Title

Satellite Remote Sensing of Savannas – Current Status and Emerging Opportunities

Short Title: Present and Future of Observing Savannas from Space

Authors

Abdulhakim M. Abdi¹²*, Martin Brandt¹, Christin Abel¹, Rasmus Fensholt¹

Affiliations

¹ Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark
² Centre for Environmental and Climate Science, Lund University, Lund, Sweden

*Corresponding author: Email: hakim.abdi@cec.lu.se

Abstract

Savannas cover a wide climatic gradient across large portions of the Earth’s land surface and are an important component of the terrestrial biosphere. Savannas have been undergoing changes that alter the composition and structure of their vegetation such as the encroachment of woody vegetation and increasing land-use intensity. Monitoring the spatial and temporal dynamics of savanna ecosystem structure (e.g., partitioning woody and herbaceous vegetation) and function (e.g., aboveground biomass) is of high importance. Major challenges include misclassification of savannas as forests at the mesic end of their range, disentangling the contribution of woody and herbaceous vegetation to aboveground biomass, and quantifying and mapping fuel loads. Here, we review current (2010 – present) research in the application of satellite remote sensing in savannas at regional and global scales. We identify emerging opportunities in satellite remote sensing that can help overcome existing challenges. We provide recommendations on how these opportunities can be leveraged, specifically (1) the development of a conceptual framework that leads to a consistent definition of savannas in remote sensing; (2) improving mapping of savannas to include ecologically relevant information such as soil properties and fire activity; (3) exploiting high-resolution imagery provided by nanosatellites to better understand the role of landscape structure in ecosystem functioning; (4) using novel approaches from artificial intelligence and machine learning in combination with multi-source satellite observations, e.g. multi/hyperspectral, synthetic aperture radar (SAR), and light detection and ranging (Lidar), and data on plant traits to infer potentially new relationships between biotic and abiotic components of savannas that can be either proven or disproven with targeted field experiments.
1. Introduction

1.1. The ecogeography of savannas

The savanna biome is characterized by the coexistence of woody (trees and shrubs) and herbaceous (grasses) plants (Figure 1). Savannas are biodiversity hotspots and habitat for several iconic species of megafauna, and provide a home and livelihood for millions of people who live within or adjacent to them [1-3]. Savannas have a broad climatic scope as they receive mean annual rainfall amounts ranging from 300 to 2500 mm [4] and have mean annual temperatures between 14 to 30 °C [5].

FIGURE 1: Observing multiple drivers and processes in a savanna from space. Savannas have a distinct seasonality with alternating dry and wet seasons. They have a continuous layer of grass with sporadic woody cover. The grass layer provides fuel for extensive fires that coincide with the dry season, and there is widespread herbivory from grazing and browsing animals. Savannas are also subject to human activity such as livestock rearing by pastoral nomads, cropland expansion and fuelwood harvesting. The
illustrated satellites exemplify the main types of Earth observation systems that have been used to study savannas at large scales.

Savannas annually store between 36 and 42 tonnes of carbon per hectare [6], representing around a third of terrestrial net primary production, and comprise a critical regulating component of the land carbon sink [7, 8]. Much of this carbon is stored as aboveground biomass (AGB), which is the pool of photosynthesized carbon embodied within vegetation and is both an essential climate variable [9] as well as an important ecosystem service in savannas.

The climate of savannas is distinctly seasonal and alternates between a growing season during which most of the photosynthesis takes place [10] and a non-growing season in which most of the herbaceous vegetation senesces and cures [11]. Tropical savannas have seasonal rainfall patterns with alternating wet and dry seasons. The dry season serves as the curing period for the grass layer, which accumulates and forms the fuel source for future fires [1]. However, unlike their tropical counterparts, temperate savannas do not have a distinctly alternating rainfall seasonality [1], and accumulation of fuel loads (i.e. combustible material) takes place in the winter when the temperatures are too low for grass growth [12].

Savannas contribute substantially to the seasonal variability of global greenhouse gas emissions from fires, and the biome represents more than 80 percent of the 4 million km² of land that is burned globally every year [13]. The carbon that is emitted from savanna fires is sequestered by post-fire vegetation growth [14, 15] and recent evidence points to a limited long-term impact on global atmospheric carbon concentrations [16]. Several species of savanna plants are adapted to fires through evolutionary mechanisms such as thick barks [17] and require this disturbance in order to not be overtaken by other species. Fire allows mesic savannas to persist at the higher end of their rainfall range despite the climate being conducive to forests [18]. For example, at the border between tropical savannas and forests, dry C₄ grasses are replaced by moister leaf litter in the forest understory, thus forming a barrier to the spread of the fire [19]. The intensity and extent of fires can be amplified during periods of extreme drought that have been preceded by above-average rainfall [20] as well as by land management such as fire suppression.

Savannas have been undergoing changes that are altering the composition and structure of their vegetation such as increased rates of woody encroachment [21]. A combination of factors are responsible for these changes and include CO₂ fertilization [22], land use [23], increasing aridity [24], more intense rainfall [25], and changes in livestock densities [26]. Although recent assessments using global ecosystem models have linked savannas to the interannual variability of the global carbon cycle [7, 8], they are still fairly overlooked [27] and eclipsed by forests [28] with which they are sometimes combined [19]. This primarily stems from ongoing ecological debates about the nature of savannas [29] but also from definitions used within remote sensing research that often overlap and are at times inconsistent with one another. Here, we review current use of satellite remote sensing in savannas, provide an overview of the challenges of terminology and large-scale mapping, summarize methods used to extract vegetation information from satellite data, and recommend several areas of emerging opportunities that will help reshape our understanding of savannas.

1.2. Overlapping definitions of savannas and the challenges they pose
Mapping the distribution of savannas is a longstanding challenge [30, 31] as there is little consensus on how to best to define them. For example, semi-arid savannas that receive less than 700 mm of rainfall a year have been conflated with drylands, which includes semi-arid ecosystems. While there is some geographic overlap between the two (Figure 2), the term drylands denotes a climatic construct based on the availability, or lack thereof, of water. On the one hand, savannas comprise dryland ecosystems at the xeric end of their distribution, e.g. in the Sahel. On the other hand, some savanna ecosystems, such as the Cerrado and wet Miombo, receive between 1300 and 1700 mm of rainfall per year, which creates conditions that are favorable for a denser tree canopy than xeric savannas. Thus, at the mesic end of their distribution, savannas are sometimes included within the definition of forests. Characterizing some savannas, e.g. those in the tropics, as forests is problematic because their tree canopy cover, which is sparse relative to adjacent rainforests, leads to the perception that the forest had undergone some form of degradation [32] and requires restoration of the tree canopy density through tree planting programs.

FIGURE 2: Overlap between global grasslands, savannas, forests and drylands. The brown color shows the distribution of temperate, subtropical and tropical grasslands, savannas and shrublands (referred to here collectively as “savannas”) in the Ecoregions 2017 dataset [33]. The distribution of forests (light blue color) is derived from the Global Ecological Zones (GEZ) dataset [34] and includes boreal, temperate, subtropical, and tropical forests. The dark blue color represents the apparent overlap between the Ecoregions 2017 and GEZ datasets, i.e. savannas and forests. The extent of drylands (dotted pattern) is based on the aridity index (AI) according to UNEP [35] where $0.05 \leq AI \leq 0.65$ (i.e. including arid, semi-arid and dry subhumid ecosystems). The green dots represent examples of these ecosystems using imagery from Google Earth Pro: (a) Western shortgrass prairie, (b) Southeast US conifer savannas, (c) Eastern Anatolian steppe, (d) Daurian steppe, (e)
Cape York Peninsula tropical savanna, (f) Central Zambezian wet Miombo, (g) Uruguayan savanna, (h) Brazilian Cerrado.

Figure 3 exemplifies the diversity of both land-cover types and the nomenclatures used to map them in a part of Africa that is predominantly composed of savannas. Out of the five satellite-derived land-cover maps in Figure 3, only one (MOD12Q1) contains classes identified as savannas, others divide the region into a variety of different land-cover types, including forests. This inconsistency has led to a fundamental misunderstanding of savannas [27], which has, in part, caused the biome to be a major target of international initiatives [36-38] aimed at restoring “degraded landscapes” through tree-planting for mitigating greenhouse gas emissions. Indeed, recent studies by Bastin et al. [39] and Bastin et al. [40] suggested that foresting the world’s sparsely vegetated areas, which includes savannas, would help mitigate the effects of climate change by sequestering most of the carbon that is emitted from fossil fuel burning. These studies generated considerable debate [41-47] and criticism for ignoring, among other things, the fact that ecosystem functioning in savannas determines the spatial arrangement of trees and grasses allowing for their coexistence [48].

Most assessments that target the restoration of degraded forests employ end-user products derived from satellite data, such as tree- or land-cover maps. However, these products can contribute to the misconception that areas with sparse tree canopies are in need of reforestation [40]. Tree-cover maps are generally designed to focus on the percentage of a pixel that is covered by a closed canopy, and scattered trees in drier areas that do not form a canopy are largely excluded [49]. Similarly, the delineated classes in land-cover maps are entirely dependent on the land-cover classification system, and if that system does not include savannas as distinct ecosystems, it could support the narrative to prioritize them for reforestation [50]. This problem is illustrated in Figure 4, where the same savanna landscapes are classified differently depending on the land-cover product. There is little agreement in remote sensing literature on how to best define the tree-grass ecosystems we refer to here as savannas. This ambiguity leads to a patchwork of definitions as a result of different land-cover mapping initiatives adopting different nomenclatures. For example, the sparse savanna of the Serengeti is classified as grassland, herbaceous vegetation, shrubland, and tree-covered area (Figure 4). The denser savanna of the wet Miombo, by contrast, is classified as a forest, (closed) deciduous broad-leaf forest, mixed forest, and tree-covered area. Despite the differences in tree canopy cover and community composition, these savannas share the same basic ecological characteristics that define savannas such as tree-grass coexistence, prevalence of fire disturbance and herbivory, seasonal rainfall with similar mean annual amounts (600-1000 mm yr\(^{-1}\), in this particular case), and an extended dry season lasting several months.
FIGURE 3: Example of the diversity in the characterization of a savanna region according to continental-scale maps derived from satellite, climatic, and botanical stratification. The Great Lakes region provides an ideal setting to illustrate the differences in nomenclature. The numbered panels are based on PlanetScope satellite imagery at 3 meter spatial resolution and represent a gradient of woody canopy cover from dense (a) to sparse (d). The bottom panel represents five land cover classifications derived from satellite data: Global Land Cover SHARE (GLC-SHARE) [51], Moderate Resolution Imaging Spectroradiometer Land Cover Type (MCD12Q1) [52], European Space Agency Climate Change Initiative Land Cover (ESA CCI LC) [53], Copernicus Global Land Services (CGLS) [54], GlobeLand30 [55]. Also shown is a map derived from botanical classification based on dominant species, White 1983 [56], and one derived from climatic stratification Koeppen-Geiger [57]. The overview figure in the lower left shows the delineation of subtropical and tropical grasslands, savannas and shrublands in Africa (pink color) according to the Ecoregions 2017 dataset [33]. This delineation is also shown as a black line border in each of the seven maps in the center of the figure.

1.3. Large area monitoring of savannas using satellite remote sensing

To a certain degree, the theoretical concepts that underpin savanna ecology across large scales remain unchanged as much of our current understanding of savannas comes from field observations [58]. The distinct advantage of satellite technology is that it can extend field observations at regular temporal intervals across large areas that are deficient in, or devoid of, field data. For example, AGB measurements from field inventories have a limited spatial and temporal footprint and can be upscaled using the mathematical
relationship between field-collected variables and satellite observations [59] to estimate AGB at different scales [60]. Savannas have played a critical role in the 50-year track record of satellite remote sensing [61] by, for example, serving as test locations for satellite missions such as the Moderate Resolution Imaging Spectroradiometer (MODIS). Furthermore, satellite remote sensing has been widely used in studies of savanna ecology such as drivers of woody plant encroachment [62] and mapping long-term changes in vegetation composition [63]. This has led to satellite remote sensing becoming an indispensable tool for monitoring the composition, structure and functioning of global savannas.

A rapid advance in satellite sensor technology has occurred over the past decade and several satellites have been launched with the capability of observing the Earth at unprecedented levels of detail. This section provides a synthesis of the current (2010 – present) state-of-the-art in satellite remote sensing of savannas. The following subsections discuss developments in the mapping and separation of woody and herbaceous components, estimating aboveground carbon stocks, detecting fires and quantifying fuel loads, and estimating land surface phenology and ecosystem-scale plant hydraulics. Some of the global-scale studies in this review use data and methods that are not necessarily specific to savannas, and their inclusion hinges on whether their results include information about

FIGURE 4: Similarity and contrast in local-scale definitions of an African savanna along a canopy cover gradient. Close-up of a 400 x 500 m area within each map from Figure 3 shows inconsistencies in how tropical savannas along a woody canopy cover gradient are characterized in global products. The numbers within the parenthesis under the title of each satellite-derived product indicates the native spatial resolution in which the product was provided. The (Planet Labs) PlanetScope 3-meter resolution satellite imagery in panels a – d are in the same order as in Figure 3. Each map panel covers identical dimensions as the corresponding PlanetScope satellite image.
savannas. In this way, the studies presented here help, in one way or another, improve our understanding of savannas and will contribute to a better characterization of the biome going forward.

2.1. Separation of woody and herbaceous components

A longstanding technical challenge in the remote sensing of savannas has been the separation of woody and herbaceous components. Insufficient representation of the spatial structure of the woody component (i.e. trees and shrubs) [64, 65] indicates that the heterogeneity and variability of trees in savannas is not fully understood [66]. A likely reason is due to the fact that non-forest trees have previously been ignored as the pixels of commonly used satellite sensors such as MODIS and Landsat are coarser than the spatial resolution of the woody canopy [67]. It is conceivable that this is what led to the woody component receiving less attention relative to the more easily observed herbaceous component. However, the gap in knowledge about the true distribution of woody and herbaceous vegetation led to several large-scale Earth observation studies being conducted with primary focus on woody cover [23, 62, 68–70], and our current understanding is that trees outside forests are more ubiquitous than previously thought [71].

2.1.1. Spatial and temporal partitioning of coarse resolutions

As savannas are highly heterogeneous, the spatial resolutions of most optical sensors are too coarse to capture the diversity of the landscape, and pixels become a mixture of different land-cover types. The spatial partitioning of different signals within coarse-resolution pixels – termed spectral unmixing – requires information about the spectral properties of the contributing materials called endmembers, which in turn requires the collection of pure spectral signatures within a study area as reference samples [72]. Although there are diverse methods for collecting endmembers from coarse-scale satellite imagery, the availability of time series data helps distinguish between the temporal signals of vegetation embodied within each pixel [73]. However, moisture-driven variability in phenology of both the grass and wood layers can pose a challenge to this approach. Moisture conditions in savannas vary considerably within and between years, which alters the response of the woody leaves and the grass layer [74]. This ultimately influences the signal received by optical satellite sensors.

The dependence of vegetation on moisture availability is a characteristic of savannas, and the strong empirical relationship between them is well-known [75]. Soil moisture controls the phenology of the herbaceous component [76] whereas trees obtain moisture from a range of depths that are inaccessible to grasses [77]. The separation of woody and herbaceous components of savannas based on their sensitivity to moisture conditions is an active area of research. Early work was based on the relationship between vegetation greenness, as indicated by the normalized difference vegetation index (NDVI), and rainfall [78], and it was assumed that trees have a high long-term greenness and low sensitivity to rainfall, whereas grasses have a high sensitivity to rainfall and moderate greenness. This approach was later expanded by Anchang et al. [79] using the rain-use efficiency concept to infer the long-term rainfall sensitivity of woody and herbaceous vegetation. In another study, Kahi and Hanan [80] leveraged rainfall sensitivity of savanna vegetation using an allometric relationship that included mean annual rainfall and maximum canopy leaf area index (LAI) of dominant tree species to partition MODIS LAI 250 m pixels over Africa into woody and herbaceous components.
Methods to disentangle woody and herbaceous signals, based on temporal trajectories, include time-series decomposition of NDVI data using the slow varying trend component as a baseline for the tree layer and the rapidly varying seasonal component as indicative of the grass layer [81, 82]. These approaches required an evergreen tree canopy as a baseline as well as noise-free data, conditions that are not applicable in many savannas. Furthermore, the signal of tree foliage in an area with a low canopy cover is relatively small and often overshadowed by other factors such as variations in background reflectance from bare ground or the grass layer [83]. Other studies relied on the phenological separability of woody and herbaceous components based on their phenological traits such as the senescence of the grass layer during the dry season, allowing for better detection of the signal from trees [84-86]. However, the uncertainty of these methods can be considerable due to, for example, the presence of deciduous trees that completely shed their canopy during the dry season.

2.1.2. Mapping at an ecologically-relevant scale

Supervised approaches to endmember extraction have relied on field collection [73] or very-high-resolution (VHR) imagery. Over the past decade, an increasing number of VHR imagery in Google Earth and cloud computing services such as Earth Engine have streamlined spatial endmember collection. The customization afforded by these services provides a range of possibilities – from manual selection of reference samples through visual interpretation [87] to near-automated approaches that utilize VHR imagery to select endmembers in medium-resolution imagery [88] and further to a fully operational software suite such as the Food and Agriculture Organization’s Open Foris Collect Earth [89]. A potential caveat to these approaches is that the heterogeneity in leaf density and greenness caused by variability in rainfall and soil moisture can introduce uncertainties into the mapping process. VHR imagery provide more detailed opportunities for mapping savanna vegetation components. For example, data from the WorldView satellite constellation, which provide spatial resolutions of less than 0.5 m, have been successfully used to identify individual tree species [90] and distinguish trees, shrubs, and grasses [91]. However, these data are commercial and proprietary, and the associated costs and large data volumes have largely limited their use to local scales. An exception is recent work by Brandt et al. [71], who created a wall-to-wall map of individual tree crowns over a sizable region from multitemporal mosaics of VHR satellite data provided under the NextView licensing agreement [92].

Eddy covariance flux towers collect data on micrometeorology, greenhouse gas and energy exchange between the atmosphere and the terrestrial biosphere from a footprint of around 1 km² around each tower [93]. Although the distribution of these towers in savanna-dominated areas is low relative to other biomes [94], existing sites provide opportunities for coupling satellite and flux tower data [93] to further enable the partitioning and mapping of plant functional types into their constituent trees and grasses [74]. Some flux towers are equipped with phenological cameras (PhenoCams) that provide site-level repeat photography [95, 96]. The so-called “regions of interest” delineated within the field of view of PhenoCam images comprise areas of relatively pure grass and tree canopies that enable the separation of their respective phenology [97] and gross primary production [98]. These offer yet another opportunity for upscaling field measurements with satellite observations [99].

2.1.3. Overcoming limitations inherent to optical remote sensing
Multispectral indices used to delineate vegetation greenness saturate in areas with relatively high green biomass [100] such as the savanna of the wet Miombo. This also applies to indices that were designed to be more sensitive to high green biomass such as the enhanced vegetation index [101]. Areas with high biomass often receive large amounts of rainfall and can thus have persistent cloud cover during the wet season, which obscures the optical signal. In contrast, synthetic aperture radar (SAR) microwaves can penetrate the woody canopy regardless of meteorological (cloud) conditions and interact with its structural components such as leaves, branches, trunks. Direct comparison has shown that L-band (\(\lambda = 0.24\) m) SAR outperforms optical data in mapping woody cover [102]. Despite its complimentary advantages, the use of SAR trails that of optical data. The reasons for this are both historical and technical. Optical data for land applications developed a wider user base early on with the launch of Landsat-1 in 1972 and it has comparatively fewer processing steps before the data are ready for analysis. In contrast, SAR data are difficult to interpret [103] and require a relatively good understanding of several parameters, such as horizontal and vertical polarization, in order to be effectively exploited [104]. However, the present-day availability of analysis-ready SAR data [105] may help narrow this gap in usage. SAR remains a viable complement to multispectral optical data, particularly with the free availability of Sentinel-1 and second generation Phased Array type L-band Synthetic Aperture Radar (PALSAR-2) data, as well as recent efforts to popularize the technology such as the online SAR Handbook [106].

### 2.2. Estimation of aboveground biomass

It has long been known that the radar backscatter signal at long wavelengths (\(\lambda \geq 0.24\) m) is sensitive to biomass and interacts with the woody components [107] of the vegetation canopy. Longwave active and passive microwave sensors provide adequate sensitivity to woody structure in savannas and do not saturate in these ecosystems where average AGB densities are less than 100 Mg ha\(^{-1}\) [69, 108, 109] (Figure 5). Furthermore, microwave observations at longer wavelengths represent the entire vegetation column and include both woody and herbaceous components of savannas [110-112]. In passive microwave systems, the surface energy reaching the sensor is small compared to optical sensors, causing the former to have large spatial footprints (i.e. low spatial resolution). For example, Brandt et al. [66] used 25-km spatial resolution passive microwave L-band data from the Soil Moisture and Ocean Salinity mission to estimate short-term trends in AGB over African ecosystems. They found considerable a loss of AGB (\(-0.05\) petagrams of carbon per year) in savannas that was associated with low rainfall. Comparatively, active microwave systems emit their own signal and thus can provide L-band data at a considerably higher spatial resolution. For example, data from the PALSAR family of sensors are provided as 25 m dual polarisation (HH/HV) global mosaics. These have recently been combined with field observations to produce continental-scale baseline maps of aboveground savanna biomass [113].

Spaceborne light detection and ranging (Lidar) is a relatively new technology (2003 – present) and its applicability in savannas has not been thoroughly investigated. Most studies using lidar data in savannas employ airborne laser scanning as it provides high accuracy and detail [114], but at a considerably higher overall cost and smaller spatial coverage than spaceborne lidar [115]. In dense savannas, lidar has the potential to map horizontal and vertical attributes of woody vegetation such as such as canopy height [116], which can be included in models estimating savanna AGB [117]. The usability of lidar technology in savannas is intricately tied to the presence of woody cover as there is a tradeoff between pulse density and accuracy of the estimated attributes. This means that the
herbaceous understory in savannas is not well represented in spaceborne lidar observations, which leads to that component being overlooked in AGB maps based on this technology.

**FIGURE 5**: Variability in aboveground biomass (AGB) estimates from eight tropical savanna ecosystems. The savanna delineation in the central panel (*pink color*) depicts subtropical and tropical grasslands, savannas and shrublands according to the Ecoregions 2017 dataset [33]. The lettered panels show examples of different ecosystems within this area and their corresponding ecosystem-wide AGB in megagrams per hectare (Mg ha\(^{-1}\)): a. Yucatan Dry Forests (47±23), b. Sahelian Acacia Savanna (4±3), c. Central Deccan Plateau Dry Deciduous Forests (15±18), d. Central Indochina Dry Forests (43±55), e. Guianan Savanna (72±105), f. Cerrado (26±43), g. Eastern Miombo Woodlands (29±44), h. Kimberly Tropical Savanna (41±21). The AGB data in this figure is based on Avitabile et al. [118], who harmonized two earlier AGB maps by Baccini et al. [119] and Saatchi et al. [120] with field data at a 1 km spatial resolution.

One of the recently launched spaceborne lidars, the Advanced Topographic Laser Altimeter System (ATLAS) onboard the Ice, Cloud and Land Elevation Satellite-2, has three pairs of laser beams with a 90 m beam separation and a 3.3 km distance between pairs. This is an insufficient sampling density for mapping vegetation structure in savannas [121] and other open ecosystems [122] because the pulses are likely to miss trees scattered across the landscape. The other operational lidar, the Global Ecosystem Dynamics Investigation (GEDI) onboard the International Space Station (ISS), has three lasers that altogether emit eight beams and measure vegetation structure within a 25 m footprint. Footprints are separated by 60 m along the ISS path and beams are separated by 600 m [123]. Even though
GEDI has a denser sampling rate and unlike ATLAS was specifically designed to measure the structure of temperate and tropical forests, it will directly measure only about 4% of Earth’s land surface [123] during the expected operation lifetime of the mission. That said, preliminary results in South Africa [124] seem promising but further research is required to fully assess its applicability in the diverse vegetation structure of savannas.

2.3. Detection and mapping of fire and burned area

Mapping of fire and burned area from space commenced soon after the launch of the first Landsat satellite in 1972 [125]. Thermal infrared imagery represents outgoing longwave radiation and reflects variability in the land surface energy balance due to loss of vegetation cover caused by fire [126, 127]. Satellite observations in the middle and shortwave infrared region are generally used to detect active fires and subsequent estimation of fire radiative power [128]. The estimation of fire radiative power is based on the notion that heat produced by a fixed quantity of biomass is invariant to the type of vegetation (e.g. woody or herbaceous) consumed by the fire [129], which causes savanna fires to have relatively high fire radiative power [130] due to their spatial and structural heterogeneity [131]. Fires in savannas are generally small [132] and evidence shows that they go largely undetected by MODIS [133] fire products, which have spatial thresholds of 10 ha for active fire detection and 169 ha for burned area mapping [134] (Figure 6). Since savanna fires typically occur during the dry season under fairly clear skies, the Sentinel-2 Multispectral Instrument (MSI) can conceivably detect active flaming fires as small as 1 m² [135] through a combination of the red band (centered at 665 nm) and the shortwave infrared band (centered at 2190 nm) due to the contrasting spectral response of these wavelengths to fire [136]. However, on days with sporadic cloud cover and during intense fires that produce thick smoke, the optical signal can be obscured. In such cases, data from Sentinel-1 SAR can be a complement for mapping active fires by penetrating through clouds and smoke [137] as well's providing observations at night.

Burned area mapping relies on a high temporal resolution in order to capture recently burned areas before vegetation regeneration. Thus, the period between the end of a fire and image acquisition is critical to sense newly burned areas [138]. Progress in mapping burned area has opened opportunities for locally adaptive burned area algorithms using complementary datasets at spatial, temporal and radiometric resolutions appropriate for small burned areas. The combination of Sentinel-2 MSI 10 m spatial resolution and thermal bands of the Landsat-8 Operational Land Imager (OLI) provide both a high temporal revisit frequency, and higher spatial resolutions and radiometric sensitivity to capture smaller burned areas [139, 140]. There has also been progress in cross-sensor fusion, for example between Sentinel-1 SAR and Sentinel-2 MSI [141], that overcomes limitations such as meteorological conditions and poor sensor performance. The remote sensing of fire activity under tree canopies is challenging using optical satellite sensors, since woody canopies can obscure the view of burned areas [142]. Here, lidar can prove useful in providing three-dimensional information on vegetation structure to differentiate between different height thresholds and the ground surface [143]. An caveat for using spaceborne lidar to map burned area is that measurements need to be made before and after the fire in order to infer its impact [138]. This may not be possible as post-fire smoke haze or cloud cover might deflect or block the lidar signal, therefore assessments of fire impact are done by comparing burned and unburned areas that are close to one another [143].

For more detailed monitoring of the strong diurnal variability of savanna fires [144], geostationary satellites provide near-real time (e.g. 5 – 15 minutes) observations at spectral
resolutions that are comparable to polar-orbiting satellites. Although the spatial resolution of geostationary satellites is coarser than most polar orbiting satellites, they can still detect fires down to 1 km$^2$ [145]. The use of geostationary satellites for fire detection is an active area of research and currently operational sensors provide global coverage – the *Spinning Enhanced Visible and InfraRed Imager* on board the Meteosat Second Generation satellites covers Africa [146]; the *Advanced Himawari Imager* on board Himawari-8 and Himawari-9 covers Australia and Asia [147]; and the *Advanced Baseline Imager* on board the Geostationary Operational Environmental Satellites covers the Americas [148].

**FIGURE 6: Number of fire ignitions relative to fire size in savannas.** This figure was created by dividing mean monthly fire sizes and ignitions for 2016 from the MODIS-based World Fire Atlas [149]. The black outline indicates the boundary of temperate, tropical and subtropical grasslands, savannas and dry forests according to the Ecoregions 2017 dataset [33].

The herbaceous component of xeric savannas is dominated by C$_4$ grasses and the spatial coverage of these grasses increases with fire frequency at a continent scale [150]. The herbaceous component comprises the main fuel source for fires in savannas [11] as standing biomass (e.g. thick-barked trees) is not necessarily affected by fires. Therefore, the fuel load is the biomass amount that is available for burning [151]. Knowledge of the amount and extent of the fuel load is important from a land management perspective but studies that map savanna fuel load at large scales have been relatively few compared to those that detect fires or map burned area [152]. Recently launched satellite sensors provide new opportunities for mapping fuel loads. For example, in a Brazilian study, spectrally unmixed Landsat-8 OLI and Sentinel-2 MSI data [153] explained 86% of the variation in fuel loads at a sub-pixel level using a Mixture Tuned Matched Filtering method. Spaceborne lidar is particularly promising for large area mapping fuel load because of its ability to estimate canopy height, which can be used to identify different fuel types. This potential of lidar is exemplified in another Brazilian study [154] where the authors used a GEDI-based modeling framework to predict fuel loads of multiple vegetation layers in a savanna at accuracies of 88% for woody fuels and 71% for the total fuel load.

### 2.4. Land surface phenology and plant hydraulics
The estimation of vegetation phenology using satellite data, termed land surface phenology [155], is widely used to study spatial and inter-annual patterns in growth, maturity and senescence of vegetation. It is a useful indicator of climatic and edaphic conditions that influence vegetation growth and the overall response of ecosystems to global change [156]. Woody and herbaceous components of savannas have distinct phenologies whereby grasses are mostly annual, and trees can be either deciduous, evergreen, semi-evergreen, and rain-stimulated [157]. Leaf deployment strategies also differ amongst the deciduous trees and grasses with the former capable of deploying leaves considerably earlier than the latter [158]. Difference in leaf deployment can be used to estimate the timing of green-up and brown-down between trees and grasses that enables their separation using satellite time series data. For example, seasonal metrics extracted from MODIS and SPOT (Satellite pour l’Observation de la Terre) observations were used map woody cover in the Sahel based on the fact that signal of evergreen and certain deciduous trees is more readily detectable in the early dry season with the absence of the underlying grass layer [85].

Photosynthetic activity in savannas during the dry season is primarily due to the woody component, because the herbaceous component generally undergoes full senescence. The woody component employs a continuum of strategies to deal with seasonal drought ranging from drought avoidance (isohydric) to drought tolerance (anisohydric) [20]. Isohydric species exhibit deciduous behavior by closing their stomata and shedding their leaves thereby shutting down photosynthesis and transpiration to avoid losing water [159]. Anisohydric species keep their stomata open and continue photosynthesis and transpiration during periods of drought. SAR and passive microwave observations can estimate water content within the leaves, branches and trunks through an index known as vegetation optical depth (VOD) [160, 161]. The slope of regression between predawn and midday VOD has been shown to approximate an/isohydric conditions across large scales [162], and VOD values for savannas indicate moderate anisohydricity.

The prevalence of deciduous trees in savannas [163] complicates the idea that woody vegetation in savannas is anisohydric. Deciduousness in response to the dry season suggests isohydricity and MODIS-based satellite observations recently revealed large-scale green-up in African savannas well before the onset of the rainy season, a phenomenon termed as “pre-rain green-up” [164, 165]. The advantage of an early leaf flush is the availability of a longer period in which trees can assimilate CO₂ and grow with little competition for resources from grasses and late-greening trees [166, 167]. Site-level observation and plant growth models show that trees must consume stored carbon and nitrogen for maintenance and growth respiration to fuel leaf flush and new shoots as well as have sufficient reserves of water through either stem and root storage, or deep root systems [159, 168, 169].

The exchange of water between the land and atmosphere is a crucial factor controlling the overall storage of water in the surface, subsurface soil, aquifers, and within the savanna vegetation itself. This moisture component can be approximated using observations from the Gravity Recovery and Climate Experiment (GRACE) satellites through terrestrial water storage (TWS) [170]. TWS is an estimate of groundwater, soil moisture, and wet biomass inferred from small changes in Earth’s gravitational field, and is tightly linked to the interannual variability of atmospheric CO₂ growth rate [171]. Using TWS Madani et al. [172] found that the interannual variability of CO₂ uptake in the moist savannas of Africa is controlled by subsurface water. In another study, GRACE TWS and L-band VOD observations suggest that water stored within the stem and trunk plays a crucial role in sustaining early leaf flush in tropical savannas [173].
Geostationary satellites offer unique opportunities for studying plant hydraulics due to the high temporal resolution of their thermal and optical bands that can capture diurnal processes and be used for assessing evapotranspiration and soil moisture status [174]. Initial work on the use of geostationary data in savannas explored the diurnal canopy water status, daily evapotranspiration, surface soil moisture, and phenology [175-179]. Despite this progress, however, there have been surprisingly few recent studies on the use of geostationary data to study the diurnal progression of ecosystem water use and moisture stress in savannas.

3. Recommendations and emerging opportunities

3.1. A consistent definition of savannas in remote sensing research

There is need for a conceptual framework to aid the process of mapping savannas from space. Perhaps the elephant in the room is the issue of how best to define savannas for application in remote sensing and at what scale that definition should hold [180]. A potential mitigating strategy to the myriad of definitions for savannas is to dissect the biome along key characteristics that are observable from space, e.g. rainfall seasonality, distinct wet and dry seasons, tree-dominated overstory with a grass-dominated understory, prevalence of fire as a key disturbance mechanism, and so on. Using this as a baseline, savannas can conceivably be subdivided along a moisture (e.g. mesic, xeric) and canopy cover (e.g. dense, sparse) gradient along with any region-specific characteristics (e.g. Mopane, Eucalypt). Ideally, the term “forest” should be reserved to ecosystems with a closed or near-closed tree canopy and an understory devoid of grasses, and the term “woodland” for ecosystems with a more open tree-dominated overstory that similarly do not have a grass-dominated understory. But ecosystems are far from being representations of the abovementioned exemplary cases, and there will invariably be cases where these definitions are inapplicable. There needs to be a conceptual space [181] were the disparate characteristics that make up savannas and their relative advantages and weaknesses at different scales are discussed. In light of this, the delineation of savannas as distinct biomes that are not degraded forests [19, 182] but an alternative stable state to forests [183] is crucial to keep them ecologically intact.

3.2. Improving land-cover classification workflows

Uncertainties in tree [184] and land-cover maps [185] are caused by misclassification, inadequate spatial resolution and minimum mapping unit, thematic inconsistencies or a combination of all. Incorporating geospatial datasets on key savanna characteristics in land-cover classification workflows can help produce more accurate maps of savanna distribution. Recent evidence points to the fact that ecoregion-based mapping of the terrestrial biosphere is an effective way to delineate biodiversity patterns [186]. We also know that the inclusion of environmental variables increases the accuracy of land-cover maps in complex savanna ecosystems [187, 188]. As is evident from chapter 2 of this review, there has been substantial remote sensing research on the structure, composition and ecophysiology of savannas. Traditional land-cover classification of savannas should be enhanced to include ecologically relevant information such as land surface phenology [189], soil properties [190], rainfall seasonality [191], fire activity [192], and even plant canopy traits and functional diversity [193, 194].
Another approach towards consolidating land-cover classification with ecosystem functioning is the establishment of long-term plots within diverse savannas globally [195] such as the ones that have been instituted in forest science (e.g. forestplots.net [196]). A widely distributed global network will be able to capture climatic and edaphic gradients as well as community biogeography and vegetation structure. The integration of such information with satellite observations via machine learning or mechanistic models will invariably lead to better models that are capable to classify differences between savannas and distinguish them from other similar ecosystems at large scales.

3.3. Incorporating higher resolutions and multi-source observations

Nanosatellites (satellites weighing between 1 and 10 kg) provide opportunities to overcome the traditional tradeoff between high temporal and high spatial resolutions. Operational Earth observation nanosatellites have spatial resolutions ranging from 0.5 m to 5 m and daily revisit times across the globe [197]. According to the Nanosats Database (nanosats.eu), out of the 1674 nanosatellites in orbit (as of June 2021), 620 are dedicated to Earth observation. The company Planet Labs (San Francisco, CA, USA, hereafter Planet) operates 436 of the Earth observation nanosatellites. Savanna ecosystem studies using commercial high resolution satellite data have typically covered areas less than 250 km² using one or two scenes [90, 198, 199] due to the prohibitive cost of procuring the satellite data. In 2020, Norway’s International Climate and Forest Initiative recently partnered with Planet [200] (www.planet.com/nicfi/) to provide free access to analysis-ready monthly mosaics of multispectral satellite imagery covering savannas across most of the tropics. This data will invariably provide insights into the role of landscape structure in the ecosystem functioning of savannas as well as unique opportunities for consistent monitoring of a substantial portion of the global savanna biome.

Data from geostationary satellites have been generally underused in savanna vegetation studies compared to polar-orbiting satellites due to their relatively lower spatial resolution. The large number of spectral bands provided by these geostationary are comparable to those of MODIS [201] (which will soon be retired) and thus offer unique opportunities for the retrieval of ecosystem-scale metrics. Furthermore, the very-high temporal resolution of geostationary satellites can capture diurnal changes in photosynthetic capacity in response to short-term climatic events, which are likely unobserved by polar-orbiting satellites. There is, however, indication of a revived interest in geostationary satellites, for example for studying vegetation seasonality [202], and vegetation-water interactions at the continental scale [203].

Hyperspectral satellite sensors have the potential to provide synoptic and holistic insight into plant functional processes due to their very narrow spectral bands [204, 205] but have been relatively underused due to limited availability compared to multispectral sensors. Hyperspectral sensors provide a wide range of spectral bands providing a full reflectance spectrum between 400 and 2500 nm spaced at intervals ranging from 1 to 10 nm. There are a handful of such sensors currently in orbit (e.g. Environmental Mapping and Analysis Program (EnMAP), Project for On-Board Autonomy-I (PROBA-I), Hyperspectral Imager Suite (HISUI), Precurseor IperSpettrale della Missione Applicativa (PRISMA)), and more are planned in the near future (e.g. Spaceborne Hyperspectral Applicative Land and Ocean Mission (SHALOM)). The fine-scale sampling of spectral bands and medium spatial resolution makes hyperspectral data useful for providing diverse information about the
Two SAR missions with complementary radiometric resolutions relevant for savannas are scheduled for launch in 2023. The European Space Agency’s BIOMASS mission [210] will operate a P-band (f = 435 MHz, λ = 0.68 m) SAR capable of penetrating the canopy of denser savannas, such as parts of the Miombo, and provide estimates of biomass at a spatial resolution of 200 m. The NISAR mission is a collaboration between NASA and the Indian Space Research Organization and will operate an L-band (f = 1.3 GHz, λ = 0.24 m) and an S-band (f = 3.2 GHz, λ = 0.12 m) SAR focused on areas with AGB of < 100 Mg ha\(^{-1}\), which encompasses a most global savannas. These missions will provide more enhanced opportunities to synergize data not only across SAR platforms, but also including lidar and optical sensors, for improved mapping of savanna vegetation structure [211]. For example, because longwave SAR can penetrate through to the understory, it can be fused with observations that capture woody biomass such as GEDI (or Sentinel-1 SAR), which will open up an opportunity to spatially separate herbaceous biomass from woody biomass.

### 3.4. Exploiting machine learning and big data analytics

The efficiency of machine learning (ML) to extract patterns from satellite data made it an attractive tool for the Earth observation community. Breakthroughs in mapping woody vegetation in savannas using ML include detecting individual trees [71], identifying them down to the species level [91], and assessing reforestation effects in savannas [212]. The stage is currently set for the study of ecosystem-scale interactions using ML, which has become increasingly popular in ecology [213], in part due to its relatively few *a priori* assumptions. Interpreting ecological ML models to understand underlying relationships is a challenge [214] as there is generally no direct assessment of how the predictions are generated. ML models often have higher predictive accuracies than traditional statistical learning methods, such as (non-)linear regression. The implication here is that ML models embody functional relationships between the parameters, provided they are not spurious [215]. This is exemplified in Schwieder et al. [216], where the authors used a random forest regression algorithm to link savanna land surface phenology to aboveground carbon and identify the key phenological metrics that define this relationship at each of their study sites.

A noteworthy advantage of ML and big data is scalability and repeatability, which could help break the entry barrier for developing countries to use satellite remote sensing for monitoring savanna resources. Such a framework was proposed by Anchang et al. [49], who combined Sentinel-1 and Sentinel-2 observations using a random forest regression to map percent woody canopy cover in savannas. More recent studies have exploited capabilities of convolutional neural networks to detect individual savanna trees [71] and elephants [217] from VHR satellite observations. These developments represent a research frontier when it comes to satellite remote sensing of savannas and open the door for hyperlocal analysis of the role of large herbivory in these ecosystems. These models, however, are rather opaque, but there are ongoing efforts to make their predictions more transparent, including the development of interpretable machine learning frameworks [218] as an alternative to black-box models.
A challenge for the use of ML with satellite observations to infer ecological relationships is model transferability [219]. Ecological models should ideally be generalizable in order to be applied outside the spatiotemporal space of the data used to train them. In savannas, the large spatial and temporal heterogeneity within ecosystems can limit model generalizability [220]. A potential solution here is model-data assimilation whereby satellite observations and process-based models, such as dynamic global vegetation models, are integrated to improve parametrization and lead to better representation of ecosystem functioning [221]. This is particularly pertinent for savannas where process-based models have generally underperformed [65]. Similarly, outputs from process-based models can be used to constrain ML predictions, creating a bidirectional benefit that could help reduce uncertainties in both ML and process-based model outputs.

The increasing availability of large volumes of data (termed “big data”), many of which are ecologically relevant for savannas, has opened up new avenues for research [222]. For example, plant trait data provide a strong response to ecosystem functioning and various plant traits show stable relationships with both optical reflectance [223] and SAR backscatter [224]. These could be parameterized using ML to develop trait-backscatter-spectral (TBS) models capable of differentiating plant functional types and allowing for their global mapping. The global plant trait database (ver. 5) [226] has nearly 12 million trait records covering 280,000 plant taxa, which can be coupled with both raw satellite observations and derived products (vegetation properties, climate, soil, elevation, disturbance, etc.) within an ML framework to infer undiscovered relationships between biotic and abiotic components of savannas. In essence, allowing for the discovery of unknown unknowns that can then be either proven or disproven with targeted field experiments.

In 2020, the Copernicus Open Access Hub (scihub.copernicus.eu) published 6.3 petabytes of Sentinel-1 SAR and Sentinel-2 MSI data. These large volumes of satellite data pose new challenges to data curation, storage and computing capacity. Hosting of data and analytical tools on public cloud infrastructures such as the Copernicus Data and Information Access Services (copernicus.eu/en/access-data/dias) as well as private cloud computing services such as Amazon Web Services, Microsoft Azure, and Google Earth Engine provide timely opportunities for continental-scale analysis [227] of savannas. For example, recent work has shown the potential of cloud-based analysis of Sentinel-1 SAR and Sentinel-2 MSI time series to map woody canopies in savannas [68]. A key advantage of cloud-based services is the potential for efficient operational workflows since both the data and algorithms are stored in the same infrastructure. This not only allows for near-real time monitoring of savannas, but it is also scalable and can be replicated as per user needs [49]. Cloud-based services such as Google’s Earth Engine (earthengine.google.com) platform are also equitable because they level the playing field for remote sensing scientists in developing countries by eliminating the need for costly computational infrastructure to run large-scale analyses.

4. Conclusions

Our understanding of large-scale savanna ecology and its response to global change has dramatically improved in recent decades, and satellite remote sensing has played a key role in this process. The marked advantage of satellite technology is that it provides continuous observations across large areas. However, it is clear that traditional methods of
land-cover mapping, whereby training samples for specific classes are collected from field surveys or photointerpretation, are insufficient for delineating savannas. More holistic methods that include the unique climatic, edaphic and disturbance characteristics of savannas are necessary for a more complete description of the role of savannas in the terrestrial biosphere. These include, for example, incorporating big geospatial data on vegetation, rainfall, fire, soil, and topography as covariates in classification models along with satellite data.

The past twelve years (2010 – 2021) have witnessed remarkable progress in new satellites and analytical methods that have helped improve our understanding of savannas. Better representation of woody and herbaceous components remains a crucial aspect that underlines many of the structural and functional attributes of savannas. Earth observation studies of savannas have generally used coarse to medium resolution imagery with more attention paid to one or the other component. This will likely change going forward as higher resolution imagery becomes more readily available enabling distinct spatial separation between trees and grasses.

Although new satellite sensors will invariably add to the ecological synthesis of savannas, currently available geostationary, hyperspectral, synthetic aperture radar (SAR), and light detection and ranging (Lidar) technologies remain relatively underused. This potentially hampers synergies of new and old data, particularly where one technology can complement or extend the other in one or more dimensions. To better understand savanna ecosystem responses to climate change and disturbances, both historical and current data need to be utilized. Despite their limitations, we are bound to using historical multi-sensor satellite observations for insight into change dynamics, whereas current and new sensor systems provide high spatial, temporal, and radiometric resolutions for more detailed analysis of ecosystem response. Indeed, the complexity of the savanna biome necessitates spatial (VHR), spectral (hyperspectral), temporal (geostationary), and structural (Lidar, SAR) synergies.

Advancements in the observation, mapping and ecological synthesis of savannas from space will gain momentum in the future, which is a promising outlook for the improved understanding of the large-scale processes that occur in the biome. As remote sensing scientists and ecologists continue to develop ecological frameworks to understand savanna vegetation functioning at large scales, improved communication between disciplines and flexible ontologies are needed to link satellite observations and ecological data. Such cross-disciplinary communication will result in enhanced maps of savanna distribution as well as better predictions of savanna response to climate change and anthropogenic disturbance. This will ultimately lead to more informed conservation efforts of these special ecosystems.

Acknowledgments: We thank Dr. Nicola Stevens, Dr. Nathalie Pettorelli, and Dr. Torbern Tagesson for their helpful comments during the drafting of the manuscript.

Author contributions: A.M.A. conceived the study, led the writing, and designed the graphics (except Figure 1, which was commissioned). All authors (A.M.A., M.B., C.A., R.F.) contributed substantial discussion of content, and extensive edits to the manuscript, its revisions, and the final production.

Funding: This work was supported by the Swedish Research Council – Vetenskapsrådet (grant 2018-00430 to A.M.A); Independent Research Fund Denmark – DFF Sapere Aude (grant 9064-00049B to M.B.); Villum Foundation through the project “Deep Learning and Remote Sensing for Unlocking Global Ecosystem Resource Dynamics” (DeReEco to R.F.).

Conflicts of Interest: The authors declare no conflict of interest.

5. References


