2096 https://doi.org/10/ghfn2v.
Brandt, M., Bäumler, R., Samimi, C., 2009. Agricultural suitability of dune system and Limpopo Basin soils near Xai-Xai. Mozambique. South Afr. J. Plant Soil 26, 206–212. https://doi.org/10.1080/02571862.2009.10639956.
de Bont, C., 2019. Modernisation and African Farmer-Led Irrigation Development: Ideology. Policies and Practices. Water Altern. 12, 23.
Dea, . DEA GeoMAD [WWW Document]. accessed 9.6.22. https://docs.digitalearthafrica.org/en/latest/data_specs/GeoMAD_specs.html#Triple-Median-Absolute-Deviations-(MADs. .
Elwan, E., Le Page, M., Jarlan, L., Baghdadi, N., Brocca, L., Modanesi, S., Dari, J., Quintana Seguí, P., Zribi, M., 2022. Irrigation Mapping on Two Contrasted Climatic Contexts Using Sentinel-1 and Sentinel-2 Data. Water 14, 804. https://doi.org/10.3390/w14050804.
Espey, J., 2019. Sustainable development will falter without data. Nature 571, 299. https://doi.org/10.1038/d41586-019-02139-w.
Gao, Q., Zribi, M., Escorihuela, M., Baghdadi, N., Segui, P., 2018. Irrigation Mapping Using Sentinel-1 Time Series at Field Scale. Remote Sens. 10, 1495 https://doi.org/10/gfnvvx.
Gella, G.W., Bijker, W., Belgiu, M., 2021. Mapping crop types in complex farming areas using SAR imagery with dynamic time warping. ISPRS J. Photogramm. Remote Sens. 175, 171–183 https://doi.org/10/gjh2bs.
Gitelson, A.A., Viña, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll content in crops. Geophys. Res. Lett. 32 https://doi.org/10.1029/2005GL022688.
Gumbo, A.D., Kapangaziwiri, E., Chikoore, H., Pienaar, H., Mathivha, F., 2021. Assessing water resources availability in headwater sub-catchments of Pungwe River Basin in a changing climate. J. Hydrol. Reg. Stud. 35, 100827 https://doi.org/10.1016/j.ejrh.2021.100827.
Hasenbein, K., Abdel-Rahman, E.M., Adan, M., Gachoki, S.M., King'ori, E., Dubois, T., Landmann, T., 2022. Availability of Sentinel-2-based time-series observations: which vegetation phenology-based metrics perform best for mapping farming systems in complex landscapes? Remote Sens. Lett. 13, 695–707. https://doi.org/10.1080/2150704X.2022.2068985.
Izzi, G., Denison, J., Veldwisch, G.J., 2021. The farmer-led irrigation development guide: A what, why and how-to for intervention design. World Bank, Washington, DC.
Jennewein, J.S., Lamb, B.T., Hively, W.D., Thieme, A., Thapa, R., Goldsmith, A., Mirsky, S.B., 2022. Integration of Satellite-Based Optical and Synthetic Aperture Radar Imagery to Estimate Winter Cover Crop Performance in Cereal Grasses. Remote Sens. 14, 2077. https://doi.org/10.3390/rs14092077.
Kajisa, K., Payongayong, E., 2011. Potential of and constraints to the rice Green Revolution in Mozambique: A case study of the Chokwe irrigation scheme. Food Policy 36, 615–626. https://doi.org/10.1016/j.foodpol.2011.07.002.
Khatami, R., Southworth, J., Muir, C., Caughlin, T., Ayana, A.N., Brown, D.G., Liao, C., Agrawal, A., 2020. Operational Large-Area Land-Cover Mapping: An Ethiopia Case Study. Remote Sens. 12, 954 https://doi.org/10/gpbxqv.
Kuhn, M., 2008. Building Predictive Models in R Using the caret Package. J. Stat. Softw. 28, 1–26 https://doi.org/10.18637/jss.v028.i05.
Kuhn, M., 2019. The caret Package.
Kumar, M., Phukon, S.N., Paygude, A.C., Tyagi, K., Singh, H., 2022. Mapping Phenological Functional Types (PhFT) in the Indian Eastern Himalayas using machine learning algorithm in Google Earth Engine. Comput. Geosci. 158, 104982 https://doi.org/10.1016/j.cageo.2021.104982.
Lebourgeois, V., Dupuy, S., Vintrou, É., Ameline, M., Butler, S., Bégué, A., 2017. A Combined Random Forest and OBIA Classification Scheme for Mapping Smallholder Agriculture at Different Nomenclature Levels Using Multisource Data (Simulated Sentinel-2 Time Series, VHRS and DEM). Remote Sens. 9, 259 https://doi.org/10/gk8kgr.
Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J.M., McNairn, H., Rao, Y.S., 2020. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. Remote Sens. Environ. 247, 111954 https://doi.org/10.1016/j.rse.2020.111954.
Marín Del Valle, T., Jiang, P., 2022. Comparison of common classification strategies for large-scale vegetation mapping over the Google Earth Engine platform. Int. J. Appl. Earth Obs. Geoinformation 115, 103092. https://doi.org/10.1016/j.jag.2022.103092.
Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: an applied review. Int. J. Remote Sens. 39, 2784–2817 https://doi.org/10/gfghjc.
Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Nauss, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environ. Model. Softw. 101, 1–9 https://doi.org/10/gc2tsg.
Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS J. Photogramm. Remote Sens. 66, 247–259 https://doi.org/10/frc3mg.
ODK collect, 2022. ODK Collect [WWW Document]. URL https://docs.getodk.org/ (accessed 9.6.22).
Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57 https://doi.org/10/f55gjk.
Pflugmacher, D., 2022. mapac: Map accuracy and area estimation.
Ramezan, C., Warner, T., Maxwell, A., 2019. Evaluation of Sampling and Cross-Validation Tuning Strategies for Regional-Scale Machine Learning Classification. Remote Sens. 11, 185 https://doi.org/10/gk8kd5.
Roberts, D., Mueller, N., Mcintyre, A., 2017. High-Dimensional Pixel Composites From Earth Observation Time Series. IEEE Trans. Geosci. Remote Sens. 55, 6254–6264 https://doi.org/10/gcgqn3.
Roberts, D., Dunn, B., Mueller, N., 2018. Open Data Cube Products Using High-Dimensional Statistics of Time Series. In: IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, pp. 8647–8650.
Segarra, J., Buchaillot, M.L., Araus, J.L., Kefauver, S.C., 2020. Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications. Agronomy 10, 641. https://doi.org/10.3390/agronomy10050641.
Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., Homayouni, S., 2020. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 6308–6325 https://doi.org/10/gkfsb5.
Thanh Noi, P., Kappas, M., 2017. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors 18, 18 https://doi.org/10/gf55h8.
Veldwisch, G.J., Venot, J.-P., Woodhouse, P., Komakech, H.C., Brockington, D., 2019. Re-introducing Politics in African Farmer-Led Irrigation Development: Introduction to a Special Issue 12, 12.
Venot, J.-P., Bowers, S., Brockington, D., Komakech, H., Ryan, C., Veldwisch, G.J., Woodhouse, P., 2021. Below the Radar: Data, Narratives and the Politics of Irrigation in Sub-Saharan. Africa 14, 27.
Weemstra, H., Oord, A.L., de Boer, F.S., Beekman, P.W., 2014. Baseflow prediction in a data-scarce catchment with Inselberg topography, Central Mozambique. Phys. Chem. Earth Parts ABC 76–78, 16–27. https://doi.org/10.1016/j.pce.2014.09.005.
Wellington, M.J., Renzullo, L.J., 2021. High-Dimensional Satellite Image Compositing and Statistics for Enhanced Irrigated Crop Mapping. Remote Sens. 13, 1300 https://doi.org/10/gmfbrm.
Xie, Y., Lark, T.J., Brown, J.F., Gibbs, H.K., 2019. Mapping irrigated cropland extent across the conterminous United States at 30 m resolution using a semi-automatic training approach on Google Earth Engine. ISPRS J. Photogramm. Remote Sens. 155, 136–149 https://doi.org/10/ghgc6p.
Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.-E., Xu, P., Ramoino, F., Arino, O., 2022. ESA WorldCover 10 m 2021 v200. https://doi.org/10.5281/zenodo.7254221.