Remote Sensing Applications: Society and Environment 36 (2024) 101282

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

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

Pixels to pasture: Using machine learning and multispectral remote sensing to predict biomass and nutrient quality in tropical grasslands

Mike Zwick [a], Juan Andres Cardoso [b], Diana María Gutiérrez-Zapata [b,c], Mario Cerón-Muñoz [c], Jhon Freddy Gutiérrez [d], Christoph Raab [e], Nicholas Jonsson [f], Miller Escobar [b], Kenny Roberts [a], Brian Barrett [a,*]

[a] School of Geographical & Earth Sciences, University of Glasgow, G12 8QQ, Scotland, UK
[b] International Center for Tropical Agriculture (CIAT), Cali, Colombia
[c] GAMMA Research Group, Faculty of Agrarian Science, Universidad de Antioquia (UdeA), Medellín, Colombia
[d] Sustainable Livestock Technical Service Corporation (GANSO), Villavicencio, Colombia
[e] Department of Geography, University of Hildesheim, Germany
[f] School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, G12 8QQ, Scotland, UK

ARTICLE INFO

Keywords:
Grassland monitoring
Biomass
Multispectral remote sensing
Machine learning
Colombia

ABSTRACT

The livestock sector in rural Colombia is critical for employment and food security but is heavily affected by climate and its change. There is a need for solutions to address key challenges arising from vulnerabilities that impact the productivity and sustainability of forages and the livestock sector. Increasing the yields of forage crops can improve the availability and affordability of livestock products while also easing the pressure on land resources. This study aims to develop remote sensing-based approaches for forage monitoring and biomass prediction in Colombia to support decision-making towards increased productivity, competitiveness and reduction of environmental impacts. Ten locations were sampled between 2018 and 2021 across climatically distinct areas in Colombia, comprising five farms in Patía in Cauca department, four farms in Antioquia department, and one research farm at Palmira in Valle de Cauca department. Ash content (Ash), crude protein (CP %), dry matter content (DM g/m^2) and in-vitro digestibility (IVD %) were measured from Kikuyu and *Brachiaria* grasses during the field sampling campaigns. Multispectral bands from coincident Planetscope acquisitions along with various derived vegetation indices (VIs) were used as predictors in the model development. For each site and forage parameter, the importance of specific predictors varied, with the NIR band and Red-Green ratio generally performing best. To determine the optimum models, the effects of using a 1) averaging kernel, 2) feature selection approaches, 3) various regression algorithms and 4) meta learners (simple ensembling and stacks) were explored. Algorithms belonging to classes of commonly used models; Decision Trees, Support Vector Machines, Neural Networks, distance-based methods, and linear approaches were tested. The performance evaluation based on unseen test data revealed that CP and DM prediction performed moderately well for all three sites (R^2 0.52–0.75, RMSE 1.7–2 % and R^2 0.47–0.65, RMSE 182–112 g/m^2 respectively). The best performing models varied by site and response variable, with Regularized Random Forest, Partial Least Squares, Random Forests, Bagged Multivariate Adaptive Regression and Bayesian Regularized Neural Networks being the top performing algorithms and Random Forest Stack being the best performing meta learner. The

* Corresponding author.
E-mail address: brian.barrett@glasgow.ac.uk (B. Barrett).

https://doi.org/10.1016/j.rsase.2024.101282
Received 29 January 2024; Received in revised form 31 May 2024; Accepted 21 June 2024
Available online 22 June 2024

workflow and thorough analysis of performance affecting factors presented in this study can benefit timely grassland monitoring and biomass prediction at the local level and help contribute to the sustainable management of tropical grasslands in Colombia.

1. Introduction

Grasslands cover approximately 40% of the Earth's land surface and almost 70% of global agricultural land area (Gibson, 2009; Reinermann et al., 2020). They provide a range of benefits, supporting biodiversity, livestock grazing and food production. Moreover, grasslands are critical for maintaining a variety of ecosystem services (Masenyama et al., 2022) such as carbon storage and mitigation, pollination, and cultural services (O'Mara, 2012; Ali et al., 2016a; Bengtsson et al., 2019; Morais et al., 2021; Bardgett et al., 2021). In Colombia, approximately 34–38% of the land area is grassland and the livestock sector accounts for 19% of agricultural employment (Tapasco et al., 2019; Rodríguez-Hernández et al., 2023). Forage grass forms the largest component of cattle diets; hence its quantity and quality are key factors influencing the productivity and stocking rates in dairy and beef production systems. Food security and the Sustainable Development Goals (e.g. SDG 2 – Zero hunger, SDG 13 – Climate action, SDG 15 – Life on Land) are directly related to these production systems (Viana et al., 2022; Bonilla-Cedrez et al., 2023), and for efficient and sustainable management of both cattle and pasture alongside expected changes in climate (Ray et al., 2015; Tapasco et al., 2015), and land use and condition (Zuluaga et al., 2021), reliable information on forage quantity and quality is needed (Ferner et al., 2015). Forage quantity (biomass) is most often determined by visual observations in the field, using rising plate meters, or by harvesting plants at different growth stages for posterior analyses (Tucker, 1980). Forage quality is an equally important attribute of grasslands concerning their productivity and management (Ali et al., 2016a). In the context of animal feed, forage quality refers to nutritional parameters such as crude protein, digestibility, and Ash content (Pullanagari et al., 2013). Ash is the total mineral (non-organic) content of a forage and is important for assessing animal health and performance. When a forage contains an excess of Ash or crude protein is low, forage quality and animal productivity decreases (lower forage intake and reduced digestibility).

Laboratory-based methods to estimate forage quality include proximate biochemical analysis and near infrared spectroscopy (NIRS) (Pullanagari et al., 2013). However, laboratory analyses can be costly, time consuming and are mostly confined to a point sampling strategy which might not be representative when performed over large areas (Starks et al., 2006; Knox et al., 2012). Similar issues exist for destructive biomass sampling (Kumar and Mutanga, 2017). Compared to traditional measurements, remote sensing can extend estimation of forage quantity and quality over larger areas and cover more of the spatio-temporal variability of forage characteristics (Ali et al., 2016a). As grasslands often exhibit small scale heterogeneities, the increasing spatial (3 m) and temporal (daily) resolution provided by nanosatellites such as PlanetScope can help overcome some of the limitations of using coarse (e.g. MODIS) or medium (e.g. Landsat, Sentinel-2) resolution data (Wachendorf et al., 2018).

Machine learning algorithms have been increasingly used for biomass estimation (e.g. Shoko et al., 2016; Zhou et al., 2021). The performance of these predictive algorithms is highly variable and identifying performance controlling factors is far from trivial, leading to a lack of consensus on the optimum approach (e.g. Morais et al., 2021). Generally, ensemble learners refer to algorithms that use an ensemble of identical base learners (e.g. Random Forest (RF) uses decision trees) whereas what we refer to as simple ensembling in the realm of meta learning means a blending of multiple/different base learners. The RF algorithm is based on an ensemble of decision trees that each votes to obtain the regression/classification result (Breiman, 2001). The individual trees are trained on samples using bootstrap aggregation. This algorithm is robust to noise in the data (Kontschieder et al., 2011), takes into account correlation between predictions of base learners (Chan and Paelinckx, 2008), has relatively few hyperparameters and is much less susceptible to overfitting (Breiman, 2001). RF has been applied in numerous studies including its application for remote sensing (e.g. Gislason et al., 2006; Barrett et al., 2016; Raab et al., 2020; Schucknecht et al., 2022). The Regularized Random Forest (RRF) algorithm in essence works identical to RF but is more stringent, especially in context of used features. It contains a penalization coefficient (λ) causing a feature only to be selected if the gain is substantial (Deng and Runger, 2012; Sylvester et al., 2018). The algorithm xgbTree (extreme gradient boosting) employs Randoms Forests together with a boosting algorithm using fewer splits and shallower trees. It has similarity to gradient boosting machines (GBM) but overfits less due to regularized boosting (Chen et al., 2018; Chen and Guestrin, 2016; Just et al., 2018; Levick et al., 2021). Substantially different to these ensemble learners is Partial Least Squares (PLS) regression. It combines elements from principal component analysis and multiple linear regression (MLR) via extraction of those orthogonal factors (latent variables) from predictors that have the highest predictive power in the subsequent MLR (Höskuldsson, 1988; Abdi, 2003, 2010). This approach makes PLS particularly powerful in cases of low number of samples and high number of features and in cases where in the latter multicollinearity occurs (Abdi, 2010).

Applying a meta learner ('level-1') in form of hierarchical blending (Simple Ens) or stacking of several different base learners ('level-0') combines advantages of respective base learners often yielding superior generalizability and hence modelling accuracy (Wang, 2018; Wu et al., 2021). This is not yet commonly applied despite its high performance (Wu et al., 2021) likely due to its novelty, its increased complexity or substantial processing requirements. The difference in blending and stacking is marginal. In blending, no meta-feature dataset is created via k-fold cross-validation but instead the leave-one-out method is used. This approach makes processing faster but also more susceptible to overfitting compared to stacking (Wu et al., 2021).

Most studies implementing machine learning approaches focus on measuring grassland quantity, and grassland quality has typically received less attention (e.g. Ramoelo et al., 2015; Wachendorf et al., 2018; Irisarri et al., 2022), due largely to additional expense of field and laboratory sampling or poor performance of approaches. Consequently, the value of grasslands, determined by both quantity and quality, and their potential through appropriate management for meeting farm sustainability and profitability is not realised. The aim of this research was to assess different feature and classifier selection approaches to estimate tropical forage quantity and quality using high resolution PlanetScope multispectral satellite data.

2. Materials and methods

2.1. Study areas

The study was performed in three regions across the Colombian Andes Mountain range. The Colombian Andes are characterised by high elevation variability causing steep temperature and precipitation gradients (Echeverría-Londoño et al., 2016) (see Fig. 1). Its connectivity or lack thereof with major moisture supply trajectories, water balance, elevation, substrate, and land management determine the types of vegetation cover and land use. This interplay of factors leads to climatically distinct regions from which we selected study areas. These vary in forage type and management practices and include a) four farms in municipalities of Antioquia department, b) one research farm in the municipality of Palmira in Valle de Cauca department and c) five farms in the municipality of Patía in Cauca department (see Fig. 1a and section 2.2.1 for information on forage species). The Antioquia sites (municipalities Belmira, Donmatías and Bello) are 17–59 km NW to NE from Medellín at elevations of 2320–2620 masl. It has a temperate oceanic climate (Köppen: Cfb) with an average temperature of 15 °C and the driest season occurs from December to February. The wet season lasts ~9 months from March to November with May being the wettest month (Fig. 1c).

The Palmira site is a forage test plot maintained by the Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT) and is 6 km SW of the city Palmira at an elevation of 970 masl. This site experiences tropical savanna climate with dry-winter characteristics (Köppen: Aw) and two dry seasons which are most pronounced from June–August and December–February. Wet seasons are March–May and September to November, with September being the wettest month. The Patía sites are 3–6 km south of the city Patía at elevations of 600–750 masl. This region has a tropical monsoon climate (Köppen: Am) and a dry season from June to September. The wet season runs from October to May with November being the wettest month.

2.2. Ground truth data collection

2.2.1. Forage quantity and quality measurements

During the sampling campaign, we sampled ten locations to determine Ash content (Ash), crude protein (CP %), dry matter content (DM g/m^2) and in-vitro digestibility (IVD %) for each sample in laboratory duplicates with the mean used in further analyses (see Table 1). Analyses were conducted at the CIAT laboratory in Palmira (for CIAT and Patía sites) and Agrosavia laboratory in Medellín

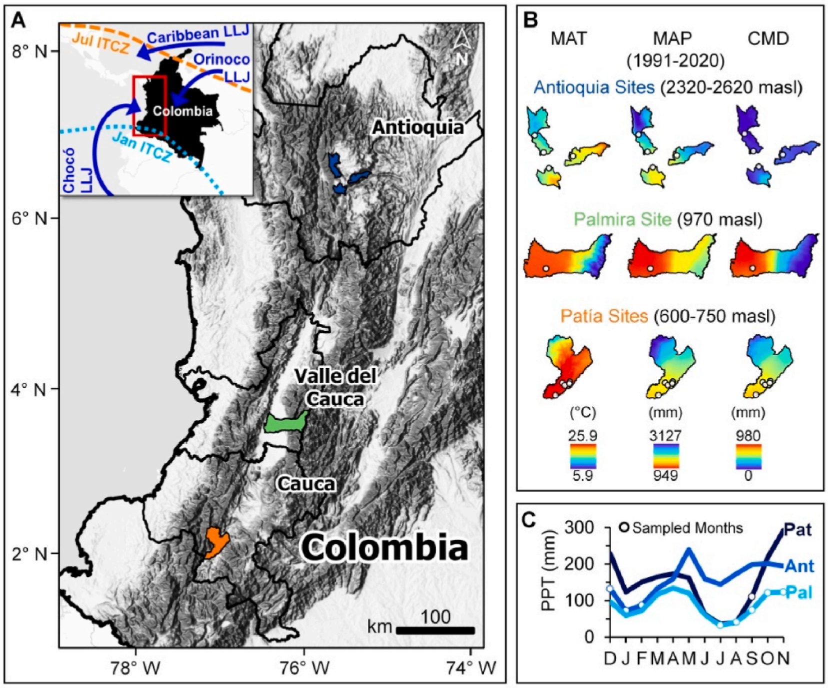

Fig. 1. A) Study areas shown on the left with air moisture sources (Poveda and Mesa, 1997) and seasonal (January and July) migration of the Intertropical Convergence Zone (ITCZ) [redrawn after (Jiménez-Sánchez et al., 2019; Asmerom et al., 2020)] illustrated in the inlet. The meridional oscillation of the ITCZ modulates the supply of precipitation via three major sources: i) easterly winds as Orinoco Low-level Jet (-LLJ) with moisture from the tropical Atlantic Ocean, ii) Caribbean LLJ with moisture from the Caribbean Sea, and iii) the low- Chocó LLJ providing moisture from the South Pacific Ocean. B) Spatial distribution of climate variables and site locations. C) Seasonal changes in average monthly precipitation (1991–2020) shown for the three study sites and sampled months. MAT = Mean Air Temperature, MAP = Mean Annual Precipitation, CMD = Climate Moisture Deficit, PPT = Precipitation (monthly), LLJ = Low-level Jet. DEM data from Farr et al. (2007), climate variables including seasonality from ClimateSA (Hamann et al., 2013).

Table 1
Summary of field sample sites. CP = Crude Protein, DM = Dry Matter, IVD = In-vitro Digestibility.

Site (variables)	No. Of Samples	Farm	Lat	Lon	Grass species/cultivars	Sampling
Patía (Ash, CP, DM, IVD)	65	El Eden	2.037	−77.076	*Megathyrsus maximus* cv. Mombaza	Sep 2018
		California	2.036	−77.111	*M. maximus* cv. Mombaza; *Urochloa brizantha* cv. Toledo; *U.* híbrido cv. Cayman	
		Villa Paola	2.070	−77.057	*M. maximus* cv. Mombaza; *U.* híbrido cv. Cayman	
		Atoca	2.052	−77.065	*M. maximus* cv. Mombaza; *U. brizantha* cv. *Toledo;* *U.* híbrido cv. Cayman; *Dichanthium aristatum*	
		Pinar del Rio	1.958	−77.124	*D. annulatum; U. decumbes; U. humidicola;*	
Antioquia (Ash, CP, DM, IVD)	120 (40 per month)	Don Bosco	6.620	−75.681	*C. clandestinus*	Dec 2020–Feb 2021
		La Tablaza	6.514	−75.624	*C. clandestinus*	
		La Fe	6.490	−75.440	*C. clandestinus*	
		Los Recuerdos	6.415	−75.591	*C. clandestinus*	
Palmira (CP, DM, IVD)	665 (Jul: 36, Aug: 178, Sep: 142, Oct: 137, Nov: 172)	CIAT	3.506	−76.353	*Mulato hybrid: U. ruziziensis x U. brizantha); U. brizantha cv. Marandu; U. humidicola*	Jul–Nov 2018

(for Antioquia sites) following forage analysis protocols described in Mazabel et al. (2020). Ash content values were not available for the Palmira site.

Before sampling, we recorded the locations of each sample spot with a South G1 differential GPS receiver (*South* Surveying & Mapping Instrument *Co.*, Ltd., China). Subsequently, DM was determined by a) cutting the forage to a height of 10 cm within an area of 0.5 × 0.5 m followed by b) drying at 60 °C for 72 h and c) weighing and normalisation of its dry mass to 1 m^2 (see Figs. S1–3). We obtained the Ash content by burning DM in a muffle furnace at 550 °C for 8 h and subsequent weighing of residue. That weight was then converted to a percentage relative to DM. We analysed CP with a FOSS Kjeltec 8100 (Foss Company) using the protocol provided by the Association of Official Analytical Chemists AOAC (method 2001.11). The IVD was estimated following the method published in Tilley and Terry (1963).

2.3. PlanetScope acquisitions

The image acquisitions for this study were obtained from PlanetScope (PS) satellites. PlanetScope is a constellation of over 200 nanosatellites developed by PlanetLabs Inc. that can provide an unprecedented combination of daily global coverage at a high spatial resolution of 3 m. The nanosatellites are called *Doves* and are equipped with sensors that can acquire data at four spectral bands covering the visible to the near-infrared parts of the electromagnetic spectrum. For Patía and Palmira, these were supplied by PS Dove-Classic (PS2) and for Antioquia by PS Dove-R (PS2. SD) satellites. All acquisitions were delivered as atmospherically corrected bottom-of-atmosphere surface reflectance. The selection criteria for suitable acquisitions were image quality and time offset between available satellite imagery and sampling dates. Images that showed cloud, glare or shadow contamination over sampling locations were discarded. Moreover, we excluded forage quantity and quality data from field sampling dates for which no suitable acquisitions within ±5 days were available.

When using PS acquisitions two major sources of uncertainty need to be considered. They are radiometric and geolocation uncertainty. Despite radiometric corrections via a) in-lab sensor calibrations, b) monthly moon imaging (Stone et al., 2020), and c) crossovers with well-calibrated satellites (Helder et al., 2020), an uncertainty of 5–6% (1σ) remains (Wilson et al., 2017). This causes inconsistency between individual PS satellites and hence their acquisitions (Le Roux et al., 2021). The issue of geolocation error is variable and dependent on the quality of the digital elevation model (DEM) and density of ground control points (GCPs) which PlanetLabs uses for geolocation and orthorectifications. The study by Leach et al. (2019) reported reasonably consistent geolocation between PS acquisitions. Nevertheless, PlanetLabs reports a geolocation RMSE of 10 m and other authors reported a high geolocation error as well for their sites (e.g. Cooley et al. (2017): 10.5 ± 3.2 m and Houborg and McCabe (2018): 19 m). Forage cover at both Patía and Antioquia was relatively homogenous. In contrast, the test plots (each 20 × 20 m) at CIAT in Palmira were substantially dissimilar to each other. To compensate for this uncertainty, we tested the model performances using a) no buffer, b) 3 m buffer and c) 6 m buffer from plot boundaries based on which data entries were excluded (see Figs. S4 and S5). Based on that preliminary analysis, we found that applying a 3 m buffer yielded the optimum results for the test plot samples.

2.3.1. Landcover dataset construction

We generated datasets of known landcover from PS satellite acquisitions as the basis for machine learning predictions of landcover and hence non-forage masks. These were constructed by creating training samples of known landcover and other cover using a high-resolution satellite basemap in ArcGIS Pro 2.7 for eight classes (clouds, shadows, forests, grasslands, built land, glares, waterbodies and bare soil) and subsequently reduced to a binary forage and non-forage classification. The number of items per class was adapted to the image size with >100 data points per class. The region-specific statistics are shown in the confusion matrices (see Fig. S6). The apparent class imbalance in this constructed dataset reflects the approximate ratio of forage to non-forage for a given region.

2.4. Data analysis

The data analysis followed a workflow that consists of three main parts: a) pre-processing, b) model development and c) model application and refinement (see Fig. 2). Analyses were carried out using R-project software (R Core Team 2023).

2.4.1. Pre-processing

The pre-processing required the harmonisation of coordinate reference systems (CRS) between sample data and the unstacked images from which we then extracted the spectral values for visible blue, green, red, and near-infrared wavelengths (package:*raster*) (Hijmans, 2023). To test if using a kernel (3 × 3 pixel spatial average) reduces the effect of geolocation error, we additionally created sets of kernel-smoothed rasters from which we extracted the surface reflectance values. Based on these, we calculated various vegetation indices (VIs) that are summarised in Table 2.

The use of VIs is part of the feature engineering process that is common practice in machine learning and well-established in remote sensing of vegetation indicators such as quantity and quality (e.g. Loozen et al., 2019; Raab et al., 2020). The key advantage of using VIs is that they reduce the scatter of site-specific characteristics like illumination angle, the atmosphere the light travels through (e.g. Raab et al., 2020) or radiometric error. The dataset was then split into train (70%) and test (30%) subsets in a way that each subset represents the range and distribution of the original dataset. For regression tasks, we sorted the data from low to high and for clas-

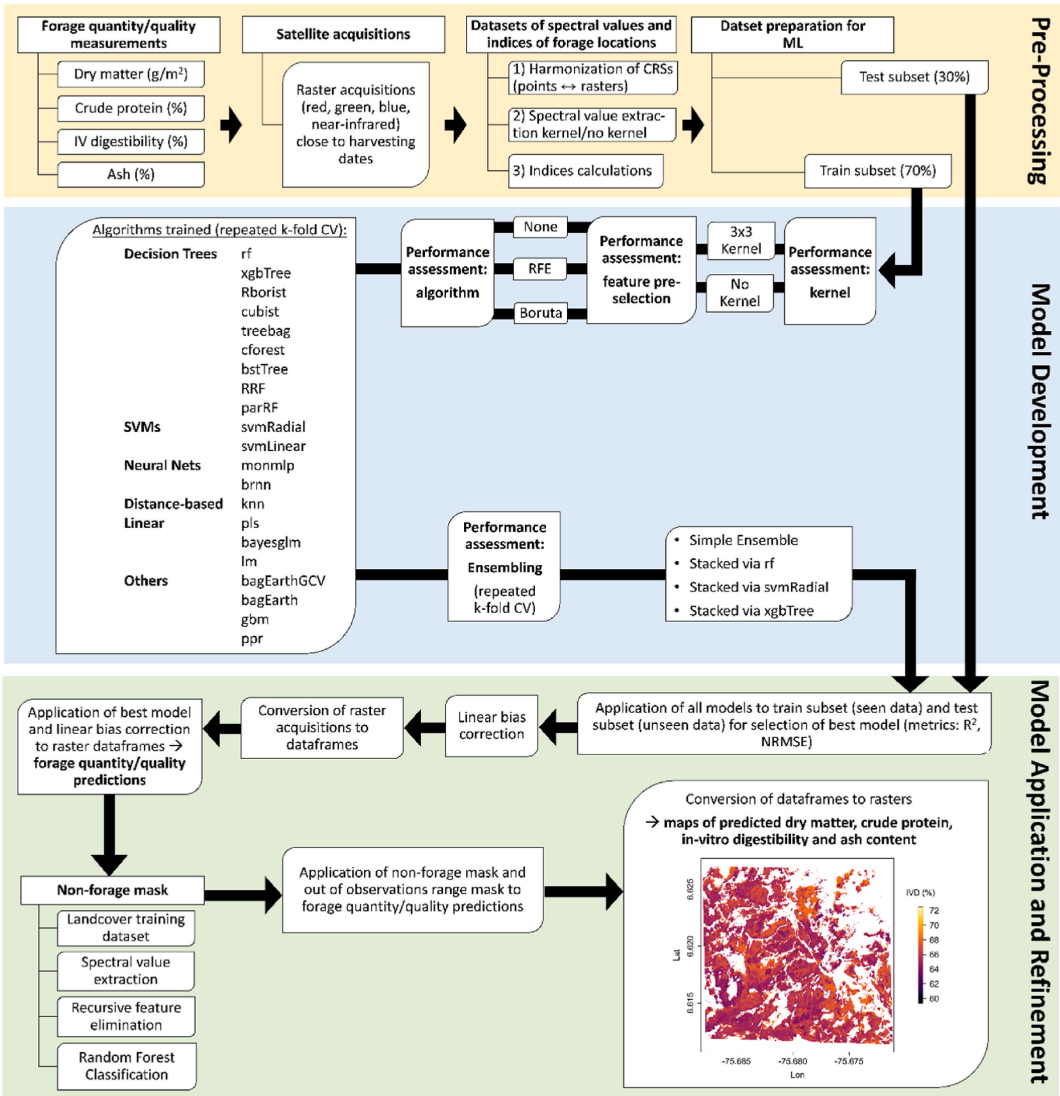

Fig. 2. Workflow of predictive model development. CRS = Coordinate reference system, ML = Machine Learning, RFE = Recursive Feature Elimination, CV = Cross-validation.

Table 2
Vegetation Indices and simple ratios used as additional predictors.

Equations for Vegetation Indices		
[a] $NDVI = \frac{(NIR-Red)}{(NIR+Red)}$	[j] $EVI = 2.5\left[\frac{NIR-Red}{(NIR+6Red-7.5Blue+1)}\right]$	$GR = \frac{Green}{Red}$
[b] $SAVI = \frac{1.5\,(NIR-Red)}{NIR+Red+0.5}$	[k] $LAI = 3.618EVI-0.118$	$GN = \frac{Green}{NIR}$
[c] $MSAVI2 = \frac{2NIR+1-\sqrt{(2NIR+1)^2-8(NIR-Red)}}{2}$	[l] $GLI = \frac{(Green-Red)+(Green-Blue)}{(2Green+Red+Blue)}$	$RB = \frac{Red}{Blue}$
[d] $GOSAVI = \frac{NIR-Green}{NIR+Green+0.16}$	[m] $MSR = \frac{(NIR/Red)-1}{\sqrt{(NIR/Red)}+1}$	$RG = \frac{Red}{Green}$
[e] $RVI = \frac{Red}{NIR}$	[n] $NLI = \frac{NIR^2-Red}{NIR^2+Red}$	$RN = \frac{Red}{NIR}$
[f] $IPVI = \frac{NIR}{NIR+Red}$	[o] $VARI = \frac{Green-Red}{Green+Red-Blue}$	$BR = \frac{Blue}{Red}$
[g] $DVI = NIR\text{-}Red$	[p] $RDVI = \frac{NIR-Red}{\sqrt{(NIR+Red)}}$	$NB = \frac{NIR}{Blue}$
[h] $SR = \frac{NIR}{Red}$	[q] $GNDVI = \frac{(NIR-Green)}{(NIR+Green)}$	$BG = \frac{Blue}{Green}$
[i] $GRVI = \frac{NIR}{Green}$	$GB = \frac{Green}{Blue}$	$BN = \frac{Blue}{NIR}$

[a] Rouse et al. (1974).
[b] Huete (1988).
[c] Qi et al. (2002).
[d] Sripada et al. (2005).
[e] Sripada et al. (2005).
[f] Crippen (1990).
[g] Tucker (1977).
[h] Birth and McVey (1968).
[i] Sripada et al. (2005).
[j] Huete et al. (2002).
[k] Boegh et al. (2002).
[l] Louhaichi et al. (2001).
[m] Chen (1996).
[n] Goel and Qin (1994).
[o] Gitelson et al. (2002).
[p] Roujean and Breon (1995).
[q] Gitelson and Merzlyak (1998).

sification the createDataPartition function (package: *caret*) (Kuhn, 2008) was used before subsampling. For the CIAT (Palmira) site, we excluded data points that fell within the plot boundary buffer of 3 m (see section 2.3. *PlanetScope Acquisitions*) after which 282 data points per response variable remained.

2.4.2. Model development

To determine the optimum models, we tested the improvement capabilities of the 3 × 3 kernel, feature selection approaches, algorithms, and meta learners (simple ensembling and stacks). Several of the applied algorithms have built-in best feature selection functions (see Table 3). To test model improvement capabilities of an independent feature selection approach, we ran all models; a) with no feature pre-selection, b) with Recursive Feature Elimination (RFE, package: *caret*) and c) with Boruta (package: *Boruta*) feature selection. The RFE is a form of backwards selection and in this study combined with 10-fold cross-validation with three repeats. It determines the best features by first partitioning the data into test and train subsets according to the set number of folds followed by tuning and training on the training subset. Afterwards, it was applied to the held-back test subset which forms the basis of predictor evaluation. Initially, it fits a model with all predictors. The feature importance is then ranked and only the best features are retained. Based on that new subset of predictors the process is repeated until only two predictors remain. After all these iterations, the set of predictors that performed best on the test subset are chosen (Kuhn, 2008). The Boruta feature selection algorithm is a wrapper around the random forest algorithm (Kursa et al., 2010). It selects the best features in an iterative process. First, it adds shuffled copies of all features (shadow attributes) and performs random forest classifications/regressions on these copies to obtain Z-scores. Secondly, it identifies the shadow attribute with the highest Z-score and “assigns a hit” to every original attribute that performs better than the shadow attribute with the highest Z-score. This process classifies each feature as either important or unimportant and removes them from the information system together with all shadow attributes. This process is then repeated with the reduced information system until the importance of each feature is determined or the set limit of runs (set to 500) is reached (Kursa et al., 2010).

The sets of best features determined via RFE or Boruta were then stored. For all algorithms including stacking, we used either a) all features, b) RFE or c) Boruta selected features while employing 10-fold cross-validation with 3 repeats for each of the regression algorithms shown in Table 3. The regression algorithms that we tested belong to the classes of a) decision trees, b) support vector machines, c) neural networks, d) distance-based methods, e) linear approaches and f) other unclassified algorithms. In the ML-Ensembles, we either performed a simple ensembling via a meta learner that determines the optimal weights of the predictions of each algorithm or by stacking where the meta learner uses the predictions of every single algorithm and trains on them ideally to outperform them.

2.4.3. Model application and refinement

Fromv all these kernel/no kernel applications, feature selection approaches, algorithms, and meta learners, the best model was selected based on R^2, NRMSE and residuals in the performance on unseen data (test subset). The equations used are shown below:

Table 3
Applied algorithms. Methods with built-in feature selection marked by asterisks.

Algorithm Class	Algorithm	Method	Packages (Version)	Reference
ML-Ensembles	Stacking with SVMRadial superlearner	SVMRadial Stack	caretEnsemble (2.0.1)	Deane-Mayer and Knowles (2016)
	Stacking with eXtreme Gradient Boosting superlearner	XGBTree Stack*	caretEnsemble (2.0.1)	
	Stacking with Random Forests	RF Stack*	caretEnsemble (2.0.1)	
	Simple ensembling with linear regression	Simple Ens	caretEnsemble (2.0.1)	
Decision Trees	Random Forest	rf*	randomForest (4.6–14)	Liaw and Wiener (2002)
	eXtreme Gradient Boosting	xgbTree*	xgboost (1.4.1.1), plyr (1.8.6)	Chen and Guestrin (2016) Wickham (2011)
	Random Forest	Rborist*	Rborist (0.2–3)	Seligman (2015)
	Cubist	cubist*	Cubist (0.3.0)	Kuhn et al. (2012)
	Bagged CART	treebag	ipred (0.9–11), plyr (1.8.6), e1071	Peters et al. (2009) Wickham (2011) Meyer et al. (2019)
	Conditional Inference Random Forest	cforest*	Party (1.7–7)	Hothorn et al. (2010)
	Boosted Trees	bstTree	bst (0.3–23), plyr (1.8.6)	Wang et al. (2020b) Wickham (2011)
	Regularized Random Forest	RRF*	randomForest (4.6–14), RRF (1.9.1)	Liaw and Wiener (2002) Deng (2013)
	Parallel Random Forest	parRF*	e1071 (1.7–7), randomForest (4.6–14), foreach (1.5.1), import (1.2.0)	Meyer et al. (2019) Liaw and Wiener (2002) Calaway et al. (2015) Bache and Torfason (2020)
SVMs	Support-Vector Machines with Radial Basis Function Kernel	svmRadial	kernlab (0.9–29)	Karatzoglou et al. (2004)
	Support-Vector Machines with Linear Kernel	svmLinear	kernlab (0.9–29)	Karatzoglou et al. (2004)
Neural Nets	Multi-Layer Perceptron Neural Network	monmlp	monmlp (1.1.5)	Cannon and Cannon (2017)
	Bayesian Regularized Neural Networks	brnn	brnn (0.8)	Pérez-Rodríguez et al. (2013)
Distance-based	k-Nearest Neighbors	knn	R Base (4.1.0)	R Core Team (2023)
Linear	Partial Least Squares	pls	pls (2.7–3)	Mevik et al. (2015)
	Bayesian Generalized Linear Model	bayesglm	arm (1.11–2)	Gelman et al. (2021)
	Linear Regression	lm	R Base (4.1.0)	R Core Team (2023)
Others	Bagged Multivariate Adaptive Regression Splines using Generalized Cross-Validation Pruning	bagEarthGCV*	earth (5.3.1)	Milborrow et al. (2020)
	Bagged Multivariate Adaptive Regression	bagEarth*	earth (5.3.1)	Milborrow et al. (2020)
	Stochastic Gradient Boosting	gbm*	gbm (2.1.8), plyr (1.8.6)	Greenwell et al. (2019)
	Projection Pursuit Regression	ppr	R Base (4.1.0)	R Core Team (2023)

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{n} (y_i - \check{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{1}$$

y_i = ith value of the variable to be predicted, $\check{y}_i$ = predicted value of y_i, $\overline{y}$ = mean value of a sample, RSS = Residual Sum of Squares, TSS = Total Sum of Squares Bounds: ∞,1

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\check{y}_i - y_i)^2}{n}} \tag{2}$$

n = total number of observations

$$NRMSE = \frac{RMSE}{Q_3 - Q_1} \tag{3}$$

$Q_3 - Q_1$ = Interquartile range

$$Q_1 = CDF^{-1}(0.25) \text{ and } Q_3 = CDF^{-1}(0.75) \quad (4)$$

CDF^{-1} = Quantile function (Cumulative Distribution Function).

The optimum model did sometimes systematically under- or over-estimate the response making a linear bias correction a requirement (based on the linear regression between observed response and predicted response within the training subset). This calibration was then applied to the unseen test subset to obtain the true performance after the linear bias correction was performed. The optimum model and linear bias correction were applied to the PS images. We a) converted the rasters to dataframes, b) calculated VIs, c) applied models and linear bias corrections and d) applied the non-forage mask by multiplying the forage quantity/quality predictions with 1 if underlying vegetation was forage and with 0 if it was non-forage. We also masked predictions that were outside of the training range by setting these to 0 as well. In the final step, the dataframe was converted back to raster and the CRS assigned. Since zero is outside of the training range for all response variables, we set that value as transparent in the subsequent output plots.

3. Results

3.1. Field data

The distribution of forage quantity (DM) and quality (Ash, CP and IVD) that formed the basis for predictive model development is shown for each site in Fig. 3 together with the month and year of the sampling. Antioquia sites had the highest median CP and IVD values and a continuous increase of these two variables throughout the sampling interval. The Palmira site had the highest median DM, which linearly increased throughout the sampled season while CP and IVD decreased. The farms in Patía were characterised by the highest Ash content, but lowest CP, DM and IVD medians. The higher Ash content is likely due to the drier and dusty climate here compared to the Palmira and Antioquia study sites. The correlations between forage variables can be informative in the context of model evaluation and they are shown in Fig. S7. Correlations (Spearman's Rho) between CP and IVD were very high (0.8–0.96) in Patía and Antioquia but substantially lower at the Palmira farm (0.34). The ground-truth datasets were the response values for the predictive models that were developed.

3.2. Overall model performances

In the description of performances, we focused on the median R^2 since this metric is most useful for formulating general recommendations and maximum R^2 to indicate best possible performance where the median receives a higher priority in the rankings. In general, we focused the performance evaluation on test subsets since these are the most representative indicators of true performance on new data. The overall prediction performances (prediction versus observation) of the models are shown in Fig. 4 as violin/box plots indicating distributions, supplemented with comparisons of group means via Wilcoxon rank sum test.

We show performances on train subsets (seen) and test subsets (unseen). As expected, the performance is substantially better on seen data compared to unseen subsets. Overall, the models for Antioquia performed best ($R^2_{median} = 0.33$, $R^2_{max} = 0.73$), followed by Palmira ($R^2_{median} = 0.25$, $R^2_{max} = 0.52$) and Patía ($R^2_{median} = 0.18$, $R^2_{max} = 0.86$). Assessment by response variables show that

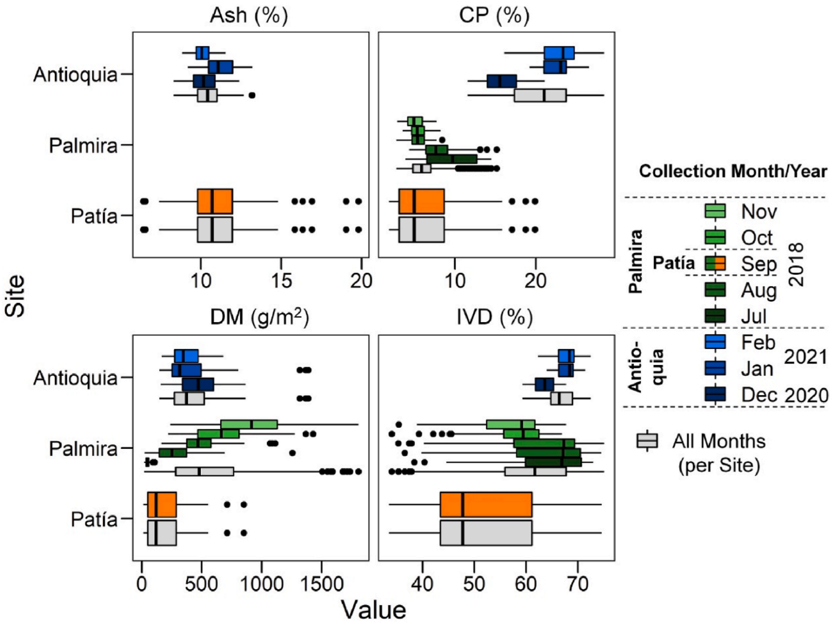

Fig. 3. Ranges of measured values shown as boxplots. CP = Crude Protein, DM = Dry Matter, IVD = In-vitro Digestibility. $N_{Antioqua} = 120$, $N_{Palmira} = 665$, $N_{Patía} = 65$.

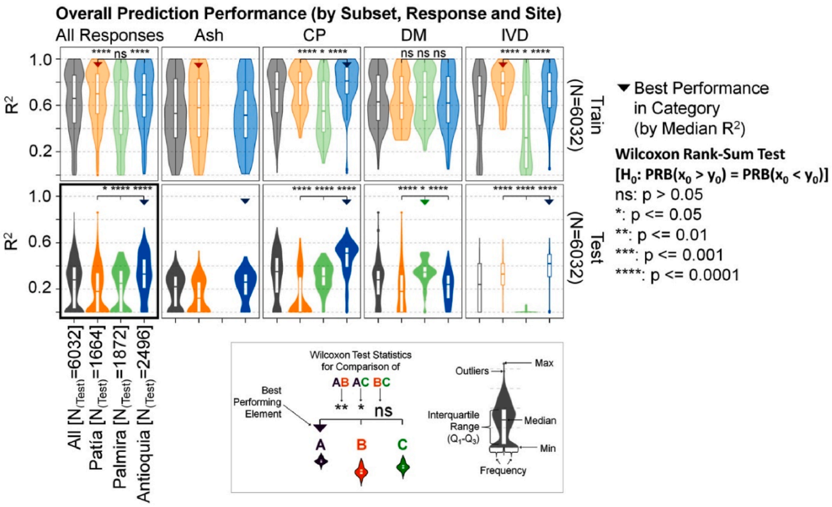

Fig. 4. Prediction performance based on all models shown by responses and train/test performance for all three sites. Key panel highlighted in black and guide to read the figure shown at the bottom. Ash not measured for the Palmira site. CP = Crude Protein, DM = Dry Matter, IVD = In-vitro Digestibility, ns = not significant, PRB = Probability.

CP and DM models performed best (CP: $R^2_{median} = 0.35$, $R^2_{max} = 0.73$; DM: $R^2_{median} = 0.28$, $R^2_{max} = 0.86$) and IVD/Ash models (IVD: $R^2_{median} = 0.24$, $R^2_{max} = 0.64$; Ash: $R^2_{median} = 0.22$, $R^2_{max} = 0.53$) performed less well.

3.3. Improvement capability of 3 × 3 kernel use and feature selection approach

We ran all models with and without a 3 × 3 kernel to determine whether applying a 3 × 3 kernel improved model performance by compensating for geolocation error. The test subset performances are shown in Fig. 5 as violin/box plots with Wilcoxon rank sum test. In general, using a kernel or not did not appear to make a difference when averaged over all sites. But since each site is unique, the use of a kernel may improve performance if surrounding vegetation is similar (Patía) but can reduce performance if adjacent vegetation is dissimilar (Palmira). That observation is in general consistent for each response variable for a given site (see Fig. S8). These findings formed the basis for selecting from kernel models for Patía and from non-kernel models for Palmira. The kernel use does not appear to make a difference for Antioquia, for which no kernel was used in the final best models.

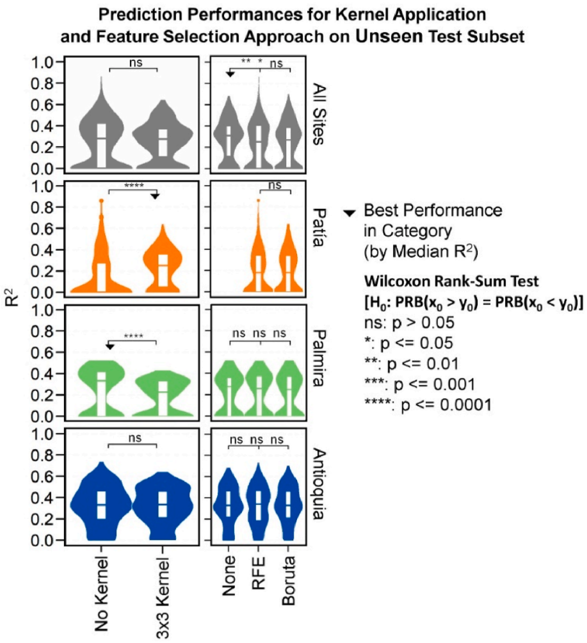

Fig. 5. Performance differences for kernel use and feature selection approach evaluated for test subset. None = no feature pre-selection, RFE = Recursive Feature Elimination.

Feature selection is essential in developing high-performing models. For Patía, we were not able to assess the performance of “no feature selection” since for that site there were too few data entries compared with the total number of features. For Palmira and Antioquia, which had more entries, we tested three approaches a) no feature selection, b) RFE and c) Boruta. We found that in general, there was no approach that systematically outperformed the others. In most cases, leaving the feature selection to algorithms that have it as an in-built function led to slightly better performing models. Nevertheless, it must be decided case-by-case which feature selection approach offers the best models. This is why we did not select one feature selection method over another when selecting the optimal model. Full values by response variables are shown in Fig. S9.

3.4. Assessing algorithms and meta learners

Regularized Random Forest (RRF, $R^2_{median} = 0.38$, $R^2_{max} = 0.65$), Partial Least Squares (pls, $R^2_{median} = 0.36$, $R^2_{max} = 0.66$), Random Forests (rf, $R^2_{median} = 0.36$, $R^2_{max} = 0.63$), Bagged Multivariate Adaptive Regression (bagEarth, $R^2_{median} = 0.35$, $R^2_{max} = 0.64$) and Bayesian Regularized Neural Networks (brnn, $R^2_{median} = 0.34$, $R^2_{max} = 0.61$) were the top performing 5 algorithms and RF Stack (Random Forest Stack, $R^2_{median} = 0.37$, $R^2_{max} = 0.73$) was the best performing meta learner (see Fig. 6). For the prediction of Ash, Support Vector Machines with Linear Kernel (svmLinear, $R^2_{median} = 0.30$, $R^2_{max} = 0.36$) performed moderately, as did Projection Pursuit Regression (ppr, $R^2_{median} = 0.39$, $R^2_{max} = 0.62$) for IVD. The top-ranked algorithms/meta learners – stacking with eXtreme Gradient Boosting superlearner (XGBTree Stack) – yielded higher performances for Patía ($R^2_{median} = 0.34$,

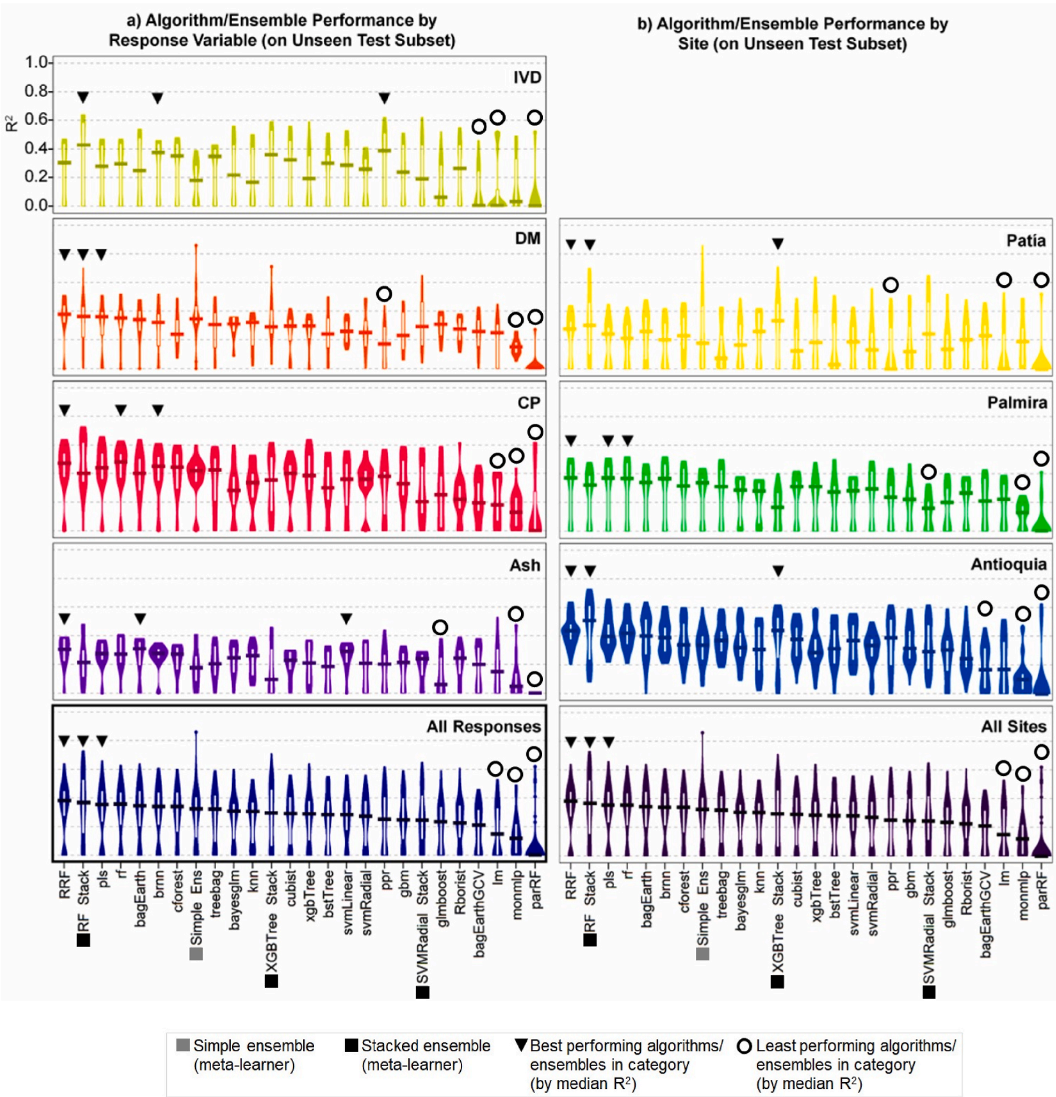

Fig. 6. Algorithm/Ensemble performance shown for responses (a) and for sites (b). Key panel highlighted in black, ensembles marked with squares, best algorithms/ ensembles indicated by black triangles, and least performing algorithms/ensembles indicated by circles.

R^2_{max} = 0.71) and Antioquia (R^2_{median} = 0.44, R^2_{max} = 0.62). Although meta learners do not always outperform single learners, there are occasional exceptions where they yield extreme maxima and hence exceptional performance.

3.5. Best performing models

The best performing kernel-feature selection-algorithm/meta learner combinations after linear correction for the unseen test subsets are displayed in Fig. 7. Since a very high R^2 in the train subset performance can indicate a good performance or a high degree of overfitting, we focus the results description on the test subset and show the train performance as Fig. S10. In the performance evaluation based on unseen test data, CP and DM were predicted relatively well for all three sites. Since the R^2 difference between train subset and test subset was not particularly high for the two response variables, it is likely that the model did not overfit.

This high ability to generalize is also visible in the train/test comparison for Ash in Antioquia and IVD in Patía and contrasts with Ash for Patía, where an initially high R^2 in the training phase (R^2 = 0.91) was substantially reduced in the testing phase, indicating a high degree of overfitting. This overfitting occurred for IVD in Antioquia as well but to a lesser extent. The predictions of IVD in Palmira were poor in the train and in the test phase and consequently thematic outputs were not generated for it. Fig. S11 displays the performance of different models for each forage parameter and study location, along with displaying the performance of individual variables or predictors in the models.

Maps with predicted target variables, shown in Fig. 8, were created by applying for each site the best-performing model (see Fig. 7) and subsequent masking of non-forage areas (see Fig. 2 for complete workflow). In their description, we focus on the timewise change of forage characteristics. At the Palmira site, CP declined slightly between July and November 2018 whereas DM increased substantially during that time. The non-forage mask seemingly performed less well in the application since several grassland plots are masked. At the Los Recuerdos farm in Antioquia, Ash did not seem to change substantially between December 2020 and February

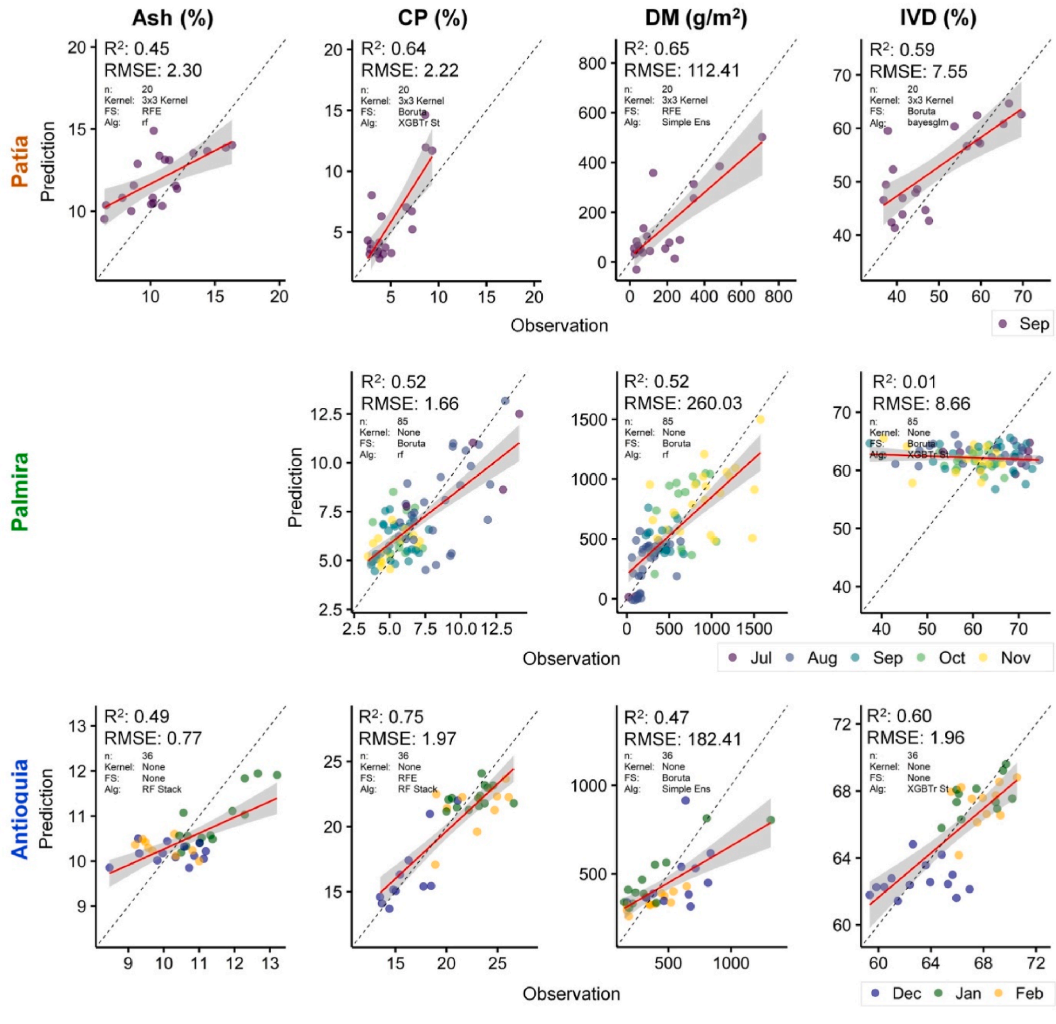

Fig. 7. Scatterplots (test subset) of best performing model predictions vs. observations (color by harvest month) after linear correction shown with performance metrics. Secondary information (small font) indicates number of datapoints, kernel use, feature selection and algorithm/meta learner. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

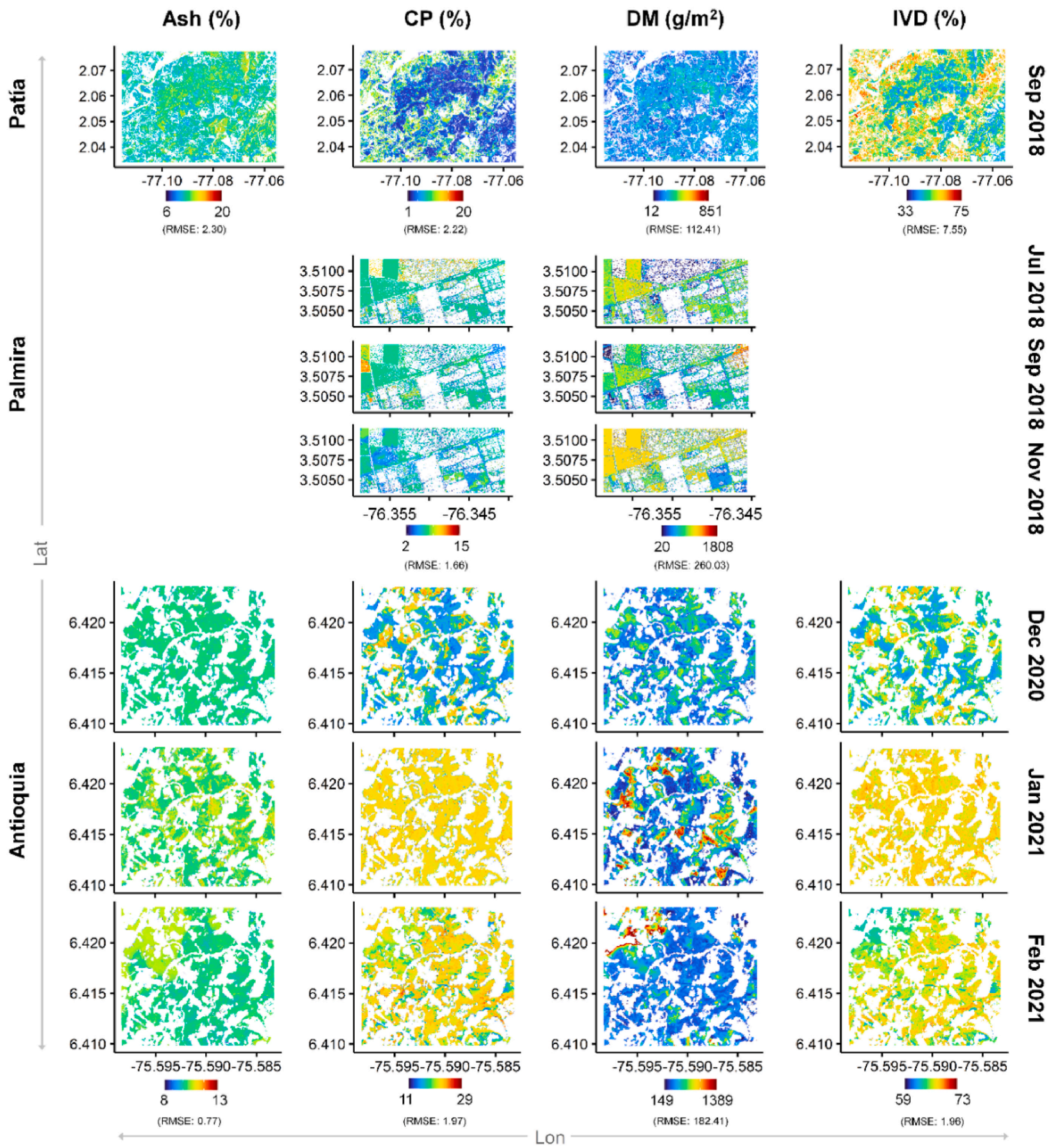

Fig. 8. Example of prediction maps for Patia (Sep 2018), Palmira (July, Sep, and Nov 2018) and Antioquia (Los Recuerdos, Dec 2020–Feb 2021). Additional maps of other months are shown in Fig. S12.

2021. In contrast, CP and IVD increase during that time while the maximum CP occurred in February and the IVD maximum in January 2021. The distribution variability of dry matter is irregular, but a maximum also appears to occur in January 2021.

4. Discussion

4.1. Environmental and anthropogenic controls on DM, CP and IVD and ash

The spatio-temporal change in forage quantity and quality is the result of a complex interplay of environmental, biological, and anthropogenic factors. A summary of these and their influence on DM, CP, IVD and ash is shown in Table 4. Since biomass amount was not always reported as dry matter amount per area, we chose the general term yield in Table 4 to accommodate slight variations in definitions. Based on this synthesis of factors, we can outline likely factors that controlled differences in forage quantity and quality between sites but also during different months for Palmira and Antioquia.

Antioquia showed the highest CP and IVD and lowest Ash values compared to the other two sites, likely due to the intensive pasture management of the dairy farms that were sampled in this region, which contrasts with the lower input systems in Palmira and Patía. Intensive dairy farming in this area is characterised by high nitrogen fertilizer inputs (~250–400 kg/Ha/y - unpublished data from our related studies on similar farms in this locality) and short-interval grazing, and is favoured by its high rainfall, high altitude (Escobar Charry et al., 2020) and latitude (Lee, 2018) leading to cooler temperatures (Ball et al., 2001; Carvajal-Tapia et al., 2021;

Table 4
Key factors influencing forage quantity and quality. Increase in factor (▲) either leads to increase (↗), decrease (↘) or both (↕) in dry matter yield (DM), crude protein content (CP), in-vitro digestibility (IVD) or ash content. For ↕, direction of change is either case or circumstance dependent.

Factor	DM	CP	IVD	Ash	References
Species ↕	↕	↕	↕	↕	(Ball et al., 2001; Carvajal-Tapia et al., 2021; Arango et al., 2014; Ortega-Gómez et al., 2011; Sierra-Montoya et al., 2017)
Polyculture▲	↗	↗	↗	↗	(Maranatha et al., 2019; Sousa et al., 2015)
Cool-season grass▲	↘	↗	↗	↗	(Ball et al., 2001; Bakker and Elbersen, 2005)
Annual grasses▲		↗	↗		Ball et al. (2001)
Maturity at harvest▲	↗	↘	↘		(Holliman, 1986; Ball et al., 2001; Insuasti et al., 2014; Springer and Aiken, 2015; Marais, 2001)
Leaf:stem ratio▲	↘	↗	↗	↗	(Ball et al., 2001; Mganga et al., 2021)
Temperature▲	↗	↘	↘	↗	(Ball et al., 2001; Enciso Valencia et al., 2021; Carvajal-Tapia et al., 2021; Berauer et al., 2020)
Diurnal temperature range▲	↗				Enciso Valencia et al. (2021)
Precipitation▲	↗	↗	↗	↗	(Larbi et al., 2010; Springer and Aiken, 2015; Enciso Valencia et al., 2021; Gándara et al., 2017; González Marcillo et al., 2021; Hare et al., 2015; Larsen et al., 2021; de Vargas Junior et al., 2012; Silva et al., 2017; Lee, 2018; Camacho-Ospina et al., 2021)
Wet season compared to dry season	↗	↗	↗	↗	(Portillo-Lopez et al., 2019; Carvajal-Tapia et al., 2021; Hare et al., 2015; Garay et al., 2017; Lee, 2018; Camacho-Ospina et al., 2021)
Climate Moisture Deficit▲	↘	↘	↘		(Springer and Aiken, 2015; Carvajal-Tapia et al., 2021; Lee, 2018)
Soil moisture▲	↗	↗	↗	↗	(Habermann et al., 2019, 2021; Seguin et al., 2002)
Soil drainage▲	↕	↕	↕	↕	(Habermann et al., 2019, 2021; Lascano et al., 2001; Enciso Valencia et al., 2021; Suescún et al., 2017; Odokonyero et al., 2017; Alvarado-Hernandez, 2015; Ortega-Gómez et al., 2011; Bakker and Elbersen, 2005)
Latitude▲	↘	↗	↗	↗	(Ball et al., 2001; Lee, 2018; Lee et al., 2017)
Altitude▲	↘	↘			(Escobar Charry et al., 2020)
Shadow▲	↕	↗			(Lin et al., 2001, Barragán-Hernández and Cajas-Girón PhD, 2019)
N-Fertilization▲	↗	↗			(Ball et al., 2001; Portillo-Lopez et al., 2019)
P/K-Fertilization▲	↗	↘	↘		Ball et al. (2001)

Berauer et al., 2020). Nevertheless, the experiment in Antioquia took place during the dry season when CP and IVD values tend to be lower (Portillo-Lopez et al., 2019; Carvajal-Tapia et al., 2021; Garay et al., 2017; Camacho-Ospina et al., 2021) and hence the difference between Antioquia and the other two sites from the perspective of these two forage quality metrics was reduced. In terms of monthly change, the maxima in CP and IVD occurred during the driest months which could mean that either previously waterlogging conditions occurred that lead to reduced forage quantity and quality (e.g. Habermann et al., 2019; Suescún et al., 2017; Odokonyero et al., 2017) or that grazing occurred which renewed the forage and hence lead to an increase in the leaf:stem ratio (Ball et al., 2001; Mganga et al., 2021) and hence quality (Holliman, 1986; Ball et al., 2001; Insuasti et al., 2014; Springer and Aiken, 2015). The co-occurrence of DM minima during January and February would suggest that grazing/cutting may be the lead cause. The Ash signal likely was dominated by soil contamination since it is does not seem to correlate substantially with any known controlling factor and because it is higher than 10 % which Ball et al. (2001) suggested is indication for soil contamination in the context of grasses.

In Palmira, the site with the lowest amount of precipitation and highest amount of drought stress, CP and IVD was much lower than in Antioquia. The continuous increase in dry matter and stepwise decrease in CP and IVD could be predominantly controlled by forage maturity related decrease in leaf:stem ratio (Holliman, 1986; Insuasti et al., 2014; Springer and Aiken, 2015) and hence reduction of CP and IVD. Additionally, the increase in precipitation towards the end of the experimental interval may have, to a minor degree, compensated the CP and IVD decrease (Hare et al., 2015; Larsen et al., 2021; Lee, 2018; Camacho-Ospina et al., 2021).

The forage quantity (DM) and quality (CP, IVD) was lowest in Patía compared to the other two sites. Although this region generally receives a similar amount of rain in September as Palmira (Hamann et al., 2013), in 2018 exceptionally little rain (31–85 mm per month) fell from June to October (see Fig. S13). Additionally, its low latitude and altitude leads to a higher temperature and evaporative loss. Prolonged drought stress due to little precipitation and generally high temperatures likely was the main reason for the exceptionally low forage quantity and quality. The Ash content in Patía is well above the 10% threshold, which indicates soil contamination. Since drought conditions occurred it likely was introduced by dust cover.

4.2. Factors influencing model performance

4.2.1. Algorithm performance in predicting forage quantity and quality

Kernel-use may be advantageous if geolocation error occurs and adjacent vegetation is relatively homogenous in its quantity and quality metrics. We showed that for Patía it improves models, for Palmira it is inadvisable due to dissimilar adjacent vegetation and that it does not make a major difference for the Antioquia farms. We also showed that feature selection approaches (RFE, Boruta) do not seem to have a major influence on model quality since many established algorithms have a built-in optimal feature selection approach that is tailored to the respective algorithm. There are feature selection approaches not considered in this study that have been found to perform well and minimize overfitting (e.g. forward feature selection (FFS), Meyer et al. (2018)) that could potentially have a stronger influence on overall model quality and in future studies these could be explored. As shown in Fig. S9, it is likely that feature selection approaches must be decided upon on a case-by-case basis to identify the optimum model for a specific site and/or parameter.

The best performing algorithms/meta learners were; a) random forests (RF), b) regularized random forests (RRF), c) extreme gradient boosting decision trees (xgbTree), d) partial least squares regression/projection to latent structures (PLS) and meta learners using e) ensembling (Simple Ens) and f) stacking via RF or xgbTree as meta learner. All but one (PLS) of the best performing algorithms are ensemble learners because these are highly suitable for non-linear modelling with relatively small number of samples compared to the optimal requirements of neural networks. These ensemble learners can be used in conjunction with bootstrap aggregation or bagging (RF), regularized (RRF) or boosted (xgbTree). Despite PLS not being an ensemble learner, it was shown that it also performs well in cases of small number of samples compared to number of features (Abdi, 2003, 2010). Although the three sites have substantially different number of samples that go into the training phase, no clear pattern is visible. To put the performances of models in this study in context with previous studies, representative examples are displayed in Table 5. When considering model performances, the influences of satellite-based factors of uncertainty and biological factors can also influence modelling capabilities. Further details on aspects of uncertainty introduced by satellite geolocation error, radiometric uncertainty, and differences in spectral band ranges/widths, especially between generations of PS satellites are discussed in section 4.2.2. Similarly, further discussion on how phenology and drought-induced grass senescence may affect the ability to predict tropical forage quantity and quality is provided in section 4.2.3.

4.2.2. *Satellite-based factors of uncertainty*

Understanding if and how a chosen satellite product introduces uncertainty and potentially limits model quality is important. In Patía, geolocation error unlikely affected the models substantially since the vegetation was relatively homogenous. We used only a

Table 5
Examples of remote-sensing-based prediction performances from the literature.

Region	Vegetation	Forage variable	Satellite/UAV (year)	Algorithms	Max R^2 (unseen subset)	Reference and notes
Germany (Alps)	Grass	YLD	UAV (2018)	*RF**	*0.25*	Schucknecht et al. (2022)
				*GBM**	*0.19*	
Australia	Grass & Crops	YLD	Various satellite-based acquisitions (2008–2011)	*PLSR*	***0.74***	Newman and Furbank (2021) model performance 200 days after sowing
				NBCRF	***0.70***	
				XGBM	***0.69***	
				xvBCRF	***0.69***	
				RPRM	*0.59*	
				SVM	*0.45*	
Belgium	Grass	YLD	UAV (2019)	*MLR**	*0.47*	Michez et al. (2020)
		Ash			*0.48*	
		CP			*0.54*	
		IVD			***0.83***	
Germany	Grass	CP	UAV (2018)	*SVM*	***0.79***	Wijesingha et al. (2020)
				CBR	***0.77***	
				RF	***0.74***	
				GPR	***0.73***	
				PLSR	*0.48*	
USA	Grass (Wheat)	YLD	MODIS (2007–2018)	*AdaBoost*	***0.86***	Wang et al. (2020a)
				RF	***0.85***	
				DNN	***0.83***	
				SVM	***0.82***	
				LASSO	***0.81***	
				MLR	***0.75***	
				*AdaBoost**	***0.72***	
Germany	Grass	YLD	Sentinel- 1 & 2 (2015–2017)	*RF**	*0.45*	Raab et al. (2020)
		CP			***0.72***	
China	Grass (Wheat)	YLD	MODIS (2001–2014)	*RF*	***0.81***	Han et al. (2020)
				GPR	***0.79***	
				SVM	***0.77***	
Brazil	Grass	YLD	PlanetScope (2019)	*XGBM**	*0.50*	Dos Reis et al. (2020)
				*RF**	*0.48*	
Brazil	Grass (Rye)	YLD	PlanetScope/UAV (2017)	*SVM**	*0.58*	Breunig et al. (2020)
USA	Grass	YLD	PlanetScope/UAV (2017–2018)	*GBM*	***0.70***	Liu et al. (2019)
Mongolia	Grass	YLD	MODIS NBAR (2006–2016)	*CRT*	***0.62***	John et al. (2018)
Republic of Georgia	Grass	YLD	RapidEye (2014)	*RF*	*0.42*	Magiera et al. (2017)
Ireland	Grass	YLD	MODIS (2001–2012)	*ANFIS*	***0.78***	Ali et al. (2016b)
				ANNs	*0.54*	
				MLR	*0.29*	

Notes: Models that only use spectral values and VIs as predictors marked with *. YLD = yield, CP = crude Protein, IVD = in-vitro digestibility, AVHRR = Advanced Very High Resolution Radiometer, GBM = stochastic gradient boosting, RF = random forests, SVM = support-vector machines, RPRM = recursively partitioned regression models, NBCRF = naïve Breiman-cutler random forests, xvBCRF = Breiman-cutler random forests cross-validated, XGBM = extreme gradient boosting models, LASSO = least absolute shrinkage and selection operator, DNN = deep neural network, MLR = multiple linear regression, *GPR* = Gaussian process regression, PLSR = partial least squared regression, CRT = Cubist Regression Trees, ANN = Artificial Neural Network, ANFIS = Adaptive Neuro-Fuzzy Inference System.

single satellite acquisition which circumvented possible issues like radiometric scatter introduced by several different satellites. Nevertheless, the employed PS Dove-Classics (PS2) with wider and in part overlapping spectral bands [in contrast to PS Dove-R (PS2.SD)] likely led to a reduction in the quality of spectral predictors/vegetation indices. The research plot in Palmira was affected by geolocation error despite our efforts to reduce its severity as the plots were relatively small (20 × 20 m) compared to the image resolution (1 pixel = 3 m). A geolocation error ranging from below 1 pixel (less than 3 m) (Leach et al., 2019) to an extreme of 19 m (Houborg and McCabe, 2018) have been reported in the literature. This paired with relatively dissimilar plots due to different grass species likely introducing scatter. Moreover, radiometric uncertainty due to various acquisitions (several per month from June to November 2018) and hence different satellites likely introduced additional uncertainty. An assessment of radiometric uncertainty showed that even after corrections, an error of 5–6% (1σ) remains (Wilson et al., 2017), leading to inconsistencies between acquisitions (Liu et al., 2017). Like Patía, only the PS Dove-Classics were available and hence a reduced quality of spectral predictors/vegetation indices due to aforementioned reasons. Antioquia benefitted from the availability of the newer PS generation PS Dove-R (PS2.SD) acquisitions. Radiometric uncertainty and differences between satellites due to the use of several acquisitions (one per month from December 2020 to February 2021) likely introduced some error. Nevertheless, it is possible that improvements in the spectral ranges and widths of this new generation PS satellite increased the quality of spectral predictors/vegetation indices. This could be a contributing factor for higher R^2 between best model predictions and observations for Antioquia sites compared to Patía and Palmira.

4.2.3. Biological aspects

In this study, the plot in Palmira allows direct insights into phenological succession and potentially problematic phases for model building since field and spectral data is available from July (planting month) to November 2018. In Antioquia, the phenological succession likely was disrupted by grazing and grasses were not observed from planting to maturity stage. The farms in Patía were exclusively observed during three days in September thus we only can make limited statements about the effect on drought-induced plant senescence and only via comparison with the other two sites. If month-wise change in residual magnitude is largely caused by a partial decoupling of spectral response and forage variable due to growth stage/phenology, then it could serve as evidence. Based on that, for DM modelling later growth stages and for CP earlier growth stages are more difficult to predict. For Antioquia and modelling of DM, a similar interpretation arises if we assume that post-grazing months (January and February) represent early growth stages. Nevertheless, the change in scatter width for CP predictions are too subtle to make statements about potentially problematic growth stages. Han et al. (2020) had similar findings for yield. They developed models with spectral predictors of winter wheat yield in China and found that the further away from the sowing date the higher the loss in accuracy. This observation was most pronounced in SVM models, and the effect of growth stage nearly absent in RF models. It is unclear whether the limiting factor is a specific growth stage or the generalization ability of a chosen algorithm. Newman and Furbank (2021) made contrasting observations in their crop yield prediction models in Australia which also included spectral predictors. Their findings showed that the further away from the sowing date, the better the models were in the prediction of crop yield. This could be caused by a stronger influence of soil spectral reflectance at initially lower vegetation density. Our findings suggest that combining all growth stages produces the best models largely caused by a wider range which agrees with Zeng and Chen (2018). They modelled forage biomass and nutrients at different growth stages in Montana and showed that their yield and CP models perform best when they include all three growth stages. Nevertheless, models from the peak growth stage also generated a high R^2 when this phase is used exclusively.

It is possible that grass senescence/dormancy as the last phase of a grass life cycle or drought-stress can cause issues for spectral prediction of forage characteristics (Royimani et al., 2021). It leads to chlorophyll degradation and carotenoid retention and is accompanied by an increase of spectral reflectance at 550 and 740 nm and reduction at 400 and 500 nm (Merzlyak et al., 1999). Moreover, senescence reduces vegetation density and the influence of the background reflectance of soil becomes stronger (Qi et al., 2002). For optimal spectral prediction of forage characteristics, vegetation should not be too dense and not too patchy as in early growth stages or senescence to minimize the influence of the soil background reflectance (Royimani et al., 2021). For tropical grasses such as those investigated in this study, soil moisture is the key controlling factor of plant senescence. In contrast, at higher latitudes temperature and day length control phenology (Archibald and Scholes, 2007; Castro and Sanchez-Azofeifa, 2008). Soil moisture was very low at the Patía farms due to ongoing moisture deficit (see Fig. S13) that caused a relatively low forage quantity and quality compared to the other two sites but may also have reduced model prediction accuracy due to grass senescence. In Palmira, sufficient rain fell during most of the investigated months and continuous senescence-related issues were unlikely. Nevertheless, during August, the amount of precipitation was reduced, coinciding with increased scatter of CP predictions of the same month. The farms at Antioquia were also investigated during the dry season but since the evaporative loss is much lower at these high altitudes (2320–2620 masl) and low temperatures, no major moisture deficit occurred. In addition to all aforementioned mechanisms that may reduce predictive accuracy, a higher diversity of grass species/cultivars may also make it more difficult to model forage characteristics. In Patía and Palmira our models are based on a relatively high grass diversity (see Table 1) whereas our Antioquia models are based on a single grass species (kikuyo *(Cenchrus clandestinus* (Hochst. Ex Chiov.)). This may provide additional explanation for the slightly better performing models of Antioquia.

5. Conclusions

This study presented a workflow for retrieving forage quantity and quality information from high resolution satellite imagery across multiple sites that could help contribute to the sustainable management of tropical grasslands in Colombia. A comprehensive analysis of the performance of different feature and classifier selection approaches (n = 26) was assessed across sites with different climates, forage types and management practices. The study shows that overall, the models for Antioquia outperform those for

Palmira and Patía and the models for DM and CP perform best. The best performing models varied by site and response variable, with Regularized Random Forest, Partial Least Squares, Random Forests, Bagged Multivariate Adaptive Regression and Bayesian Regularized Neural Networks being the top performing algorithms and Random Forest Stack being the best performing meta learner. For each site and forage parameter, the importance of specific predictors varied, with the NIR band and Red-Green ratio generally performing best. Incorporating management information into remote sensing-based quantity and quality predictions could lead to enhanced understanding and predictive capabilities. Cooperation with smallholder farmers to determine their attitudes and potential constraints to mainstreaming such technologies and their outcomes on the ground can also prove beneficial. Through improving communication between earth observation and agricultural communities and the successful integration of satellite-based technologies, future strategies can be implemented for increasing production and improving forage management while maintaining ecosystem attributes and services across tropical grasslands. The development of these approaches is important for operationalising high spatial and temporal resolution satellite data in grassland monitoring.

Ethical statement

I declare that all ethical practices have been followed in relation to the development, writing, and publication of the article: *"Pixels to Pasture: Using Machine Learning and Multispectral Remote Sensing to Predict Biomass and Nutrient Quality in Tropical Grasslands"*.

CRediT authorship contribution statement

Mike Zwick: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Juan Andres Cardoso:** Writing – review & editing, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization. **Diana María Gutiérrez-Zapata:** Writing – review & editing, Resources, Investigation, Data curation. **Mario Cerón-Muñoz:** Writing – review & editing, Resources, Data curation. **Jhon Freddy Gutiérrez:** Writing – review & editing, Investigation. **Christoph Raab:** Writing – review & editing, Software, Resources. **Nicholas Jonsson:** Writing – review & editing, Investigation, Funding acquisition, Conceptualization. **Miller Escobar:** Writing – review & editing, Investigation. **Kenny Roberts:** Writing – review & editing, Investigation. **Brian Barrett:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

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.

Data availability

Data will be made available on request.

Acknowledgements

This study was supported by funding received from UKRI through the RCUK-CIAT Newton-Caldas Fund Sustainable Tropical Agricultural Systems programme (project numbers BB/R022879/1 and BB/S01893*X*/1). This work was carried out as part of the CGIAR Initiatives: Sustainable Animal Productivity and Livestock & Climate. We thank all donors who globally support our work through their contributions to the CGIAR System. The views expressed in this document may not be taken as the official views of these organizations. CGIAR is a global research partnership for a food-secure future. Its science is carried out by 15 Research Centers in close collaboration with hundreds of partners across the globe. The authors would also like to acknowledge Johanna Mazabel and Rodrigo Peña for their assistance in the collection of field samples, and we thank the two anonymous reviewers whose comments and suggestions have helped to improve the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rsase.2024.101282.

References

Abdi, H., 2003. Partial least square regression (pls regression). Encycl. Res. Methods Social Sci. 6, 792–795.
Abdi, H., 2010. Partial least squares regression and projection on latent structure regression (PLS Regression). Wiley Interdiscip. Rev. Comput. Stat. 2 (1), 97–106.
Ali, I., Cawkwell, F., Dwyer, E., Barrett, B., Green, S., 2016a. Satellite remote sensing of grasslands: from observation to management. J. Plant Ecol. 9 (6), 649–671.
Ali, I., Cawkwell, F., Dwyer, E., Green, S., 2016b. Modeling managed grassland biomass estimation by using multitemporal remote sensing data—a machine learning approach. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 10 (7), 3254–3264.
Alvarado-Hernandez, A., 2015. Plant nutrition in tropical forestry. Tropical forestry handbook 2, 1113–1202.
Arango, J., Moreta, D., Núñez, J., Hartmann, K., Domínguez, M., Ishitani, M., Miles, J., Subbarao, G., Peters, M., Rao, I., 2014. Developing methods to evaluate phenotypic variability in biological nitrification inhibition (BNI) capacity of Brachiaria grasses. Trop. Grasslands-Forrajes Tropicales 2 (1), 6–8.
Archibald, S., Scholes, R.J., 2007. Leaf green-up in a semi-arid African savanna-separating tree and grass responses to environmental cues. J. Veg. Sci. 18 (4), 583–594.
Asmerom, Y., Baldini, J.U., Prufer, K.M., Polyak, V.J., Ridley, H.E., Aquino, V.V., Baldini, L.M., Breitenbach, S.F., Macpherson, C.G., Kennett, D.J., 2020. Intertropical convergence zone variability in the Neotropics during the common Era. Sci. Adv. 6 (7), eaax3644.
Bache, S.M., Torfason, M.T., 2020. Import: an Import Mechanism for R.
Bakker, R.R., Elbersen, H.W., 2005. Managing ash content and quality in herbaceous biomass: an analysis from plant to product. 14th Eur. Biomass Conf. 17, 21.

Ball, D.M., Collins, M., Lacefield, G.D., Martin, N.P., Mertens, D.A., Olson, K.E., Putnam, D.H., Undersander, D.J., Wolf, M.W., 2001. Understanding forage quality. Am. Farm Bureau Federation Publication 1 (1), 1–15.
Bardgett, R.D., Bullock, J.M., Lavorel, S., Manning, P., Schaffner, U., Ostle, N., Chomel, M., Durigan, G., L Fry, E., Johnson, D., Lavallee, J.M., 2021. Combatting global grassland degradation. Nat. Rev. Earth Environ. 2 (10), 720–735.
Barragán-Hernández, W.A., Cajas-Girón PhD, Y.S., 2019. Cambios bromatológicos y estructurales en Megathyrsus maximus bajo cuatro arreglos silvopastoriles. Cienc. Techno. Agropecu. 20 (2), 231–258.
Barrett, B., Raab, C., Cawkwell, F., Green, S., 2016. Upland vegetation mapping using Random Forests with optical and radar satellite data. Remote Sens. Ecol. Conserv. 2 (4), 212–231.
Bengtsson, J., Bullock, J.M., Egoh, B., Everson, C., Everson, T., O'connor, T., O'farrell, P., Smith, H.G., Lindborg, R., 2019. Grasslands—more important for ecosystem services than you might think. Ecosphere 10 (2), e02582.
Berauer, B.J., Wilfahrt, P.A., Reu, B., Schuchardt, M.A., Garcia-Franco, N., Zistl-Schlingmann, M., Dannenmann, M., Kiese, R., Kühnel, A., Jentsch, A., 2020. Predicting forage quality of species-rich pasture grasslands using vis-NIRS to reveal effects of management intensity and climate change. Agric. Ecosyst. Environ. 296, 106929.
Birth, G.S., McVey, G.R., 1968. Measuring the color of growing turf with a reflectance spectrophotometer. Agron. J. 60 (6), 640–643.
Boegh, E., Soegaard, H., Broge, N., Hasager, C.B., Jensen, N.O., Schelde, K., Thomsen, A., 2002. Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. Rem. Sens. Env. 81 (2–3), 179–193.
Bonilla-Cedrez, C., Steward, P., Rosenstock, T.S., Thornton, P., Arango, J., Kropff, M., Ramirez-Villegas, J., 2023. Priority areas for investment in more sustainable and climate-resilient livestock systems. Nat. Sustain. 1–8.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Breunig, F.M., Galvão, L.S., Dalagnol, R., Dauve, C.E., Parraga, A., Santi, A.L., Della Flora, D.P., Chen, S., 2020. Delineation of management zones in agricultural fields using cover–crop biomass estimates from PlanetScope data. Int. J. Appl. Earth Obs. Geoinf. 85, 102004.
Calaway, R., Weston, S., Calaway, M.R., 2015. Package 'foreach'. R package 1–10.
Camacho-Ospina, J.A., Cuesta-Peralta, A., García-Contreras, G.A., Sanjuanelo-Corredor, D.W., Arevalo-Obando, L.M., 2021. Nutritional composition, nitrate and nitrite levels of Kikuyo grass and its influence on milk production. Rev. MVZ Córdoba 27, e2119. . e2119.
Cannon, A.J., Cannon, M.A.J., 2017. Package 'monmlp'.
Carvajal-Tapia, J.I., Mazabel, J., Vivas-Quila, N.J., 2021. Classification of Megathyrsus maximus accessions grown in the Colombian dry tropical forest by nutritional assessment during contrasting seasons. Front. Sustain. Food Syst. 362.
Castro, K.L., Sanchez-Azofeifa, G.A., 2008. Changes in spectral properties, chlorophyll content and internal mesophyll structure of senescing Populus balsamifera and Populus tremuloides leaves. Sensors 8 (1), 51–69.
Chen, G., Li, S., Knibbs, L.D., Hamm, N.A., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M.J., Guo, Y., 2018. A machine learning method to estimate PM2. 5 concentrations across China with remote sensing, meteorological and land use information. Sci. Total Environ. 636, 52–60.
Chan, J.C.W., Paelinckx, D., 2008. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Rem. Sens. Env. 112 (6), 2999–3011.
Chen, J.M., 1996. Evaluation of vegetation indices and a modified simple ratio for boreal applications. Can. J. Rem. Sens. 22 (3), 229–242.
Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. Proc. 22nd ACM SIGKDD Int. Conf. KDD, San Francisco, CA, USA 785–794.
Cooley, S.W., Smith, L.C., Stepan, L., Mascaro, J., 2017. Tracking dynamic northern surface water changes with high-frequency planet CubeSat imagery. Rem. Sens. 9 (12), 1306.
Deane-Mayer, Z.A., Knowles, J.E., 2016. caretEnsemble: ensembles of caret models. R package version 2 (0).
Deng, H., Runger, G., 2012. Feature selection via regularized trees. In: 2012 IJCNN. IEEE, pp. 1–8.
Deng, H., 2013. Guided random forest in the RRF package. arXiv preprint arXiv:1306.0237.
Crippen, R.E., 1990. Calculating the vegetation index faster. Rem. Sens. Env. 34 (1), 71–73.
de Vargas Junior, F.M., Socorro, M.M., Setti, J.D.A., Pinto, G., Martins, C.F., da Costa, J.A.A., Magrin, M.N., Camilo, F.R., Montagner, D.B., 2012. Disponibilidade e valor nutritivo de gramíneas tropicais sob pastejo com ovinos. Arch. Zootec. 62 (238), 295–298.
Dos Reis, A.A., Werner, J.P., Silva, B.C., Figueiredo, G.K., Antunes, J.F., Esquerdo, J.C., Coutinho, A.C., Lamparelli, R.A., Rocha, J.V., Magalhães, P.S., 2020. Monitoring pasture aboveground biomass and canopy height in an integrated crop–livestock system using textural information from PlanetScope imagery. Rem. Sens. 12 (16), 2534.
Echeverría-Londoño, S., Newbold, T., Hudson, L.N., Contu, S., Hill, S.L., Lysenko, I., Arbeláez-Cortés, E., Armbrecht, I., Boekhout, T., Cabra-García, J., Dominguez-Haydar, Y., 2016. Modelling and projecting the response of local assemblage composition to land use change across Colombia. Divers. Distrib. 22 (11), 1099–1111.
Enciso Valencia, K.J., Rincón Castillo, Á., Ruden, D.A., Burkart, S., 2021. Risk reduction and productivity increase through integrating Arachis pintoi in cattle production systems in the Colombian Orinoquía. Front. Sustain. Food Syst. 5, 666604.
Escobar Charry, M.A., Cárdenas Rocha, E.A., Carulla Fornaguera, J.E., 2020. Effect of altitude and defoliation frequency in the quality and growth of Kikuyu grass (Cenchrus clandestinus). Rev. Fac. Nac. Agron. Medellín 73 (1), 9121–9130.
Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., 2007. The shuttle radar topography mission. Rev. Geophys. 45 (2).
Ferner, J., Linstädter, A., Südekum, K.H., Schmidtlein, S., 2015. Spectral indicators of forage quality in West Africa's tropical savannas. Int. J. Appl. Earth Obs. Geoinf. 41, 99–106.
Gándara, L., Borrajo, C.I., Fernández, J.A., Pereira, M.M., 2017. Efecto de la fertilización nitrogenada y la edad del rebrote sobre el valor nutritivo de Brachiaria brizantha cv." Marandú". Rev. Fac. Cienc. Agrar. Univ. Nac. Cuyo 49 (1), 69–77.
Garay, J.R., Cancino, S.J., Zárate Fortuna, P., Ibarra Hinojosa, M.A., Martinez Gonzalez, J.C., González Dávila, R.P., Cienfuegos Rivas, E.G., 2017. Dry matter accumulation and crude protein concentration in Brachiaria spp. cultivars in the humid tropics of Ecuador. Trop. Grasslands-Forrajes Tropicales 5 (2), 66–76.
Gelman, A., Su, Y.S., Yajima, M., Hill, J., Pittau, M.G., Kerman, J., Zheng, T., Dorie, V., 2021. Data Analysis Using Regression and Multilevel/Hierarchical Models. [R package arm version 1.12-2].
Gibson, D.J., 2009. Grasses and Grassland Ecology. Oxford University Press.
Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27 (4), 294–300.
Gitelson, A.A., Merzlyak, M.N., 1998. Remote sensing of chlorophyll concentration in higher plant leaves. Adv. Space Res. 22 (5), 689–692.
Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Algorithms for remote estimation of vegetation fraction. Remote Sens. Environ. 80 (1), 76–87.
Goel, N.S., Qin, W., 1994. Influences of canopy architecture on relationships between various vegetation indices and LAI and FPAR: A computer simulation. Rem. Sens. Rev. 10 (4), 309–347.
González Marcillo, R.L., Castro Guaman, W.E., Guerrero Pincay, A.E., Vera Zambrano, P.A., Ortiz Naveda, N.R., Guaman Rivera, S.A., 2021. Assessment of Guinea grass panicum maximum under silvopastoral systems in combination with two management systems in Orellana Province, Ecuador. Agric. For. 11 (2), 117.
Greenwell, B., Boehmke, B., Cunningham, J., Developers, G.B.M., Greenwell, M.B., 2019. Package 'gbm'. R package version 2 (5).
Habermann, E., de Oliveira, E.A.D., Delvecchio, G., Belisario, R., Barreto, R.F., Viciedo, D.O., Rossingnoli, N.O., de Pinho Costa, K.A., de Mello Prado, R., Gonzalez-Meler, M., Martinez, C.A., 2021. How does leaf physiological acclimation impact forage production and quality of a warmed managed pasture of Stylosanthes capitata under different conditions of soil water availability? Sci. Total Environ. 759, 143505.
Habermann, E., Dias de Oliveira, E.A., Contin, D.R., Delvecchio, G., Viciedo, D.O., de Moraes, M.A., de Mello Prado, R., de Pinho Costa, K.A., Braga, M.R., Martinez, C.A., 2019. Warming and water deficit impact leaf photosynthesis and decrease forage quality and digestibility of a C4 tropical grass. Physiol. Plantarum 165 (2), 383–402.
Hamann, A., Wang, T., Spittlehouse, D.L., Murdock, T.Q., 2013. A comprehensive, high-resolution database of historical and projected climate surfaces for western North America. Bull. Am. Meteorol. Soc. 94 (9), 1307–1309.
Han, J., Zhang, Z., Cao, J., Luo, Y., Zhang, L., Li, Z., Zhang, J., 2020. Prediction of winter wheat yield based on multi-source data and machine learning in China. Rem. Sens. 12 (2), 236.

Hare, M.D., Phengphet, S., Songsiri, T., Sutin, N., 2015. Effect of nitrogen on yield and quality of Panicum maximum cvv. Mombasa and Tanzania in Northeast Thailand. Trop. Grasslands-Forrajes Tropicales 3 (1), 27–33.
Helder, D., Anderson, C., Beckett, K., Houborg, R., Zuleta, I., Boccia, V., Clerc, S., Kuester, M., Markham, B., Pagnutti, M., 2020. Observations and recommendations for coordinated calibration activities of government and commercial optical satellite systems. Rem. Sens. 12 (15), 2468.
Hijmans, R., 2023. Raster: geographic data analysis and modeling. R package version 3, 6–21. https://rspatial.org/raster.
Holliman, M.C. (Ed.), 1986, *Forage-animal Management Systems*. Virginia Agricultural Experiment Station. Virginia Polytechnic Institute and State University.
Höskuldsson, A., 1988. PLS regression methods. J. Chemom. 2 (3), 211–228.
Hothorn, T., Hornik, K., Strobl, C., Zeileis, A., 2010. Party: A Laboratory for Recursive Partytioning.
Houborg, R., McCabe, M.F., 2018. A cubesat enabled spatio-temporal enhancement method (cestem) utilizing planet, landsat and modis data. Remote Sens. Environ. 209, 211–226.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Rem. Sens. Env. 25 (3), 295–309.
Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Rem. Sens. Env. 83 (1–2), 195–213.
Insuasti, S., Silva, A., Jader, J., Andrade, J., 2014. Producción de materia seca y calidad del pasto kikuyo P. clandestinum en diferentes niveles de fertilización nitrogenada y en asocio con aliso alnus acuminata en el trópico alto colombiano. In: Anales del 1° Simposio Int. de Arborización de pastos en Regiones Subtropicales.
Irisarri, J.G.N., Durante, M., Derner, J.D., Oesterheld, M., Augustine, D.J., 2022. Remotely sensed spatiotemporal variation in crude protein of shortgrass steppe forage. Rem. Sens. 14 (4), 854.
Jiménez-Sánchez, G., Markowski, P.M., Jewtoukoff, V., Young, G.S., Stensrud, D.J., 2019. The Orinoco low-level jet: an investigation of its characteristics and evolution using the WRF model. J. Geophys. Res. Atmos. 124 (20), 10696–10711.
John, R., Chen, J., Giannico, V., Park, H., Xiao, J., Shirkey, G., Ouyang, Z., Shao, C., Lafortezza, R., Qi, J., 2018. Grassland canopy cover and aboveground biomass in Mongolia and Inner Mongolia: spatiotemporal estimates and controlling factors. Remote Sens. Environ. 213, 34–48.
Just, A.C., De Carli, M.M., Shtein, A., Dorman, M., Lyapustin, A., Kloog, I., 2018. Correcting measurement error in satellite aerosol optical depth with machine learning for modeling PM2. 5 in the Northeastern USA. Rem. Sens. 10 (5), 803.
Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A., 2004. kernlab-an S4 package for kernel methods in R. J. Stat. Software 11, 1–20.
Knox, N.M., Skidmore, A.K., Prins, H.H., Heitkönig, I.M., Slotow, R., van der Waal, C., de Boer, W.F., 2012. Remote sensing of forage nutrients: combining ecological and spectral absorption feature data. ISPRS J. Photogrammetry Remote Sens. 72, 27–35.
Kontschieder, P., Bulo, S.R., Bischof, H., Pelillo, M., 2011. Structured class-labels in random forests for semantic image labelling. In: 2011 Int. Conf. Comp. Vision. IEEE, pp. 2190–2197.
Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Software 28, 1–26.
Kuhn, M., Weston, S., Keefer, C., Coulter, N., 2012. Cubist models for regression. R package Vignette R package version 0.0 18, 480.
Kumar, L., Mutanga, O., 2017. Remote sensing of above-ground biomass. Rem. Sens. 9 (9), 935.
Kursa, M.B., Jankowski, A., Rudnicki, W.R., 2010. Boruta–a system for feature selection. Fundam. Inf. 101 (4), 271–285.
Larbi, A., Hassan, S., Kattash, G., Abd El-Moneim, A.M., Jammal, B., Nabil, H., Nakkul, H., 2010. Annual feed legume yield and quality in dryland environments in north-west Syria: 1. Herbage yield and quality. Anim. Feed Sci. Technol. 160 (3–4), 81–89.
Larsen, R.E., Shapero, M.W., Striby, K., Althouse, L., Meade, D.E., Brown, K., Horney, M.R., Rao, D.R., Davy, J.S., Rigby, C.W., Jensen, K.B., 2021. Forage quantity and quality dynamics due to weathering over the dry season on California annual rangelands. Rangel. Ecol. Manag. 76, 150–156.
Lascano, C.E., Schmidt, A., Barahona, R., 2001. Forage Quality and the Environment.
Le Roux, J., Christopher, S., Maskey, M., 2021. Exploring the use of PlanetScope data for particulate matter air quality research. Rem. Sens. 13 (15), 2981.
Leach, N., Coops, N.C., Obrknezev, N., 2019. Normalization method for multi-sensor high spatial and temporal resolution satellite imagery with radiometric inconsistencies. Comput. Electron. Agric. 164, 104893.
Lee, M.A., 2018. A global comparison of the nutritive values of forage plants grown in contrasting environments. J. Plant Res. 131, 641–654.
Lee, M.A., Davis, A.P., Chagunda, M.G., Manning, P., 2017. Forage quality declines with rising temperatures, with implications for livestock production and methane emissions. Biogeosciences 14 (6), 1403–1417.
Levick, S.R., Shendryk, Y., Johnson, S., Setterfield, S., Rossiter-Rachor, N., 2021. Remote Sensing of Gamba Grass in Northern Australia: an Assessment of Methods and Guidelines for Use.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R. News 2 (3), 18–22.
Lin, C.H., McGraw, M.L., George, M.F., Garrett, H.E., 2001. Nutritive quality and morphological development under partial shade of some forage species with agroforestry potential. Agrofor. Syst. 53, 269–281.
Liu, Y., Ma, L., Wang, N., Qian, Y., Qiu, S., Li, C., Tang, L., 2017. Vicarious radiometric calibration/validation of Landsat-8 operational land imager using a ground reflected radiance-based approach with Baotou site in China. J. Appl. Remote Sens. 11 (4), 44004. 044004.
Liu, H., Dahlgren, R.A., Larsen, R.E., Devine, S.M., Roche, L.M., O'Geen, A.T., Wong, A.J., Covello, S., Jin, Y., 2019. Estimating rangeland forage production using remote sensing data from a small unmanned aerial system (sUAS) and planetscope satellite. Rem. Sens. 11 (5), 595.
Loozen, Y., Karssenberg, D., de Jong, S.M., Wang, S., van Dijk, J., Wassen, M.J., Rebel, K.T., 2019. Exploring the use of vegetation indices to sense canopy nitrogen to phosphorous ratio in grasses. Int. J. Appl. Earth Obs. Geoinf. 75, 1–14.
Louhaichi, M., Borman, M.M., Johnson, D.E., 2001. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. Geocarto Int. 16 (1), 65–70.
Magiera, A., Feilhauer, H., Waldhardt, R., Wiesmair, M., Otte, A., 2017. Modelling biomass of mountainous grasslands by including a species composition map. Ecol. Indicat. 78, 8–18.
Marais, J.P., 2001. Factors affecting the nutritive value of kikuyu grass (Pennisetum clandestinum)-a review. Trop. grasslands 35 (2), 65–84.
Maranatha, G., Manu, A.E., Sobang, Y.U.L., Samba, F.D., Pelokilla, M.R., 2019. The evaluation of nutritive value and in vitro digestibility of Mulato grass (Brachiaria hybrid cv. Mulato) grown under mixed culture system with legume and horticulture plants on dry land. In: IOP Conf. Series: Earth Environ. Sci., 387. IOP Publishing, 012032. 1.
Masenyama, A., Mutanga, O., Dube, T., Bangira, T., Sibanda, M., Mabhaudhi, T., 2022. A systematic review on the use of remote sensing technologies in quantifying grasslands ecosystem services. GIScience Remote Sens. 59 (1), 1000–1025.
Mazabel, J., Worthington, M., Castiblanco, V., Peters, M., Arango, J., 2020. Using near infrared reflectance spectroscopy for estimating nutritional quality of Brachiaria humidicola in breeding selections. Agrosyst. Geosci. Environ. 3 (1), e20070.
Mevik, B.H., Wehrens, R., Liland, K.H., Mevik, M.B.H., Regression, P.L.S., Suggests, M.A.S.S., 2015. Package 'pls'.
Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C.C., Lin, C.C., 2019. e1071: misc functions of the department of statistics, probability theory group (formerly: E1071), TU Wien. R package version 1 (2).
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. Software 101, 1–9.
Merzlyak, M.N., Gitelson, A.A., Chivkunova, O.B., Rakitin, V.Y., 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. Physiol. Plantarum 106 (1), 135–141.
Mganga, K.Z., Ndathi, A.J., Wambua, S.M., Bosma, L., Kaindi, E.M., Kioko, T., Kadenyi, N., Musyoki, G.K., Van Steenbergen, F., Musimba, N.K., 2021. Forage value of vegetative leaf and stem biomass fractions of selected grasses indigenous to African rangelands. Anim. Prod. Sci. 61 (14), 1476–1483.
Michez, A., Philippe, L., David, K., Sébastien, C., Christian, D., Bindelle, J., 2020. Can low-cost unmanned aerial systems describe the forage quality heterogeneity? Insight from a timothy pasture case study in southern Belgium. Rem. Sens. 12 (10), 1650.
Milborrow, S., Hastie, T., Tibshirani, R., Fortran, A.M., Lumley's, T., 2020. Earth: Multivariate adaptive regression splines. R package version 5.1.2. Available online at: https://CRAN.R-project.org/package=earth.

Morais, T.G., Teixeira, R.F., Figueiredo, M., Domingos, T., 2021. The use of machine learning methods to estimate aboveground biomass of grasslands: A review. Ecol. Ind. 130, 108081.
Newman, S.J., Furbank, R.T., 2021. Explainable machine learning models of major crop traits from satellite-monitored continent-wide field trial data. Nat. Plants 7 (10), 1354–1363.
Odokonyero, K., Acuña, T.B., Cardoso, J.A., Jimenéz, J.D.L.C., Rao, I.M., 2017. Effect of endophyte association with Brachiaria species on shoot and root morpho-physiological responses under drought stress. J. Plant Biochem. Physiol. 5 (1), 1–10.
O'Mara, F.P., 2012. The role of grasslands in food security and climate change. Ann. Bot. 110 (6), 1263–1270.
Ortega-Gómez, R., Castillo-Gallegos, E., Jarillo-Rodríguez, J., Escobar-Hernández, R., Ocaña-Zavaleta, E., de la Mora, B.V., 2011. Nutritive quality of ten grasses during the rainy season in a hot-humid climate and ultisol soil. Trop. Subtrop. Agroecosyst. 13 (3), 481–491.
Pérez-Rodríguez, P., Gianola, D., Weigel, K.A., Rosa, G.J.M., Crossa, J., 2013. An R package for fitting Bayesian regularized neural networks with applications in animal breeding. J. Anim. Sci. 91 (8), 3522–3531.
Peters, A., Hothorn, T., Hothorn, M.T., 2009. Package 'ipred'. R Package 2009.
Portillo-Lopez, P.A., Meneses-Buitrago, D.H., Morales-Montero, S.P., Cadena-Guerrero, M.M., Castro-Rincon, E., 2019. Evaluation and selection of forage grass and legume species in Narino, Colombia. Pastos Forrajes 42 (2), 87–97.
Poveda, G., Mesa, O.J., 1997. Feedbacks between hydrological processes in tropical South America and large-scale ocean–atmospheric phenomena. J. Clim. 10 (10), 2690–2702.
Pullanagari, R.R., Dynes, R.A., King, W.M., Yule, I.J., Thulin, S., Knox, N.M., Ramoelo, A., Michalk, D., Millar, G., Badgery, W., 2013. Remote sensing of pasture quality. Proc. 22nd Int. Grasslands Congr. 2013, 6.
Qi, J., Marsett, R., Heilman, P., Bieden-bender, S., Moran, S., Goodrich, D., Weltz, M., 2002. Ranges improves satellite-based information and land cover assessments in southwest United States. Eos, Trans. Am. Geophys. Union 83 (51), 601–606.
R core team, 2023. R: a Language and Environment for Statistical Computing. R foundation for statistical computing, vienna, austria. Available at: https://www.r-project.org/.
Raab, C., Riesch, F., Tonn, B., Barrett, B., Meißner, M., Balkenhol, N., Isselstein, J., 2020. Target-oriented habitat and wildlife management: estimating forage quantity and quality of semi-natural grasslands with Sentinel-1 and Sentinel-2 data. Remote Sens. Ecol. Conserv. 6 (3), 381–398.
Ramoelo, A., Cho, M.A., Mathieu, R., Madonsela, S., Van De Kerchove, R., Kaszta, Z., Wolff, E., 2015. Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data. Int. J. Appl. Earth Obs. Geoinf. 43, 43–54.
Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6 (1), 5989.
Reinermann, S., Asam, S., Kuenzer, C., 2020. Remote sensing of grassland production and management—a review. Rem. Sens. 12 (12), 1949.
Royimani, L., Mutanga, O., Dube, T., 2021. Progress in remote sensing of grass senescence: a review on the challenges and opportunities. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 14, 7714–7723.
Rodríguez-Hernández, N.S., Arango, M., Moreno, L., Argüello, O., Bernal, J., Pérez, O., 2023. Grassland management effect on ecosystems services in the livestock system in an oxisol from the Eastern high plains of Colombia. Front. Environ. Sci. 11, 1107466.
Roujean, J.L., Breon, F.M., 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sens. Environ. 51 (3), 375–384.
Rouse, J.W., Haas, R.H., Deering, D.W., Schell, J.A., Harlan, J.C., 1974. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. No. E75-10354).
Schucknecht, A., Seo, B., Krämer, A., Asam, S., Atzberger, C., Kiese, R., 2022. Estimating dry biomass and plant nitrogen concentration in pre-Alpine grasslands with low-cost UAS-borne multispectral data–a comparison of sensors, algorithms, and predictor sets. Biogeosciences 19 (10), 2699–2727.
Seguin, P., Mustafa, A.F., Sheaffer, C.C., 2002. Effects of soil moisture deficit on forage quality, digestibility, and protein fractionation of Kura clover. J. Agron. Crop Sci. 188 (4), 260–266.
Seligman, M., 2015. Rborist: extensible, parallelizable implementation of the random forest algorithm. R package version 0, 1. 0.
Sierra-Montoya, E., Barahona-Rosales, R., Ruiz-Cortés, Z.T., 2017. Reproductive behavior of crossbred dairy cows grazing an intensive silvopastoral system under tropical dry forest conditions. Arq. Bras. Med. Vet. Zootec. 69, 1–9.
Silva, T.V.S., Sousa, L.F., Santos, A.C.D., Ferreira, A.C.H., Cardoso, R.R., Sousa, J.T.L.D., Jardim, W.C., Ricardo, A.D.S., Almeida, J.S.D., Carvalho, J.B.D., 2017. Nutritional quality of massai grass fertilized with phosphorus and nitrogen and its influence on intake and weight gain of sheep under rotational grazing on quartzipsamment soil. Semina Ciências Agrárias 1427–1438.
Shoko, C., Mutanga, O., Dube, T., 2016. Progress in the remote sensing of C3 and C4 grass species aboveground biomass over time and space. ISPRS J. Photogrammetry Remote Sens. 120, 13–24.
Sousa, L.F., Maurício, R.M., Paciullo, D.S.C., Silveira, S.R., Ribeiro, R.S., Calsavara, L.H., Moreira, G.R., 2015. Forage intake, feeding behavior and bio-climatological indices of pasture grass, under the influence of trees, in a silvopastoral system. Trop. Grasslands-Forrajes Tropicales 3 (3), 129–141.
Springer, T.L., Aiken, G.E., 2015. Harvest frequency effects on white clover forage biomass, quality, and theoretical ethanol yield. Biomass Bioenergy 78, 1–5.
Sripada, R.P., Heiniger, R.W., White, J.G., Weisz, R., 2005. Aerial color infrared photography for determining late-season nitrogen requirements in corn. Agron. J. 97 (5), 1443–1451.
Starks, P.J., Zhao, D., Phillips, W.A., Coleman, S.W., 2006. Development of canopy reflectance algorithms for real-time prediction of bermudagrass pasture biomass and nutritive values. Crop Sci. 46 (2), 927–934.
Stone, T.C., Kieffer, H., Lukashin, C., Turpie, K., 2020. The Moon as a climate-quality radiometric calibration reference. Rem. Sens. 12 (11), 1837.
Suescún, D., Villegas, J.C., León, J.D., Flórez, C.P., García-Leoz, V., Correa-Londono, G.A., 2017. Vegetation cover and rainfall seasonality impact nutrient loss via runoff and erosion in the Colombian Andes. Reg. Environ. Change 17, 827–839.
Sylvester, E.V., Bentzen, P., Bradbury, I.R., Clément, M., Pearce, J., Horne, J., Beiko, R.G., 2018. Applications of random forest feature selection for fine-scale genetic population assignment. Evol. Appl. 11 (2), 153–165.
Tapasco, J., Martínez, J., Calderón, S., Romero, G., Ordóñez, D.A., Álvarez, A., 2015. Impactos Económicos del Cambio Climático en Colombia: Sector Ganadero. Banco Interamericano de Desarrollo. Monografía No. 254. . Washington, DC.
Tapasco, J., LeCoq, J.F., Ruden, A., Rivas, J.S., Ortiz, J., 2019. The livestock sector in Colombia: toward a program to facilitate large-scale adoption of mitigation and adaptation practices. Front. Sustain. Food Syst. 61.
Tilley, J.M.A., Terry, D.R., 1963. A two-stage technique for the in vitro digestion of forage crops. Grass Forage Sci. 18 (2), 104–111.
Tucker, C.J., 1977. Asymptotic nature of grass canopy spectral reflectance. Appl. Opt. 16 (5), 1151–1156.
Tucker, C.J., 1980. A critical review of remote sensing and other methods for non-destructive estimation of standing crop biomass. Grass Forage Sci. 35 (3), 177–182.
Viana, C.M., Freire, D., Abrantes, P., Rocha, J., Pereira, P., 2022. Agricultural land systems importance for supporting food security and sustainable development goals: a systematic review. Sci. Total Environ. 806, 150718.
Wachendorf, M., Fricke, T., Möckel, T., 2018. Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. Grass Forage Sci. 73 (1), 1–14.
Wang, R., 2018. Significantly improving the prediction of molecular atomization energies by an ensemble of machine learning algorithms and rescanning input space: a stacked generalization approach. J. Phys. Chem. C 122 (16), 8868–8873.
Wang, Y., Zhang, Z., Feng, L., Du, Q., Runge, T., 2020a. Combining multi-source data and machine learning approaches to predict winter wheat yield in the conterminous United States. Rem. Sens. 12 (8), 1232.
Wang, Z., Hothorn, T., Wang, M.Z., 2020b. Package 'bst'.
Wickham, H., 2011. The split-apply-combine strategy for data analysis. J. Stat. Software 40, 1–29.
Wijesingha, J., Astor, T., Schulze-Brüninghoff, D., Wengert, M., Wachendorf, M., 2020. Predicting forage quality of grasslands using UAV-borne imaging spectroscopy. Rem. Sens. 12 (1), 126.
Wilson, N., Greenberg, J., Jumpasut, A., Collison, A., Weichelt, H., 2017. Absolute radiometric calibration of planet dove satellites, flocks 2p & 2e. Planets: San Francisco, CA, USA. https://earth.esa.int/eogateway/documents/20142/1305226/Absolute-Radiometric-Calibration-Planet-Dove-Flocks-2p-2e.pdf.

Wu, T., Zhang, W., Jiao, X., Guo, W., Hamoud, Y.A., 2021. Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration. Comput. Electron. Agric. 184, 106039.

Zeng, L., Chen, C., 2018. Using remote sensing to estimate forage biomass and nutrient contents at different growth stages. Biomass Bioenergy 115, 74–81.

Zhou, W., Li, H., Xie, L., Nie, X., Wang, Z., Du, Z., Yue, T., 2021. Remote sensing inversion of grassland aboveground biomass based on high accuracy surface modeling. Ecol. Indicat. 121, 107215.

Zuluaga, A., Etter, A., Nepstad, D., Chará, J., Stickler, C., Warren, M., 2021. Colombia's pathway to a more sustainable cattle sector: a spatial multi-criteria analysis. Land Use Pol. 109, 105596.