DSPACE

https://dspace.org/

Optical remote sensing of crop biophysical and biochemical parameters An overview of advances in sesnor technologies and machine learning algorithms for precision agriculture

Mahlatse Kganyago; Clement Adjorlolo; Paidamwoyo Mhangara; L Tsoeleng

ELSEVIER SCI LTD
https://hdl.handle.net/10539/41798
WOS

Downloaded from a DSpace repository. https://dspace.org/

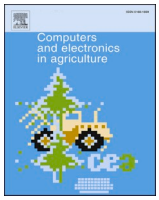
Computers and Electronics in Agriculture 218 (2024) 108730



Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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





Optical remote sensing of crop biophysical and biochemical parameters: An overview of advances in sensor technologies and machine learning algorithms for precision agriculture

Mahlatse Kganyago [a,b,*], Clement Adjorlolo [a,c], Paidamwoyo Mhangara [a], Lesiba Tsoeleng [d]

[a] School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg 2050, South Africa
[b] Department of Geography, Environmental Management and Energy Studies, University of Johannesburg, Johannesburg, South Africa
[c] African Union Development Agency (AUDA-NEPAD), 230 15th Rd, Midrand, 1685, Johannesburg, South Africa
[d] Earth Observation, South African National Space Agency, The Enterprise Building, Mark Shuttleworth Street, Pretoria 0001, South Africa

ARTICLE INFO

Keywords:
Remote sensing
Machine learning
Precision agriculture
Leaf area index
Chlorophyll content

ABSTRACT

This paper provides an overview of the recent developments in remote sensing technology and machine learning algorithms for estimating important biophysical and biochemical parameters for precision farming. The objectives are (i) to provide an overview of recent advances in remotely sensed retrieval of biophysical and biochemical parameters brought by the developments in sensor technologies and robust machine learning algorithms and (ii) to identify the sources of uncertainty in retrieving biophysical and biochemical parameters and implications for precision agriculture. The review revealed that developments in crop biophysical and biochemical parameters retrieval techniques were mainly driven by announcements and the availability of new sensors. Two ground-breaking events can be identified, i.e., the availability of Sentinel-2 and the SuperDove constellation. The two provide high temporal-high spatial resolution data relevant for site-specific management and super-spectral configuration, enabling retrieval of crop growth and health parameters. The free availability of Sentinel-2 triggered the testing of its spectral configurations and upscaling of retrieval approaches using simulated data from field spectrometers and airborne hyperspectral sensors. SuperDoves will likely reduce the cost of very high-resolution data while providing unprecedented capabilities for detailed, accurate and frequent characterisation of field variability. Studies showed that the red-edge bands and hybrid models coupling Radiative Transfer Model (RTM) and machine learning regression algorithms (MLRA) are promising for operational and accurate monitoring of stress-related crop parameters to aid time-sensitive agronomic decisions. However, such models were tested in Mediterranean climates and performed poorly in African semi-arid areas and China's temperate continental semi-humid monsoon climates. Therefore, locally-calibrated RTM models incorporating crop-type maps and other spatio-temporal constraints may reduce uncertainties when adapted to data-scarce regions. Generally, permanent experimental sites and a lack of systematic calibration data on various crops are some limiting factors to using remote sensing technologies for PA in Sub-Saharan Africa. Other complexities arise from farm configurations, such as small field sizes and mixed cropping practices. Therefore, future studies should develop generic, scalable and transferable models, especially within under-studied areas.

1. Introduction

In an era of climate variability and change, dwindling agricultural resources brought on by ineffective land management and competition from other land uses, and the need to address food insecurity, precision agriculture is one of a few viable solutions. It promises the optimisation of farm inputs (i.e., fertilisers, seeds, water and chemicals), improved efficiency and profitability of the agricultural system by averting potential losses over stressed areas, and reduced environmental impact by avoiding the excessive application of inputs (Mulla, 2013). Consequently, many studies report the benefits of integrating specific precision agriculture technologies in farm operations. For example, Bellvert et al. (2020) found cost savings of €7 090 and €9 960 over two consecutive seasons, i.e., 2016 and 2017, respectively, when using

* Corresponding author.
E-mail address: mahlatsek@uj.ac.za (M. Kganyago).

https://doi.org/10.1016/j.compag.2024.108730
Received 29 November 2022; Received in revised form 3 January 2024; Accepted 7 February 2024
Available online 13 February 2024
0168-1699/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

optimised precision irrigation based on an integrated vine water consumption model and remote sensing data in a commercial vineyard of 100 ha. However, precision agriculture faces several limitations, such as information accuracy, high data volumes, complex information to be understood by farmers, and initial implementation costs. In previous decades, high data volumes generated by PA technologies (in-situ sensors, weather station, software, aerial and satellite data) were difficult to store, access and rapidly analyse to provide timely field variability information. Recently, this is averted by adopting big data analytics on cloud computing services and platforms, but the interpretation of the complex information from PA services and their cost vis-à-vis benefits to the farmers may be another critical limitation. New studies on agricultural digitisation footprint are concerned with growing data volumes in precision agriculture, indicating exponential increases over time (Kayad et al., 2022; Marinello et al., 2019). In contrast, the issue of information accuracy poses a significant risk to precision agriculture success and controls farmers' perception of its benefits and, thus, its related technologies.

In-situ sensors such as weather stations and soil moisture sensors provide information about the variability of weather and soil parameters. In contrast, crop parameters can be measured with proximal and remote sensors. Remotely sensed leaf and canopy parameters related to crop growth, health, and yield are particularly appealing. These can be acquired non-destructively and repeatedly over vast areas and in great detail. Moreover, remote sensing systems are versatile and flexible, capable of mounting advanced imaging and non-imaging spectroradiometers deployable on airborne (or aerial systems) and space-bourne (or satellite) systems. Specifically, the spectroradiometers deployed on unmanned spacecraft have evolved rapidly from low spatial-high temporal resolution sensors in the 1970s (e.g., Advanced Very High-Resolution Radiometer, AVHRR) to medium spatial-low temporal resolution sensors in the 1980s (e.g., Landsat Thematic mapper, TM) to more sophisticated sensor constellations providing high spatial-high temporal resolution in the 2020s (e.g., Sentinel-2 and Planet SuperDove).

Critically, the availability of remotely sensed data alone is insufficient, as they are complex and do not readily provide desired information for precision agriculture. Therefore, techniques of varying sophistication are required to extract relevant crop health and growth information for farm-level management and decisions. These techniques have evolved in tandem with the developments in sensor technology and access to data from such new sensors. For example, there has been a shift from the reliance on the visible and near-infrared (VNIR) vegetation indices (VIs) to advanced red-edge indices with the advent of RapidEye and Worldview sensors (Qian et al., 2022; Xie et al., 2018) and more sophisticated physically-based techniques which use Radiative Transfer Models (RTMs). Meanwhile, advances in machine learning regression algorithms (MLRAs) have brought prospects for integrating a variety of remotely sensed, climatic and environmental variables in the retrieval of important biochemical and biophysical parameters of crops such as Leaf Chlorophyll $a + b$ Content (LC_{ab}), Canopy Chlorophyll Content (CCC) and Leaf Area Index (LAI). Coupled with RTMs, studies have shown that MLRAs can be used to develop generic and crop-specific tools for retrieving crop biochemical and biophysical parameters over various crop types, growth stages and climatic conditions (Fernandes et al., 2014; Shah et al., 2019; Yan et al., 2019).

Despite the advances mentioned above in retrieving crop biochemical and biophysical parameters from remotely sensed data, several technical challenges still need to be addressed. These include field measurement errors (such as instrument and sampling errors), poor radiometric quality due to residual errors from atmospheric correction, and parameterisation of RTMs. Unfortunately, the latest developments in sensor technology and MLRAs for biophysical and biochemical parameter retrieval in the context of precision agriculture have not been comprehensively reviewed in the last decade. To this end, we assimilate and review the literature from 2011 to 2023 to showcase recent developments in sensor technology and machine learning algorithms applications for biophysical and biophysical parameter retrieval in agricultural landscapes, as well as the technical challenges and emerging trends with the potentials to support PA. The objectives of this paper were two-fold. First, to provide an overview of recent advances in remotely sensed retrieval of crop biochemical and biophysical parameters brought by the development of sensor technologies and robust machine learning algorithms. Second, to identify the sources of uncertainty in retrieving biophysical and biochemical parameters and implications for precision agriculture. The scope of this review is limited to retrieval of LAI and Chlorophyll Content over herbaceous croplands and grasslands. The literature search was performed using relevant keywords on the Scopus Abstract and Citation Database (www.scopus.com) to select "Published" and "In Press" journal articles.

The remainder of this manuscript is organised as follows: Section 2 provides an overview of the theoretical basis for the remote sensing of biophysical and biochemical parameters, focusing mainly on leaf and canopy parameters related to crop growth and health, i.e., LAI and Chlorophyll content, section 3 provides an overview of the developments in sensor technology (i.e., multi- and hyper-spectral sensors), as well as advances in remote sensing platforms, section 4 provides a review of the advances in the applications of MLRAs for retrieving leaf and canopy parameters of crops, contrasting with traditional techniques such as inversion of RTMs and VIs, section 5 identifies the sources of uncertainty in the retrieval of crop parameters, and sections 6 and 7 summarises the lessons and makes recommendations based on, and discusses main findings of, the literature review.

2. Remote sensing of essential biophysical and biochemical parameters for precision agriculture

Remote sensing of biophysical and biochemical parameters been topical since the dawn of satellite-based remote sensing. Biophysical parameters include Fractional Vegetation Cover (FVC), Fraction of absorbed Photosynthetically Active Radiation (FaPAR), LAI, and biomass; while biochemical parameters include to the canopy water content, chlorophyll content, Nitrogen (N), Phosphorus (P), Pottasium (K), Carteronoids (Car), and proteins. The LAI is inarguably one of the most studied parameters due to its importance in characterizing vegetation, representing the energy exchange between the plant canopies and the atmosphere, and related to the productivity of plants. LAI —measured as m^2/m^{-2}(-|-)— is the ratio of photosynthetically active (PV) and non-PV (NPV) leaf area to ground area. The PV component is referred to as the green LAI (LAI_G), while the NPV component is referred to as the brown LAI (LAI_B). Among the two, LAI_G tends to receive significant attention from researchers of various disciplines and satellite data providers due to its importance for ecological, climate, and crop yield modelling (Zaroug et al., 2013). Consequently, several operational LAI_G, based on low-resolution sensors exist and have been extensively evaluated and validated (Fang et al., 2019, 2012; Sun et al., 2013) and utilised for various crop monitoring (Campos-Taberner et al., 2018), drought assessments (Zhang et al., 2020), and other applications (Rasul et al., 2020). However, only a handful of studies (Amin et al., 2021; Delegido et al., 2015) attempted to characterise LAI_B.

In precision agriculture, characterising the spatio-temporal variability of LAI_G at the field level is critical for monitoring the crop physiological development and phenological status of both erect (i.e., erectophile) and non-erect (i.e., planophile) canopies during the vegetative stages of the crop. Moreover, it can be used to model crop biomass and yield, thus providing essential insights into the productivity of the fields. Despite being limitedly studied, LAI_B is critical for detecting and evaluating the impact of heatwaves and droughts, determining the extent of conservation agriculture (or soil management) practices and modelling fire risks of senescent crops (Amin et al., 2021; Delegido et al., 2015). Moreover, it is an essential decision-support tool for optimising harvest schedules, decisions on planting dates of cover crops, and planning for transportation and storage and fire risk modelling (Hank

et al., 2019). Due to the coverage and cost limitations of destructive lab-based methods, LAI measurements can be acquired with optical field-based instruments such as DEMON, Ceptometer (Accu-Par, Decagon Devices, Pullman, WA, USA), smartphones (Pocket-LAI), digital hemispherical photography (DHP), Pastis57, multi-band-vegetation imager (MVI), TRAC (Tracing Radiation and Architecture of Canopies), and LiCOR Plant Canopy Analyzer (Li-Cor, Inc., Lincoln, NE, USA). These instruments provide an effective Plant Area Index (PAIe) as they cannot distinguish leaves from other plant tissue (e.g., branches and stems). Moreover, errors due to the non-random distribution of canopy foliage, radiation interception from other plant components, and gap fraction saturation as the LAI approaches values of 5 – 6 m^2 m^{-2} (Gower et al., 1999; Weiss et al., 2004), are carried forward to the retrieval methods, and subsequent remotely sensed LAI retrievals. Nevertheless, these instruments are scientifically accepted due to their ease of use, portability, affordability, and consistency.

Leaf chlorophyll *a* and *b* content (LC_{ab}) is a critical crop biochemical parameter for monitoring crop health status, stress and gross primary productivity (Gitelson et al., 2014). Leaf chlorophyll content is significantly correlated to N, i.e., one of the most limiting nutrient in plants, which is critical for crop growth, health, and yield (Gitelson et al., 2003). Chlorophyll, inert structural elements in cell tissue, and the carbon-fixing enzyme, ribulose biphosphate carboxylase (RuBisCo), are all products of N. Direct estimation of N content through lab-based methods is destructive, laborious and costly, thus remotely sensed proxies such as LC_{ab} are crucial for optimising N fertilisation and rapid assessments of crop N status in precision agriculture (Jia et al., 2013; Tian et al., 2013; Vincini et al., 2016). In contrast, Canopy Chlorophyll Content (CCC) —obtained as a product of LAI_G and LC_{ab}— is closely related to LAI (Boegh et al., 2013; Gitelson et al., 2005), for example, with R^2 exceeding 85 % for Maize and Barley (Ciganda et al., 2008). Generally, there is a lack of consensus in the literature regarding the relationship between CCC and N content. For example, Baret et al. (2007) and Delloye et al. (2018) showed that CCC is applicable to canopy N content assessment. Contrarily, Vincini et al. (2016) argued that the factors affecting LAI, such as stand density and water stress, make CCC inadequate for distinguishing N deficiency from other crop stressors. Meanwhile, the C_{ab}-sensitive vegetation indices (VIs) such as the red-edge green Chlorophyll Index (CI_{green}), Chlorophyll Index ($CI_{red-edge}$), and MERIS Chlorophyll Index (MTCI) tend to be the best estimators of LAI, while the LAI-sensitive ones such as NDI45 and NDVI also tend to be the best linear estimators for C_{ab} (Frampton et al., 2013; Viña et al., 2011). This non-exclusive sensitivity to various biophysical and biochemical parameters is mainly due to the interaction of various plant traits, which influences canopy spectral properties, thus making decoupling of such interacting parameters difficult (Delloye et al., 2018; Verrelst et al., 2016). Therefore, there is need for further studies to clarify these inconsistencies.

3. Developments in remote sensing sensor technologies for biophysical and biochemical parameters retrieval

3.1. Advances in remote sensing platforms

The developments in sensor technology are occurring rapidly, thus enabling the extraction of accurate and reliable actionable information promptly. In the 1970s, remote sensing systems consisted of sensors mounted on aerial platforms such as fixed-wing aeroplanes. Today, the low-Earth orbiting satellites and remotely piloted aircraft systems (i.e., RPAS, also commonly referred to as unmanned aerial vehicles, UAVs) are increasingly adopted for characterising field conditions. Inarguably, UAVs are the most powerful remote sensing platforms, with the capability to fly closest to the targets (i.e., at low altitudes), thus, offering super high spatial resolutions (<1 cm) and customisable spectral coverage and flexible revisit times (López-Granados et al., 2016). Recently, UAVs are becoming more accessible and affordable and provide essential tools to acquire data at field scales and at a relatively low cost compared to other platforms. Hence, the availability of UAVs has since driven the miniaturisation of advanced sensor technology. For example, recent studies demonstrate the capability of UAVs for carrying a range of miniaturised sensor payloads and accurately characterising crop water stress variability and status (Zhou et al., 2018), NPK fertiliser deficiency (Corti et al., 2019), weed patch detection (Zisi et al., 2018), crop diseases mapping (Abdulridha et al., 2019), above-ground-biomass of crops (Zheng et al., 2019), and yield estimation (Wan et al., 2020; Zhou et al., 2017). Compared with manned aerial (i.e., aircraft) and space (i.e., satellite) platforms, UAVs are substantially cheaper to operate and maintain, thus holding more possibilities for precision agriculture. Additionally, due to their very low altitude, images collected by UAVs do not require atmospheric correction, which limits the adoption of satellite images in precision agriculture as they introduce a more considerable delay than is necessary to support timely farm management decisions (Atzberger, 2013).

Regrettably, the adoption of UAVs in developing regions such as sub-Saharan Africa is still limited, attributable to a myriad of complex factors, among others, affordability, smaller field sizes —characteristic of extensive small-holder farming (i.e., usually < 0.5 ha)— such that economic benefits of UAVs cannot be realised, lack of awareness, technical and institutional capacities, and poor rural connectivity. Besides, the legislative restrictions (such as a requirement for a piloting license, air service license and RPAS Operator Certificate in South Africa) for flying UAVs instinctively reserve the technology to a few individuals and companies. Moreover, although the cost of UAVs may be low, the cost of the remote sensing payload may be prohibitive. It may be compounded by the repeated imaging required for crop monitoring during the season. Intrinsically, UAVs are spatially limited due to limitations posed by battery capacity; thus, they may only address a few user imaging needs at a time.

Alternatively, satellite platforms —have capabilities to remotely and repeatedly acquire data over large areas, in one pass, at medium (<30 m) to very high (<1 m) spatial resolution— are a convenient compromise and yet still relatively cheaper than manned aerial platforms. Since the launch of the first Earth Resources Satellite (Landsat 1 MSS) in the 1970s, the utilisation of space for collecting Earth observation (EO) data has increased rapidly, with numerous low-Earth orbiting satellites launched. In the last decade, various low (e.g., Sentinel-3), medium-to-high (e.g., Sentinel-2), and very-high-resolution satellites (e.g., Worldview-3) were launched into space. At the same time, ground-based systems are required to acquire calibration and validation data under various field conditions and cover types. Using proximal and hand-held sensors, information on crops' biophysical and biochemical properties, such as LAI, chlorophyll content, and spectral data, can be collected non-invasively and non-destructively. Today, myriad data collected with ground-based, aerial (manned aircraft), low-altitude UAVs, and space-based systems (i.e., satellites) are essential for cross-calibration, training and validation of biophysical and biochemical retrieval models, and product development.

3.2. Advances in sensors

Remotely sensed optical sensors can be classified according to their spatial resolution (i.e., low- to very-high-spatial-resolution sensors), temporal resolutions (i.e., low- to high-temporal-resolution sensors), and spectral or bandwidth configurations (i.e., broadband or multi-spectral sensors, and narrow-band or hyperspectral sensors).

3.2.1. Multispectral sensors

Precision agriculture applications require high spatial (<20 m) and high temporal resolution (<5 days) remotely sensed data to characterise field conditions adequately and accurately throughout the phenological stages of the crops (Atzberger, 2013). Solitarily, the remotely sensed data from low- and medium-resolution multispectral sensors are either

unreasonably coarse to obtain relevant within-field crop condition information or are unreasonably infrequent (exacerbated by cloud cover) to support site-specific farm management decisions meaningfully. This inherent compromise between spatial and temporal resolutions is well-known and evident in many heritage satellite missions, such as the MODerate Resolution Imaging Spectroradiometer (MODIS) and the Landsat programme. A summary of the significant milestones in multi-spectral sensor technology is provided in Fig. 1.

These spatial and temporal compromises presented new research problems and, consequently, the development of data fusion algorithms seeking to leverage the high temporal resolution of low-resolution sensors and the spatial detail of medium- to high-resolution (<20 m) data to generate daily or near-daily medium- or high-resolution synthetic reflectance data which are similar to the measured reflectance observations, with $R^2 > 0.85$ (Wu et al., 2012). The data fusion algorithms —consisting of the spatial and temporal adaptive reflectance fusion models and unmixing-based methods— have also been applied to increase the frequency of the VIs and, thus, biophysical parameters such as LAI (Wu et al., 2015). Using the Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM), Zhou et al. (2020) fused the reflectance imagery of Sentinel-2 and Sentinel-3 to generate high temporal-high spatial resolution reflectance data to the R^2 between 60% and > 90% depending of the date. Moreover, using the Spatial-Temporal Data Fusion Approach (STDFA), Wu et al. (2015) found that the daily LAI values generated from daily synthetic vegetation indices were better than the well-established MODIS LAI product when validated against the daily LAI measurements from the winter wheat field with $R^2 \approx 0.98$ and RMSE ≈ 0.15 m^2/m^{-2}. In another study, Houborg and Houborg & McCabe (2018) developed the Cubesat Enabled Spatio-Temporal Enhancement Method (CESTEM) to generate consistent surface reflectance with Landsat-8 but at 3 m offered by PlanetScope Dove. Therefore, one can argue that the development of data fusion algorithms reduces the shortcomings of both the low- and medium-resolution sensors while also improving their utility for precision agriculture by exploiting good qualities of each, i.e., higher repeat cycles (up to < 1 day) and detailed coverage (<20 m). Since they were comprehensively

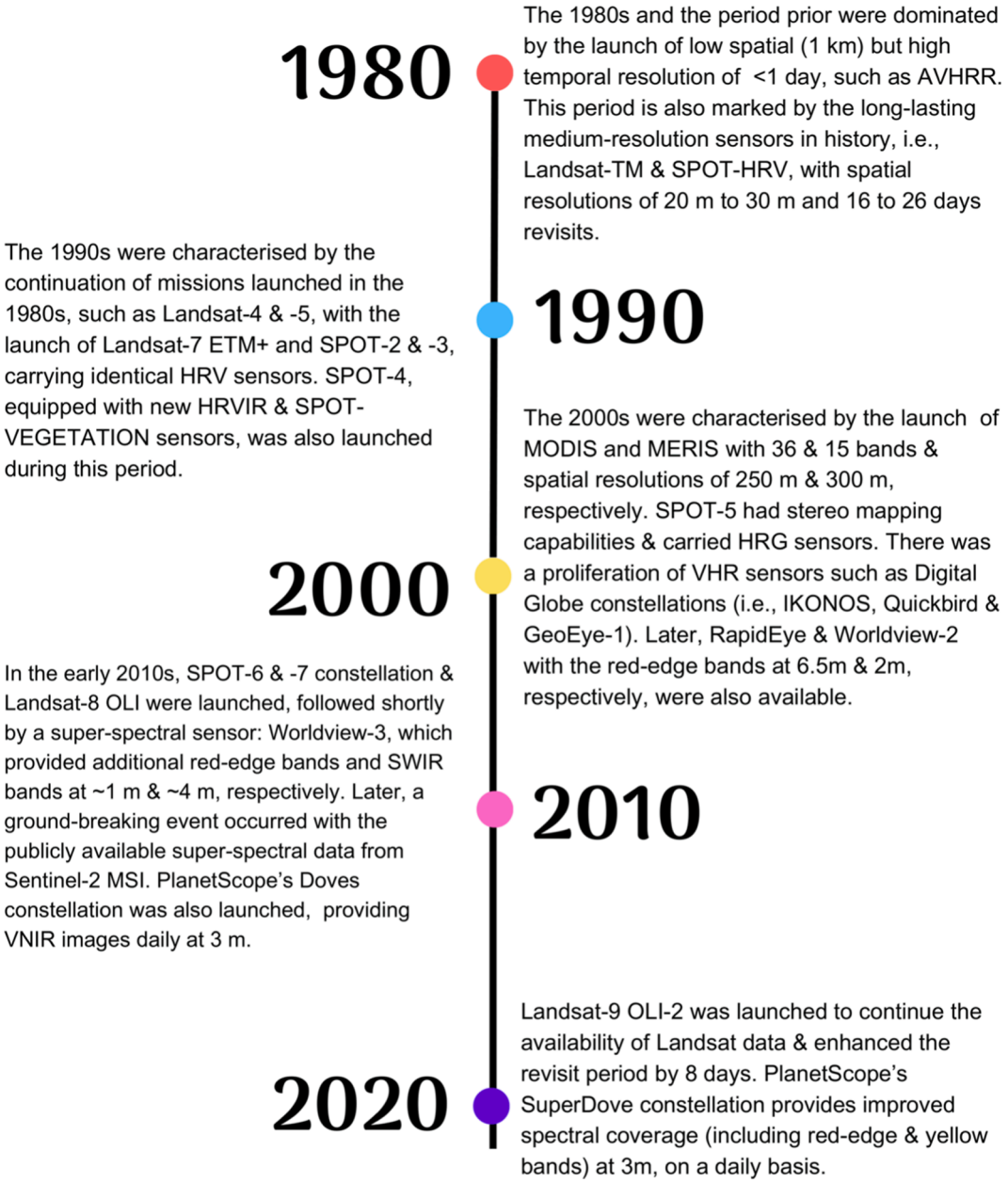


Fig. 1. Significant milestones in the Earth observation satellites with multispectral sensor payloads.

validated, provide long-term time series, and are based on well-established methods, products from low-resolution sensors are invaluable for the quality assessment of new estimation approaches using medium to high-resolution sensors (Kganyago et al., 2020; Zhou et al., 2020). Moreover, data fusion algorithms are also used for harmonising the spatial resolution of sensors offering multiple resolutions, such as MultiSpectral Instrument (MSI) (Zhang et al., 2019) or between different sensors, i.e., Operational Land Imager (OLI), HJ-1 CCD (Charge Coupled Device) or MSI and MODIS or Sentinel-3 Ocean and Land Colour Instrument (OLCI) (Chen et al., 2020; Kimm et al., 2020; Liao et al., 2019; Zhou et al., 2020).

In recent years, the advances in sensor technology have brought about the capacity to obtain high spatial and high temporal resolution multispectral data through off-nadir viewing and tasking capabilities and satellite constellations. However, most of these very-high-resolution (VHR, <5 m) satellite sensors, e.g., IKONOS, Quickbird, GeoEye-1, and Pleiades (High-resolution optical imagers, HiRI), provided only VNIR multispectral data which, in the context of biophysical and biochemical parameter retrieval, are spectrally constrained. Besides, they capture the necessary bands, i.e., red and NIR, for vegetation analysis using parametric techniques such as vegetation indices. Unfortunately, the retrieval of crop biophysical and biochemical parameters using VNIR broadbands has been shown to suffer from saturation due to the chlorophyll absorption in the red band (see detailed discussion in Section 4). Interestingly, some of the new commercial high-resolution and VHR sensor designs, such as SPOT-6 and -7 New Astrosat Optical Modular Imager (NAOMI) and PlanetScope's Dove constellation, retain the VNIR spectral configuration, whilst Copernicus Sentinel-2 constellation prioritised higher spatial resolution (i.e., 10 m) for this spectral region. This can be attributed to the popularity and wide adoption of VNIR indices in various vegetation analyses, including precision agriculture. Therefore, high spatio-temporal characteristics seem to be a priority in providing time-critical site-specific information relevant to precision agriculture. Also, because of their simplicity, vegetation indices such as NDVI are relatable to farmers, and their establishment places them as a benchmark for new indices (Li and Wang, 2013, Zhengyang et al., 2011, Wang et al., 2018, He et al., 2017).

Two recent ground-breaking events in Earth observation satellite sensor technology can be identified, which offer unparalleled capabilities and prospects for precision agriculture. First, the availability of PlanetScope's large constellation of nano-satellite sensors in 2017, consisting of hundreds of CubeSats with 10 cm by 10 cm by 30 cm dimensions and providing imagery at 3 m spatial resolution daily. The imagery from Planet CubeSats has been demonstrated to provide better estimations of biophysical and biochemical parameters using Radiative Transfer Models (RTM) and empirical approaches (Kimm et al., 2020). For example, Kimm et al. (2020) found $R^2 > 75$ % and RMSE of $\sim 1\ m^2\ m^{-2}$ using the Green Wide Dynamic Range Vegetation Index (GrWDRVI) and LUT-based RTM inversion in Illinois, USA. They attributed their results to the CubeSat's finer resolution reflectance data, consistent with the sampling area of in-situ LAI measurements. Although VHR imagery is critical to achieving greater precisions in site-specific management applications, it poses several challenges due to increased spatial variability.

Second, the announcement and subsequent launch of the Sentinel-2 constellation, i.e., Sentinel-2A (in 2016) and −2B (in 2017), carrying identical Multispectral Imager (MSI) cameras. Sentinel-2 MSI does not only guarantee data continuity and interoperability with previous missions such as Landsat, SPOT, and MERIS (Wu et al., 2019) but also provides better spatial resolutions (i.e., up to 10 m), temporal revisits (i.e., 5 days), and the red-edge bands. The red-edge bands —known to enhance the biophysical and biochemical parameter retrieval accuracy using various techniques (Dong et al., 2019; Ramoelo et al., 2012)— were previously only provided by programmable commercial missions such as RapidEye constellation (i.e., since 2008), Gaofen-6 and Worldview series (i.e., since 2009). Although RapidEye and Worldview provided higher spatial resolutions, i.e., 6.5 m and > 2 m, respectively, their prohibitive costs have impeded systematic and operational monitoring of crops using VHR sensors (Houborg et al., 2015). Therefore, extensive archives are only available for big cities worldwide, while agricultural areas have limited coverage, and data availability depends on historical customer orders over specific small areas. For Gaofen-6, its limitation mainly because it is region-specific, and inaccessible to regions out of China. Conversely, Sentinel-2 MSI is provided at no cost to users and provides three red-edge bands, centred at 705 nm, 740 nm and 783 nm; thus present many research and operational advancement prospects. From the precision agriculture perspective, the increased accuracy brought by incorporating red-edge bands is significant for increasing the reliability of retrieved stress-related parameters of crops to aid time-sensitive farm management decisions.

Consequently, several studies have exploited Sentinel-2 red-edge data for LAI, LC_{ab}, and N retrieval over various crops and environments. Before its launch, studies mainly demonstrated its potential using simulated data from leaf and canopy RTMs and hyperspectral sensors. Frampton et al. (2013) used experimental data from SEN3Exp (Barrax, Spain, 2009) and SicilyS2EVAL (Sicily, Italy, 2010) combined with synthetic Sentinel-2 MSI data generated from two airborne imaging spectrometers, i.e., Compact Airborne Spectrographic Imager (CASI) and AISA Eagle, and the PROSAIL model, to the conclusion that red-edge bands improved correlation of the VIs with LAI, LC_{ab} and CCC and averted saturation effects. Using SPARC field campaign (Barrax, Spain, 2003) and Compact High-Resolution Imaging Spectrometry (CHRIS) data, Verrelst et al., (2012b) evaluated the performance of the three MSI spectral configurations, i.e., S2-10m (4 bands), S2-20m (8 bands), and S2-60m (10 bands), with MLRAs and concluded that LAI and FVC could be accurately estimated with S2-10m configuration due to its high spatial resolution, while highest accuracies for LC_{ab} were achieved with red-edge bands. Post-launch studies such as Delloye et al. (2018) mostly confirm the findings of the previous studies, showing the contribution of red-edge bands to the accurate retrieval of LAI, LC_{ab} and CCC, particularly reduced uncertainties of the low and high values and better characterisation across various growth stages and farming practices. Meanwhile, Herrmann et al. (2011) found that Sentinel-2 MSI perform equivalently to a field hyperspectral sensor, i.e., FieldSpec Pro FR spectrometer (Analytical Spectral Devices, USA), in estimating LAI.

The significance of Sentinel-2 and Worldview sensor configurations is also reflected in the missions launched in the 2020s, such as the third generation of PlanetScope's constellation, SuperDove. They represent another revolution in the Earth observation satellite sensor technology for precision agriculture. It offers five to eight spectral bands, which include VNIR bands with two green bands (513 – 549 nm and 547 – 583 nm), yellow (600 – 620 nm) and red-edge bands (697 – 713 nm), thus providing the capability for detailed, accurate, and frequent characterisation of field variability. A summary of sensor characteristics is provided in Table 1.

3.2.2. *Hyperspectral sensors*

Hyperspectral data, characterised by hundreds of narrow (<10 nm) and contiguous spectral bands, are critical for accurately and reliably extracting the crop biophysical and biochemical parameters. Contrary to multispectral broadband sensors, hyperspectral sensors (i.e., imaging and non-imaging spectrometers) can detect minute variations in plant biochemical and biophysical traits that affect the reflected radiance. Because of this capability, hyperspectral sensors are critical for identifying significant spectral bands (or regions) where the various plant traits absorb electromagnetic energy (also called absorption features), simulating new multispectral sensor data (Delegido et al., 2011; Estévez et al., 2020; Verrelst et al., 2013b, 2012b), and detecting fine plant leaf traits and other stress factors (De Castro et al., 2012). These absorption features are then used to design the VIs and used as input variables in machine-learning regression models (Delegido et al., 2014, 2011). Despite their recognised significance for agricultural and natural

Table 1
Characteristics of multispectral sensors relevant for biophysical and biophysical parameter retrieval, classified by spatial resolution. Sensors with pixel sizes of ~ 250 m – 1 km, 20 m to ~ 250 m, <20 m to =>5 m, and <=5 m are classified as low, medium, high and very-high-spatial resolution, respectively.

	Satellite	Sensor	Spectral coverage	Spatial resolution	Revisit period	Availability
Low	SPOT-4 & −5	SPOT-VGT	VNIR	1 km	<1 day	1998 – 2014
	Terra / Aqua	MODIS	VNIR/SWIR	250 m – 1 km	<1 day	2001 – to date
	Envisat	MERIS	VNIR	300 m	<1 day	2002 – 2012
	PROBA	PROBA-V	VNIR	300 m	<1 day	2013 – 2020
	Sentinel-3	OLCI	VNIR	300 m	<1 day	2016 – Present
Medium	Landsat-5	TM	VNIR/SWIR	30 m	16 days	1984 – 2012
	SPOT 1—4	HRV	VNIR	20 m	26 days	1986 –2013
	Landsat-7	ETM+	VNIR/SWIR	30 m	16 days	1999 – 2022
	Landsat-8 & −9	OLI	VNIR/SWIR	30 m	16 days	2014 – Present
	Sentinel-2	MSI	RE/SWIR	20 m	5 days	2016 – Present
High	RapidEye	REIS	VNIR/RE	6.5 m	1 day	2008 – 2020
	SPOT 5	HRG	VNIR/SWIR	10 m	26 days	2002 – 2015
	SPOT 6/7	NAOMI	VNIR	6 m	~3 days	2013 – Present
	Sentinel-2	MSI	VNIR	10 m	5 days	2016 – Present
Very-High	IKONOS	–	VNIR	3.2 m	<3 days	1999 – 2015
	GeoEye	–	VNIR	1.84 m	<3 days	2008 – Present
	QuickBird	BGIS 2000	VNIR	2.62 m	<3 days	2001 – 2015
	Pleiades	HiRI	VNIR	2 m	1 day	2011 – Present
	Worldview-2	WV-110	VNIR/RE/Y	1.84 m	1 day	2009 – 2022
	Worldview-3	–	VNIR/RE/Y/SWIR	1.24 m – 3.7 m	1 day	2014 – Present
	PlanetScope Dove	Dove classic and Dove-R	VNIR	3 m	1 day	2017 – Present
	PlanetScope Dove	SuperDove	VNIR/RE/Y	3 m	1 day	2021 – Present

resources management applications, only a few space-based hyperspectral sensors or imaging spectrometers existed, with most being science demonstrator missions. These include sensors such as Hyperion onboard EO-1, CHRIS onboard PROBA-V, DLR Earth Sensing Imaging Spectrometer (DESIS), EnMap and PRISMA, respectively. The capability of the technology is mainly demonstrated with airborne hyperspectral sensors such as Airborne Visible and Infrared Imaging Spectrometer (AVIRIS), HyMap, and Compact Airborne Spectrographic Imager (CASI), and non-imaging field spectrometers which are seemingly cheaper to operate relative to satellite systems, that also require sophisticated processing procedures and extensive storage.

3.2.3. *Leaf and canopy radiative transfer modelling*

Studies have also used physically-based leaf and canopy RTMs to demonstrate the capability of upcoming optical sensors (Féret et al., 2017), design new spectral indices (Yi et al., 2014), and determine optimal spectral bands for retrieval of biophysical and biochemical parameters (Richter et al., 2012). RTMs simulate the interactions of solar radiation and plant biophysical and biochemical properties based on physically-sound cause-effect relationships. Given a set of leaf and canopy vegetation traits (e.g., LAI, Chlorophyll *a* and *b*, leaf water content, dead matter, and leaf angle distribution), acquisition conditions (e.g., view and illumination angles), soil background and environmental parameters, RTMs can model the full-range (i.e., 400 nm to 2500 nm) spectral reflectance of vegetation at a leaf or canopy scales. In this regard, several RTMs are at our disposal, varying in complexity, the number of parameters, and physical principles. Some examples include Invertible Forest Reflectance Model (INFORM) (Atzberger, 2000), Leaf Incorporating Biochemistry Exhibiting Reflectance and Transmittance and Yields (LIBERTY) (Dawson et al., 1998), FLUSPECT (Vilfan et al., 2018, 2016), and SCOPE (Soil Canopy Observation of Photosynthesis and Energy) (Van Der Tol et al., 2009). Blackburn (2007) and Ustin et al. (2009) previously reviewed some of these RTMs.

Inarguably, PROSPECT + SAIL (commonly called PROSAIL) —a one dimensional coupled leaf-canopy RTM consisting of PROSPECT (Jacquemoud and Baret, 1990) and SAIL (Scattering by Arbitrary Inclined Leaves) (Verhoef, 1984) models— is one the most utilised RTM in vegetation studies (Jacquemoud et al., 2009, 1995). It has been popularised by wide availability, simplicity, fewer input parameters, comparative predictive power to complex RTMs, and its relatively fast computation time. The role of PROSPECT in the PROSAIL RTM is to preliminarily simulate leaf directional-hemispherical spectral reflectance and transmittance before they are incorporated into the SAIL to simulate the Top-of-Canopy reflectance, considering various structural and optical properties of the leaves, canopy, and soil background level for a given acquisition and illumination configuration. Indeed, the two RTMs have evolved independently, predominantly in terms of the complexity of the input parameters, parameterisation (Verrelst et al., 2013a; Wang et al., 2018), coupling with other RTMs (Estévez et al., 2020; Laurent et al., 2014; Verhoef and Bach, 2003), and inversion techniques. Consequently, different version of PROSPECT exist such as PROSPECT-4, -5, -5B, -PRO, and -Dynamic (-D) (Féret et al., 2017; Feret et al., 2008), while the prominent SAIL versions are 4SAIL and 4SAIL2 (Verhoef et al., 2007; Verhoef and Bach, 2007). Generally, PROSPECT requires LC_{ab} content (C_{ab}, [μg cm^{-2}]), leaf dry matter content (C_m, [g cm^{-2}]), leaf water thickness or water content (C_w, [g cm^{-2}]), carotenoid content (C_{cx}), leaf size to crop height (*Sl*), and a leaf mesophyll structural parameter (N, [unitless]). On the other hand, SAIL requires LAI (m/m^{-2}(−|-)), average leaf inclination angle (ALIA, [°]), the fraction of diffuse incoming solar radiation (skyl, usually fixed at 0.1) and the view and illumination geometry, i.e., sun zenith angle (θ_s, [°], sensor view zenith angle (θ_v, [°]), and relative azimuth angle (ϕ_{sv}, [°]) as well as a hot spot parameter (*Sl*, [m/m]). A summary of input parameters used in literature for specific vegetation or crop types is provided in Table 2. We refer interested readers to an exhaustive account of this RTM's applications in vegetation studies by Jacquemoud et al. (2009).

Recent studies (Estévez et al., 2020; Laurent et al., 2014) demontrate the benefit of estimating crop parameters from the Top-of-Atmosphere (TOA) reflectance data by coupling the canopy RTMs with an atmospheric RTMs. By inverting these coupled models, the crop parameters can be estimated directly based on the at-sensor spectral signature. Moreover, various operational and upcoming sensor responses in the spectral domain can be used to resample the simulated spectra by RTMs, providing insights into the prospects of these sensors. For example, before the Sentinel-2 and −3 launch, several studies demonstrated its potential value for biophysical and biophysical parameter mapping using RTM-simulated data (Delegido et al., 2011; Estévez et al., 2020; Verrelst et al., 2013b, 2012b). This included exploring different inversion techniques such as MLRAs, vegetation indices, and LUT-based approaches (Atzberger and Richter, 2012; Clevers and Gitelson, 2013; Richter et al., 2009; Verrelst et al., 2015), which informed post-launch operational techniques.

Table 2
PROSAIL input parameters commonly used in literature to generate Look-Up-Tables (LUTs).

Source	PROSPECT							4SAIL						Vegetation / Crop type
	N	C_{ab}	C_m	C_w	C_{cx}	*SI*	Ver.	LAI	ALIA	Hot	θ_s	θ_v	φ_{sv}	
(Frampton et al., 2013)	1.5	5–70	0.009			–	5	0–8	35	0.01	30	10		
(Yi et al., 2014)	1.2–2.6	60	0.001–0.018	0.01–0.06	15	–	5	0.2–8	–	0.5	35	0		Cotton
(Nigam et al., 2014)	1–3	30–70	0.008–0.025	0.01–0.06		0.1–0.5	–	1–7	–	0.5	– 20–80	0–55	±120	Wheat
(Sehgal et al., 2016)	1.0	20–80	0.0046	0.01–0.04	1.0	–	5B	0.1–6	70,57,45	0.78,0.40,0.32	51,45,33	0	0	Wheat
(Kattenborn et al., 2017)	1.9	10–60	Cw/3.2–4	0.01–0.03	3–15	–	5	0.2–6		0.05	35.5	6.5	98.6	
(Fenghua et al., 2017)	1–4	20–80	0.002–2.0	0.0005–0.04	4–17	–	–	1–5	20–50	0.01–1	±50	±50		Rice
(Atzberger and Richter, 2012)	2.0	20–70	0.004–0.007	0.6–1.4	–	–	–	0.001–8.0	20–70	0.01–1.0	21	8.4	138	Various
(Si et al., 2012)	1.5–1.9	15–55	0.0025–0.005	0.01–0.02	–	–	–	0.1–4	20–70	0.05–0.1	0	0	30	Grassland
(Duan et al., 2014)	1–2	20–70	0.004–0.007	0.005–0.03	–	–	–	0.001–6	30–70	0.05–1				Maize, potato, and sunflower
(Jay et al., 2017)	1–2	20–65	0.002–0.015	0.03–0.09	5–20	–	–	0.1–3.5	10–90	0.33				Sugarbeet
(Liang et al., 2015)	1	10–90	0.002–0.02	0.003–0.05	–	–	5B	0.1–10	30–80	0.05–0.1	30	13.31	138.08	Various
(Li et al., 2017)	1.2–1.8	35–75	0.003–0.011	0.8–0.9	–	–	–	0–7	45–65	0.1–0.5	30–80	0–55	±120	Wheat
(Verrelst et al., 2016)	1.2–2.6	0–80	0.001–0.05	0.001–0.05	–	–	4	0–7	30–60	–	30	–	–	Maize and soybean rotation

Notes for Table 2.
N – Leaf structre paramter; C_{ab} – Leaf chlorophyll concentration; C_m – Dry matter content or leaf mass per area; C_w – Equivalent water content; LAI – Leaf Area Index; ALIA – Average Leaf Inclination Angle; hot – Hot spot size paramter. Skyl parameter - fraction of diffuse incoming solar radiation, is often fixed at 0.1 for all wavelengths (Richter et al., 2012).

4. Machine learning regression algorithms for biophysical and biochemical variables retrieval

Machine learning regression algorithms (MLRAs, also called nonlinear, nonparametric regression algorithms) are robust to the nonlinear functional dependence of the crop biophysical and biochemical parameters and the spectral reflectance data, small training samples, and noise and do not have normality assumptions when compared to linear algorithms such as Ordinary Least Squares Regression (OLS) and Partial Least Squares Regression (PLSR). MLRAs establish relationships with each predictor variable using all the available variations without making assumptions (Verrelst et al., 2012a). Generally, MLRAs are typically classified into three categories, i.e., tree-based (or tree ensembles), kernel-based, and deep learning (or Neural Networks), based on their architectural designs (Rivera-caicedo et al., 2017). The MLRAs were widely exploited for crop biophysical and biochemical parameters retrieval in recent studies (see Table 3) as they are more robust than traditional approaches such as vegetation indices and RTM inversion using numerical optimisation procedures. For example, the NDVI has several limitations, such as saturation with increasing biomass and LAI, sensitivity to soil background and atmospheric contamination. Despite much progress in addressing these problems, such as incorporating background adjustment and atmospheric noise compensation using the blue band (450 – 520 nm) by Liu and Huete (1995), soil-adjustment factor by (Huete, 1988), red-edge bands by (Gitelson and Merzlyak, 1994), recent studies such as Verrelst et al. (2015) show that vegetation indices are sensitive to specific ranges of crop parameters below or beyond which they perform poorly, regardless of their mathematical formulation and spectral bands. Feng et al. (2019) found that vegetation indices were sensitive to various cultivars, irrigation levels, plant densities, N rates and years.

Tree-based MLRAs are less complicated, intuitive, and popular than kernel-based and deep learning algorithms. This family of algorithms have origins in the Classification and Regression Trees (CART), which is rarely used in recent years due to the advent of more robust and updated variants such as the Random Forest (RF) and Gradient Boosted Regression Trees (GBRT), also known as Gradient Boosting Machines (GBM) (Friedman, 2001). Recently, Chen and Guestrin (2016) proposed a more advanced, scalable and sparsity-aware improvement on GBM, namely, eXtreme Gradient Boosting (XGBoost). The XGBoost algorithm aims to improve the implementation of GBM by (1) handling missing values more efficiently, (2) constructing trees quickly and building huge models utilising parallel computing, and (3) utilising highly regularised formalisation and gradient-boosted trees, thus avoiding over-fitting. Therefore, XGBoost often outperforms other algorithms (Beltran et al., 2019).

However, RF is still prominently used in literature due to good performance and relatively few required hyperparameters, with studies showing its robustness and high accuracy in predicting various crop biophysical and biochemical parameters. For example, LI et al. (2017a) found an R^2 of 88 % and an RMSE of 0.195 m^2 m^{-2} in retrieving grassland LAI Landsat TM + and OLI. Another study (Tavakoli and Gebbers, 2019) found R^2 of 56 % to 69 % in N content of Wheat in Germany using images acquired from a digital camera. Overall, tree-based algorithms are appealing due to their interpretability and comprehensibility. For example, these algorithms allow interrogation of the tree structure and variables used at each split and provide the importance (or relative influence) values for each explanatory variable. Therefore, they can be used to obtain novel insights, such as new functional relationships between the response and explanatory variables. Others (Shelestov et al., 2017) have used such variable importance measures for feature selection. The capability to interrogate the model structure is critical for troubleshooting the models to obtain robust crop parameter retrievals from satellite images (Azodi et al., 2020).

On the other hand, kernel-based and deep-learning MLRAs are

Table 3
Commonly used MLRAs for biophysical and biochemical retrieval with associated accuracy measures, i.e., Coefficient of determination, R^2, and Root Mean Squared Error, RMSE, where possible. RF, XGBoost, SVM, GPR, KRR, and ANN denote Random Forest, eXtreme Gradient Boosting, Gaussian Process Regression, Kernel Ridge Regression, and Artificial Neural Networks (ANN).

	Author(s)	MLRA	Accuracy (R^2; RMSE)	Targetparameter (s)	Crop type (s)	Sensor	Explanatory Variables	Experimental site(s)
Tree-based	(Ramoelo et al., 2015)	RF	R^2: 0.71 – 0.89; RMSE: 0.04 – 0.0.8	N	Grassland	Worldview-2	Spectral bands and Vegetation indices	Northern South Africa
	(LI et al., 2017a)	RF	R^2: 0.88; RMSE: 0.195 m^2 m^{-2}	LAI	Grassland	Landsat7 ETM+; Landsat8 OLI	Spectral bands and Vegetation indices	Hulunber, Inner Mongolia, China (49°2024N, 119°5944E)
	(Tavakoli and Gebbers, 2019)	RF	R^2: 0.56 – 0.69; RMSE: 0.27 – 0.19 %*	N	Wheat	Digital camera	Spectral bands and Vegetation indices	Marquardt experimental station, Potsdam, Germany (52°27′N, 12°57′E)
Kernel-based	(Malenovský et al., 2017)	SVM	R^2: 0.52 – 0.54; RMSE: 242.6 – 238.3 nmol g^{-1} dry weight	LC_{ab}	Antarctic mosses	Hyperspectral UAS; WorldView-2	Spectral bands	Antarctic Specially Protected Area 135 (66.282°S, 110.539°E) and Robinson Ridge (66.368°S, 110.586°E)
	(Verrelst et al., 2012b)	GPR	R^2: 0.94 – 0.96; RMSE: 0.47 – 0.55 m^2 m^{-2}	LAI	Multi-crops#	Sentinel-2 MSI*; Sentinel-3 OLCI*	Spectral bands	Barrax, La Mancha region, Spain (30°3′ N, 2°6′ W)
	(Verrelst et al., 2012b)	GPR	R^2: 0.94 – 0.99; RMSE: 1.81 – 5.36 μg cm^{-2} ¥	LC_{ab}	Multi-crops#	Sentinel-2*; Sentinel-3*	Spectral bands	Barrax, La Mancha region, Spain (30°3′ N, 2°6′ W)
	(Campos-Taberner et al., 2016)	GPR	R^2: 0.88 – 0.89; RMSE: 0.78 m^2 $m^{-2\varphi}$	LAI	Rice	Landsat-8 OLIϕ; SPOT-5ϕ	Spectral bands	Valencia, Spain; and Lomellina, rice district, Lombardy, Italy
	(Campos-Taberner et al., 2016)	KRR	R^2: 0.82 – 0.83; RMSE: 0.94 – 0.97 m^2 $m^{-2\varphi}$	LAI	Rice	Landsat-8 OLIϕ; SPOT-5ϕ	Spectral bands	Valencia, Spain; and Lomellina, rice district, Lombardy, Italy
	(Elarab et al., 2015)	RVM	RMSE: 5.31 μg cm^{-2}	LC_{ab}	Oats	VNIR & Thermal UAS	Spectral bands and Vegetation indices	Utah, USA (39o14′ N,112o6′ W)
	(Verrelst et al., 2016)	GPR	R^2: 0.79; RMSE: 72.36 mg m^{-2}	LC_{ab}	Multi-crops#	Field spectrometer†	Spectral bands	Barrax, La Mancha region, Spain (30°3′ N, 2°6′ W)
	(Verrelst et al., 2016)	GPR	R^2: 0.94; RMSE: 0.40 m^2 m^{-2}	LAI	Multi-crops#	Field spectrometer†	Spectral bands	Barrax, La Mancha region, Spain (30°3′ N, 2°6′ W)
	(Verrelst et al., 2016)	GPR	R^2: 0.95; RMSE: 0.37 m^2 m^{-2}	LAI	Multi-crops#	HyMap†	Spectral bands	Barrax, La Mancha region, Spain (30°3′ N, 2°6′ W)
	(Wen et al., 2018)	GPR	R^2: 0.85; RMSE: 0.95 μg ml^{-1}	N	Rice	Hyperspectral UAS	Spectral bands	ShenyangAgricultural University, Liaoning Province, China (41°81′63″N, 123°55′85″E)
Deep learning	(Verger et al., 2011)	ANN	R^2: -; RMSE: 0.37 m^2 m^{-2}	LAI	Multi-crops#	CHRIS/PROBA	Spectral bands	Barrax, Castilla-La Mancha region, Spain (30°3′ N, 2°6′ W)
	(Campos-Taberner et al., 2016)	ANN	R^2: 0.83 – 0.84; RMSE: 0.91 – 0.93 m^2 m^{-2} φ	LAI	Rice	Landsat-8 OLIϕ; SPOT-5ϕ	Spectral bands	Valencia, Spain; and Lomellina, rice district, Lombardy, Italy
	(Delloye et al., 2018)	ANN	R^2: 0.55 – 0.85; RMSE: 1.00 – 0.70 m^2 m^{-2} ‡	LAI	Wheat	Sentinel-2 MSI; SPOT-5	Spectral bands	Belgium
	(Delloye et al., 2018)	ANN	R^2: 0.08 – 0.31; RMSE: 13.94 – 11.03 μg cm^{-2} ‡	LC_{ab}	Wheat	Sentinel-2 MSI; SPOT-5	Spectral bands	Belgium
	(Delloye et al., 2018)	ANN	R^2: 0.46 – 0.62; RMSE: 0.51 – 0.35 g m^{-2} ‡	CCC	Wheat	Sentinel-2 MSI; SPOT-5	Spectral bands	Belgium
	(Dhakar et al., 2019)	ANN	R^2: 0.56 – 0.75; RMSE: 1.34 – 0.94 m^2 m^{-2} £	LAI	Wheat	Sentinel-2 MSI	Spectral bands	Pataudi block of Gurugram district, Haryana, India

Notes for Table 3.
¥ Accuracy ranges are for various configurations tested in the study, i.e., S2-10 m (4 bands with 10 m spatial resolution), S2-20 m (8 bands, i.e., B2 to B8a, at 20 m resolution), S2-60 m (10 bands, i.e., B1 to B9, at 60 m resolution), and S3-300 m (19 bands, i.e., BO1 to BO20, at 300 m resolution).
* Simulated from Compact High-Resolution Imaging Spectrometry (CHRIS) data with a spectral range of 400 to 1050 nm.
# Sunflower, Maize, Alfalfa, Wheat, Sugar beet, Onion, Garlic, Potato, and Vineyard.
ϕ Simulated from PROSAIL Radiative Transfer Model.
φ Accuracies are for simulated Landsat OLI and SPOT-5 data, respectively.
† Original data were reduced to fewer bands using Gaussian Process Regression-based band analysis tool (GPR-BAT).
‡ Accuracies were achieved with simulated Sentinel-2 data divided into different band subsets, i.e., SPOT-5 (4), 10-bands (3), Red-edge (7) and All bands (9).
£ Accuracies were achieved with inversion of PROSAIL LUT with ANN using Sentinel-2 MSI corrected with MODTRAN and libRadtran atmospheric correction approaches.
* Accuracies are for estimated N over various phenologies, i.e., 219 and 234 days after sowing (DAS) and combined dates.

complex, computationally expensive, and opaque (or 'black box'). One also has to parameterise several confounding hyperparameters, and there is no capability to compute variable importance measures directly and interrogate the models' inner workings. Nonetheless, these MLRAs have attracted significant attention in recent studies using simulated and actual data from various hyper- and multi-spectral sensors (see Table 3). Support Vector Machine (SVM) is perhaps the most commonly used kernel-based algorithm in remote sensing applications and supports several kernels, with popular ones being Gaussian Radial Basis Function (RBF), Sigmoid, polynomial and linear kernels (Mountrakis et al., 2011; Yuan et al., 2017). However, the Gaussian Regression Process (GPR) and its variants, such as Variational Heteroscedastic Gaussian Processes, are popular in crop biophysical and biochemical parameters retrieval studies, offering better prospects due to its high accuracy and a special capability to generate estimates of the response variable's uncertainty which enable evaluations of the crop parameters retrievals' reliability for use in operational applications (Verrelst et al., 2013b). Like tree-based algorithms, GPR provides insights into the relevant bands which indicate the relationship with crop biophysical and biochemical parameters. GPR also supports multiple kernels, the common being anisotropic squared exponential and scaled Gaussian kernels. The algorithm requires fewer hyper-parameters, i.e., $\theta = \{v, \sigma_b, \sigma_n\}$, where (v, σ_b) refers to the signal —composed of the scaling factor v and the length-scale of the explanatory variables σ_b— and σ_n refers to the standard deviation of the estimated noise. Verrelst et al., (2012b) compared the capabilities of deep learning, i.e., ANN, and three kernel-based MLRAs, i.e., SVM, Kernel Ridge Regression (KRR) and GPR, in the retrieval of LAI, LC_{ab}, and Fractional Vegetation Cover (FVC) using simulated Sentinel-2 and Sentinel-3 data from CHRIS/PROBA data. Their study found that GPR had a relatively higher processing speed (i.e., < 2 s) and produced the most accurate results for all the biophysical and biochemical parameters considered, i.e., R^2 of 0.89 – 0.99, with LC_{ab} being the most accurate and within Global Monitoring for Environmental and Security (GMES) limit of 10 % accuracy. Others (Wen et al., 2018) found an accuracy of 85 % and RMSE of 0.95 μg ml in estimating N in Rice using Hyperspectral data collected from an Unmanned Aerial System (UAS).

In contrast, Artificial Neural Network (ANN) is the most commonly used deep learning algorithm. As its name suggests, the algorithm is inspired by neurological sciences. It consists of layered and interconnected structures of artificial neurons by weights or links. It consists of various hyperparameters, such as the number of hidden layers, weights, nodes per layer, learning rate, the shape of the nonlinearity, and regularisation parameters. The ANN structure is often optimised using a learning algorithm, such as computationally-heavy Levenberg-Marquardt or a quicker optimisation algorithm, such as gradient back-propagation. Recent developments such as Recurrent Neural Networks, Convolutional Neural Networks, and Long Short-Term Memory present new prospects for improving crop parameter retrieval accuracy. For example, Albughdadi et al. (2021) found that a 2-D convolutional network, i.e., UNet, performed better than the SNAP Biophysical processor and the multilayer perceptron regressor and was computationally efficient. These are less prone to error than other machine learning techniques and have been shown to perform optimally in various environments for various applications, e.g., yield prediction (Barbosa et al., 2020), LAI (Apolo-Apolo et al., 2020), chlorophyll content (Xiaoyan et al., 2020), leaf water content (Nasir et al., 2019), crop type mapping and plant disease detection (Golhani et al., 2018).

Besides regression problems, MLRAs can also be applied to feature selection or dimensionality reduction, thus critical for reducing the dimensionality ($n < p$) and collinearity of predictors. Although some MLRA, such as SVMs, has been dubbed as robust to dimensionality, recent studies increasingly show that dimensionality reduction techniques help reduce uncertainties in crop parameters retrieval. For example, Verger et al. (2011) found that feature selection benefited the ANN performance, with additional bands causing the noise. The high dimensionality, characteristic of hyperspectral sensors and RTM-generated LUTs (often 1 nm spaced spectral bands), causes collinearity; thus, can result in over-fitting and a biased outcome (Wen et al., 2018). In such cases, although the retrieval model performs well in one area, the retrieval accuracy plummets when transfered to other environmental scenarios. Hence, the optimal number of spectral bands for LAI retrieval was 7 out of the 62 bands from CHRIS/PROBA. While Shelestov (2017) found that the same number of features were significant for LAI, FaPAR, and FCover retrieval using Landsat and SPOT-5 images collected as part of the SPOT-5 Take Five initiative. Others (Delegido et al., 2011; Verrelst et al., 2016) have shown that four to nine bands, selected through variable selection techniques, are sufficient to achieve robust retrievals of biophysical and biochemical parameters. The variable selection approaches are generally divided into filter-based (e.g., analysis of variance), wrapper-based (e.g., Recursive Feature Elimination-Support Vector Machine), and embedded (e.g., sparse Partial Least Squares) algorithms. In addition to feature selection, multivariate dimensionality reduction techniques such as Principal Component Analysis (PCA) and Partial Least Squares (PLS) have been used effectively. Rivera-caicedo et al. (2017) showed that such techniques improved the LAI retrievals when coupled with NN, KRR, and GPR, achieving R^2_{cv} of 93 % instead of all HyMap bands. Considering that the multispectral data from new generation sensors such as Sentinel-2 have increased the number of bands, i.e., more than 7, it remains questionable whether feature selection with these datasets can improve the accuracy and reduce uncertainties of retrieving crop biophysical and biochemical parameters. Generally, optimising multispectral features for specific crop parameters is limitedly studied.

Various MLRAs have different limitations, as identified in Table 4. These relate to their ability to handle different datasets, computational complexity and efficiency, transferability (i.e., portable to new areas and periods), and transparency and explainability, i.e., the algorithm's inner mechanism can be interrogated to understand their functioning in different scenarios.

5. Sources of uncertainty in the retrieval of crop biophysical and biochemical parameters and implications for precision agriculture

Uncertainties in remotely sensed biophysical and biochemical parameters are a significant concern for agronomic applications. These uncertainties emanate from several factors, such as in-situ measurement errors (including instrument and sampling errors), acquisition conditions (including atmospheric conditions and sensor and sun geometries), and parameterisation of RTMs and retrieval techniques such as machine learning algorithms. In-situ measurements are a critical part of any empirical analysis, including crop parameter retrieval using RTMs, VIs and MLRAs; therefore, they must be reasonably accurate and collected using appropriate sampling strategies. Otherwise, the quality of these measurements may adversely impact the accuracy and reliability of the retrieved crop parameters. The existence of various field instrumentation with various estimation methods is one of the factors causing measurements to be incomparable and unreproducible, particularly when there is no cross-calibration of such instruments to reduce systematic errors. For example, in-situ measurements from the Minolta SPAD-502 chlorophyll meter (i.e., commonly used for LC_{ab} measurements) and MC-100 Chlorophyll Concentration are different. While MC-100 measures the absolute chlorophyll values, the SPAD-502 chlorophyll meter measures an index related to LC_{ab}. Therefore, it requires site-specific data, i.e., lab-based chlorophyll content, for calibration over many crop types and leaf structures. Such a requirement undermines the advantages of using remotely sensed satellite data, as the instrument provides unstable and inconsistent in-situ measurements. Uddling et al. (2007) showed that SPAD values saturate at $LC_{ab} > 40$ μg cm^{-2} and differ according to growth stages, species, and distribution. Therefore, determining empirical calibration equations for every situation may be difficult, labour-intensive, and expensive. Other studies (Elarab et al.,

Table 4
Pros and cons of different Machine Learning Regression Algorithms (MLRAs), reported in the literature.

	Pros	Cons
Tree-based	• Intuitive, RF requires few parameters for parameterisation, XGBoost and RF are computationally inexpensive, RF does not overfit[4] RF is computationally efficient[4]	• RF requires a large training set for better performance[1] RF suffer from high dimensionality and collinearity,
Kernel-based	• GPR provides uncertainty/confidence intervals estimates[3], GPR has good interpolation capabilities, KRR and SVM can handle small training data[1], GPR and KRR are computationally efficient[3] GPR is transparent, provides relevant samples and bands to model accuracy[3],	• KRR and SVM are computationally expensive when training data is large[1] SVM is computationally expensive[3]
Deep Learning	• ANN is transferable, ANN is computationally efficient in application mode[2, 3]	• Computationally expensive during training[5], Opaque "Blackbox" models [3], ANN is unpredictable when presented with unseen spectra[3], ANN is complex ANN is sensitive to dimensionality[2] ANN requires sufficiently large training sets[1, 2]

Note for [1](Carter and Liang, 2019), [2] (Verger et al., 2011), [3] (Verrelst et al., 2012c), [4] (Li et al., 2015), [5] (Rivera-Caicedo et al., 2017).

2015; Kganyago et al., 2022) utilised the published calibration equations to determine in-situ LC_{ab} values, which may introduce uncertainties due to variations in crop types, growth stages, and prevailing environmental conditions in various study areas. As a result, Kganyago et al. (2022) suggest instrument cross-calibration and adjustment of systematic errors in the field measurements taken by various instruments. Moreover, uncertainties linked to optical instruments such as LAI-2000 (i.e., commonly used for LAI measurements) are inevitable and well-known. For example, Rautiainen et al. (2012) reported that LAI-2000 has 10 % – 20 % uncertainties varying by species and environment. For precision agriculture applications, uncertainties related to in-situ instruments will likely be transferred to the estimated crop biophysical and biochemical parameters. It is, therefore, critical that they are reported adequately as part of the metadata to ensure that the user of the information generated from such data knows its limitations.

Satellite and aerial images are prone to contamination by varying atmospheric constituents between the acquisition dates and residual errors after atmospheric correction. Atmospheric contamination and the inability of atmospheric correction (AC) tools to completely remove it also causes increased uncertainties in the accuracy of the crop parameters estimates. Changing atmospheric conditions during image acquisition has been a challenge since the dawn of satellite-based remote sensing. Aerosols (smoke, dust, and atmospheric gases such as CO_2) increase the atmospheric backscattering signal while attenuating the surface directional reflectance signal (Hilker et al., 2009), undermining the radiometric integrity and consistency of the measured reflectance. Consistent radiometric measurements are critical for phenological analysis and reliable estimates of crop parameters. Wilson et al. (2014) showed that aerosol optical thickness (AOT) may cause errors of about 1.7 % in the measured reflectance and a 5 % change in the Normalised Difference Vegetation Index (NDVI). Ohde (2013) determined the sensitivities of chlorophyll-a concentration estimation methods to atmospheric dust and clouds. Their results showed that while dusty skies caused an overestimation of $>$ 8 % in chlorophyll-a concentration, atmospheric conditions consisting of mixtures of clouds and dust resulted in overestimations of 7 % – 14 %. Alarmingly, as an integral input to precision agriculture, such errors manifest in remotely sensed crop parameters and may lead to false detection of stress and other crop conditions. Consequently, it is critical to utilise well-validated AC tools (de Keukelaere et al., 2018; Doxani et al., 2018; Sola et al., 2018) and standardised analysis-ready-data processed according to minimum requirements CEOS (Committee on Earth Observation Satellites) Analysis Ready Data for Land (CARD4L). A limitation here could be the delayed implementation of relevant standards by satellite data providers, which may limit the radiometric consistency between sensors and, thus, the realisation of the full information potential of satellite data for precision agriculture applications (Giuliani et al., 2017). Moreover, the delay introduced by AC, renders subsequent information extraction steps (such as biophysical and biochemical parameter retrieval) obsolete as it comes with a larger delay than required to support timely farm management decisions (Atzberger, 2013). In such cases, retrieval models which utilise the Top-of-Atmosphere (TOA) data directly (Estévez et al., 2022, 2020), are convenient. Besides the signal attenuation and AC residual errors, satellite radiometric measurements induce non-negligible sensor and sun geometry effects, explaining $>$ 30 % in reflectance variability (Kganyago et al., 2023).

Another important consideration is the limitations of simulations generated by Radiative Transfer Models (RTMs). Despite successfully implemented operational workflows using MLRAs to invert crop parameters (Baret and Weiss, 2018; Weiss and Baret, 2016), RTMs are ill-posed, i.e., the different combinations of canopy parameters can simulate similar canopy reflectance due to mutually compensating effects (Houborg et al., 2015). Moreover, they cannot represent complex canopies of various crops due to their simplified modelling assumptions (Darvishzadeh et al., 2008; Dorigo et al., 2012). Hence, contrary to the widely accepted wisdom that coupled RTM-MLRA models are universal, recent studies found poor performances in environments other than where they were trained (Fernandes et al., 2014; Kganyago et al., 2020; Xie et al., 2019). In validating the SNAP Biophysical processor, based on PROSAIL-generated LUTs and Neural Networks, Kganyago et al. (2020) found LAI uncertainties of $>$ 2 m^2/m^{-2}(−|-) using Sentinel-2 data and concluded that the method was not suitable for precision agriculture. In another study, Xie et al. (2019) found a similar result in China over winter wheat, where the SNAP biophysical processor obtained LAI uncertainties of 1.53 m^2/m^{-2}(−|-) ($R^2 > 0.5$), whilst the uncertainty for CCC was 148.58 μg cm^2. To deal with ill-posedness, several regularisation strategies are used in literature, including incorporating measured leaf and canopy parameter value ranges (Xie et al., 2019), imposing spatial constraints on the RTM model (Atzberger and Richter, 2012), incorporating land cover information (Verrelst et al., 2012d) and using multiple best solutions instead of one in LUT-based inversions (Verrelst et al., 2013a). Duan et al. (2014) found low uncertainties of 0.62 m^2 m^{-2} in estimating LAI using the PROSAIL model and UAV hyperspectral data in Inner Mongolia, China, by approximating LAI and Average Leaf Angle (ALA) parameter ranges from the local in-situ measurements. Therefore, locally calibrated and regularised RTMs combined with other MLRAs may reduce uncertainties and improve the accuracy of biophysical and biochemical parameters, thus making them fit-for-purpose for precision agriculture.

Other uncertainties may be introduced by several confounding hyperparameters required in calibrating various MLRAs. The sensitivity of MLRAs to these hyperparameters in the context of crop biophysical

and biochemical parameters retrieval has not yet been thoroughly studied. However, studies applying MLRAs often employ hyperparameter tuning strategies which evaluate all combinations of the required hyperparameters or those selected randomly. Then, hyperparameters that result in the lowest uncertainties (e.g., RMSE) are chosen to retrieve biophysical and biochemical parameters. It should be noted that these hyperparameters are site-, crop condition- (i.e., phenology), and species-specific and a slight deviation in site factors (i.e., crop, climatic and environmental conditions) may result in higher uncertainties in the retrieved parameters. This is especially important to consider when MLRA models are transferred to new areas (or sites) where the crop types and conditions, climatic (i.e., temperatures, relative humidity, and rainfall), and environmental (i.e., soil types, soil moisture, and topography) factors may be different from the original site where the model was trained. Therefore, a sound transferability framework, which considers the variability of crop types, physiological conditions, climate, and environmental factors, is needed to improve the transferability of MLRAs and reduce the need for extensive field data collection.

Finally, although the availability of sub-meter spatial resolution UAV imagery offers many prospects for precision agriculture, it may introduce some uncertainty to crop parameter estimation using various retrieval techniques, including MLRA. This uncertainty emanates from high spatial variability caused by detecting inter-row spaces and background features such as soils, weeds, litter, and shadows. These might confuse the retrieval algorithms, especially if their spectral signatures were not considered during model calibration, causing false alarms. Therefore, masking such background features is an important preprocessing step. Moreover, object-based or parcel-based approaches may be more appropriate than pixel-based approaches to deal with the 'too much' detail (i.e., <10 cm) characteristic of UAV data (Yang et al., 2022). Acquiring data at ultra-high spatial resolution also presents image registration problems which may cause resource wastage. For example, image misregistrations will result in mislocated prescription maps (Gómez-Candón et al., 2014), causing the Variable Rate Application (VRA) systems to spray fertilisers, herbicides, and pesticides in areas that are not necessarily needed. This is because the accuracy available from most GPS (Global Positioning Systems) is coarser than the pixel size provided by ultra-high-resolution systems. To deal with this problem, Hunt et al. (2018) suggested that each image should be analysed as a separate plot for monitoring. Besides, these images also present opportunities for discrimination of crop species in inter-cropping cultural systems, hence the capability to estimate crop parameters per species in such systems by accounting for different canopy configurations and leaf spectral properties. Overall, uncertainties from various sources imply that the retrieval of crop biophysical parameters may lack the required accuracy, needed for effective site-specific crop monitoring and management. Interested readers can find other important but broader technical considerations in the literature (Ali and Imran, 2021; Malenovský et al., 2009).

6. Lessons and recommendations for future studies

Several lessons could be derived from the literature review, as well as the opportunities for future studies. Concerning the relevant crop growth and health parameters, this paper established that most studies rarely focused on the estimation of LAI_B, but rather on LAI_G. However, the current retrieval methods for LAI_G cannot be directly used to retrieve LAI_B. Hence, there is an opportunity for future studies to explore techniques for its retrieval and operationalisation. It is an essential decision-support tool for assessing the impact of heat waves and, at the end of the season, for optimising harvest schedules, covering crop planting dates, and planning transportation and storage logistics. Moreover, there is a lack of consensus in the literature about the relationship between CCC and N content. Some authors, such as Baret et al. (2007), showed that CCC could be used in canopy N content, and others, such as Vincini et al. (2016), argue that it is inadequate for distinguishing N deficiency from other crop stressors. To demystify these inconsistencies, further studies are needed in different environments.

Inarguably, UAVs are the most powerful remote sensing systems, offering many advantages such as super high spatial resolutions (<1 cm) and customisable spectral coverage and flexible revisit times (López-Granados et al., 2016). However, their lack of adoption in developing regions such as sub-Saharan Africa is related to a myriad of complex factors, such as smaller field sizes (i.e., usually < 0.5 ha) such that economic benefits of UAVs cannot be realised, lack of awareness, technical and institutional capacities, and poor rural connectivity. Therefore, UAV data can be used with satellite data to enhance operational models for crop parameter retrieval at many resolutions, making such models portable to data-scarce regions. In the past, the compromise between the spatial and temporal resolutions in remote sensing sensor designs reduced the utility of the datasets for precision agriculture. Recent literature demonstrates the prospects of data fusion algorithms for reducing the shortcomings of both the low- and medium/high-resolution sensors and improving their utility for precision agriculture by exploiting good qualities of each, i.e., higher repeat cycles (up to < 1 day) and detailed coverage (<20 m). Moreover, this paper found that new high- and very-high-resolution sensors (e.g., PlanetScope's Dove constellation) maintain the broadband VNIR sensor configurations despite the well-established saturation of crop parameters retrieved within this region of the electromagnetic spectrum. This is attributed to their lower cost and the fact that VNIR indices are still popular in various vegetation analyses and easily understood by users. Moreover, high spatial and temporal resolutions seem to be the most attractive factor for operational site-specific management than the spectral bands they provide.

Two ground-breaking events which revolutionise the Earth observation satellite sensor technology for precision agriculture were identified. First, the free availability of Sentinel-2 constellation, i.e., Sentinel-2A (in 2016) and -2B (in 2017), carrying identical MSI cameras. Sentinel-2 constellation guarantees data continuity of, and interoperability with, the previous missions and provides higher spatial resolutions, 5 days revisits, and red-edge bands previously offered by commercial missions only. The new red-edge bands, centred at 705 nm, 740 nm, and 783 nm, triggered new research and presented prospects for operational and accurate monitoring of stress-related agricultural crop parameters to aid time-sensitive agricultural decisions and improved yields. Second, the CubeSat constellations, such as PlanetScope's SuperDove, will likely reduce the VHR data cost, thus increasing accessibility to researchers, product developers, governments, and farmers. In addition to standard VNIR bands, it provides two green bands (513 – 549 nm and 547 – 583 nm), yellow (600 – 620 nm) and red-edge bands (697 – 713 nm) daily and a 3 m spatial resolution. This provides the capability and opportunities for detailed, accurate, and frequent characterisation of field variability, thus increasing operational outcomes.

Machine Learning Regression Algorithms (MLRAs) have evolved significantly. However, this review found that RF is still prominently used in literature, with studies citing good performance and fewer required hyperparameters as the main attractive aspects. In general, tree-based algorithms are also appealing because they are interpretable and explainable, i.e., they allow interrogation of the tree structure and variables used and indicate influential variables. In contrast, kernel-based and deep learning algorithms are considered complex, computationally expensive, and opaque (or 'black box'). Some limitations are a requirement to tune many confounding hyperparameters, and there is no capability to directly compute variable importance and interrogate the models' inner workings. Therefore, the effect of each hyperparameter on the accuracy of these MLRAs has not been studied yet. Nonetheless, these MLRAs have attracted significant attention in recent studies using simulated and various data from hyper- and multi-spectral sensors and higher accuracies were reported. The most considerable

interest was in GPR, and its variants are evident, where most studies report relatively high accuracies and a unique capability to provide response variable uncertainty estimates that enable assessment of the reliability of the crop parameter retrievals for agronomic applications. Like tree-based algorithms, GPR provides insights into the relevant bands, which indicate the relationship with the response variables. It has been found to perform better than Artificial Neural Networks (ANN), which is somewhat computationally expensive and complex. Further studies are needed to stabilise MLRA performance across many crop types and conditions and climatic and environmental conditions by incorporating coupled RTM parametrised with crop- and site-specific in-situ data. In the future, studies should conduct a bibliometric analysis to identify which retrieval techniques and hyper-parameter values are frequently used.

Additionally, the review found that coupling MLRAs with multivariate dimensionality reduction enhances the accuracy of crop parameter retrievals even when using powerful algorithms such as ANN, KRR, and GPR (Yang et al., 2012). Considering that the data from quasi-hyperspectral sensors such as Sentinel-2, Worldview-3 and SuperDove have many bands, it is not yet known whether feature selection with these datasets can improve the accuracy and reduce uncertainties of retrieving crop growth and health parameters. Generally, the optimisation of multispectral features for specific crop parameters is limitedly studied. Therefore, future studies should address this gap, especially considering the numerous variables, such as spectral bands, vegetation indices, and textural features, that are essential for accurately predicting crop parameters.

7. Conclusions

This paper sought to provide a comprehensive review of recent advances in remotely sensed retrieval of biochemical and biophysical parameters of crops brought by the developments in sensor technologies and novel machine learning retrieval techniques. Moreover, sources of uncertainty in retrieving crop parameters were identified, and practical implications for precision farming were discussed. Overall, the review revealed that developments in MLRA crop parameter retrieval techniques were mainly driven by announcements and the availability of new sensors, with the availability of the Sentinel-2 and SuperDoves constellations being ground-breaking events. Many studies were conducted with simulated data and airborne hyperspectral sensors at specific study areas with time series of field data covering many crops. Unfortunately, such permanent experimental sites are missing in sub-Saharan Africa, and there need to be coordinated and systematic efforts targeting calibration and validation data collection. Comparatively, well-coordinated campaigns are exemplary in other regions, such as Europe, where several campaigns exist, such as SPectra bARrax Campaign (SPARC, Spain, 2003 – 2004), SENtinel-2 and FLuorescence EXperiment (SEN2FLEX, Spain, 2005) and SEN3EXP (Spain, 2009).

Moreover, other prominent field campaigns included AgriSAR (Germany, 2006), and CarboEurope/FLEx/Sentinel-2 (CEFLES2, France, 2007). Although some models were tested across many sites, in most cases, such places were in Mediterranean climates. Hence, such models may not be portable to different climates, such as semi-arid African agricultural areas (Kganyago et al., 2020) and temperate continental semi-humid monsoon climates in China (Xie et al., 2019). Another consideration is that farm sizes in developed countries such as the USA, Spain, France and Italy are significantly larger than in sub-Saharan Africa. Therefore, crop-specific models may not be relevant in regions where farm configurations are characterised by small field sizes (i.e., $<$0.5 Ha) and mixed cropping systems. In addition to addressing gaps identified here, future research should focus on the development of generic (multi-crop), scalable (multi-sensor) and transferable models (across many sites and growing stages), especially within under-studied sub-Saharan African areas.

Funding details

This work was supported by the University of Witwatersrand and University of Johannesburg URC Grant [2023URC00563].

CRediT authorship contribution statement

Mahlatse Kganyago: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Clement Adjorlolo:** Supervision, Writing – review & editing. **Paidamwoyo Mhangara:** Supervision, Writing – review & editing. **Lesiba Tsoeleng:** Data curation, Investigation.

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

No data was used for the research described in the article.

Acknowledgements

We appreciate the participation of anonymous reviewers in the peer-review process and the editorial team. Mahlatse Kganyago received funding from University of Witwatersrand and University of Johannesburg (UJ) University Research Council Grant (URC, [2023URC00563]).

References

Abdulridha, J., Ampatzidis, Y., Kakarla, S.C., Roberts, P., 2019. Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques. Precis. Agric. https://doi.org/10.1007/s11119-019-09703-4.

Albughdadi, M., Rieu, G., Duthoit, S., Alswaitti, M., 2021. Towards a massive sentinel-2 LAI time-series production using 2-D convolutional networks. Comput. Electron. Agric. 180 https://doi.org/10.1016/j.compag.2020.105899.

Ali, A., Imran, M., 2021. Remotely sensed real-time quantification of biophysical and biochemical traits of Citrus (Citrus sinensis L.) fruit orchards – A review. Sci. Hortic. https://doi.org/10.1016/j.scienta.2021.110024.

Amin, E., Verrelst, J., Rivera-Caicedo, J.P., Pipia, L., Ruiz-Verdú, A., Moreno, J., 2021. Prototyping Sentinel-2 green LAI and brown LAI products for cropland monitoring. Remote Sens. Environ. 255 https://doi.org/10.1016/j.rse.2020.112168.

Apolo-Apolo, O.E., Pérez-Ruiz, M., Martínez-Guanter, J., Egea, G., 2020. A mixed data-based deep neural network to estimate leaf area index in wheat breeding trials. Agronomy 10. https://doi.org/10.3390/agronomy10020175.

Atzberger, C., 2000. Development of an invertible forest reflectance model The INFORM-Model, in: A Decade of Trans-European Remote Sensing Cooperation. Proceedings of the 20th EARSeL Symposium.

Atzberger, C., 2013. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. Remote Sens. (Basel) 5, 949–981. https://doi.org/10.3390/rs5020949.

Atzberger, C., Richter, K., 2012. Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. Remote Sens. Environ. 120, 208–218. https://doi.org/10.1016/j.rse.2011.10.035.

Azodi, C.B., Tang, J., Shiu, S.H., 2020. Opening the Black Box: Interpretable Machine Learning for Geneticists. Trends Genet. 36, 442–455. https://doi.org/10.1016/j.tig.2020.03.005.

Barbosa, A., Trevisan, R., Hovakimyan, N., Martin, N.F., 2020. Modeling yield response to crop management using convolutional neural networks. Comput. Electron. Agric. 170, 105197 https://doi.org/10.1016/j.compag.2019.105197.

Baret, F., Weiss, M., 2018. Gio Global Land Component - Lot I "Operation of the Global Land Component" Algorithm Theoretical Basis Document.

Baret, F., Houlès, V., Guérif, M., 2007. Quantification of plant stress using remote sensing observations and crop models: The case of nitrogen management. J. Exp. Bot. 58 https://doi.org/10.1093/jxb/erl231.

Bellvert, J., Mata, M., Vallverdú, X., Paris, C., Marsal, J., 2020. Optimizing precision irrigation of a vineyard to improve water use efficiency and profitability by using a decision-oriented vine water consumption model. Precis. Agric. https://doi.org/10.1007/s11119-020-09718-2.

Beltran, J.C., Valdez, P., Naval, P., 2019. Predicting Protein-Protein Interactions based on Biological Information using Extreme Gradient Boosting. 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2019. https://doi.org/10.1109/CIBCB.2019.8791241.

Boegh, E., Houborg, R., Bienkowski, J., Braban, C.F., Dalgaard, T., van Dijk, N., Dragosits, U., Holmes, E., Magliulo, V., Schelde, K., di Tommasi, P., Vitale, L., Theobald, M.R., Cellier, P., Sutton, M.A., 2013. Remote sensing of LAI, chlorophyll

and leaf nitrogen pools of crop- and grasslands in five European landscapes. Biogeosciences 10, 6279–6307. https://doi.org/10.5194/bg-10-6279-2013.
Campos-Taberner, M., García-Haro, F.J., Busetto, L., Ranghetti, L., Martínez, B., Gilabert, M.A., Camps-Valls, G., Camacho, F., Boschetti, M., 2018. A critical comparison of remote sensing Leaf Area Index estimates over rice-cultivated areas: From Sentinel-2 and Landsat-7/8 to MODIS, GEOV1 and EUMETSAT polar system. Remote Sens. (Basel) 10. https://doi.org/10.3390/rs10050763.
Carter, C., Liang, S., 2019. Evaluation of ten machine learning methods for estimating terrestrial evapotranspiration from remote sensing. Int. J. Appl. Earth Obs. Geoinf. 78, 86–92. https://doi.org/10.1016/j.jag.2019.01.020.
Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 13-17-Augu, 785–794. https://doi.org/10.1145/2939672.2939785.
Chen, Y., Feng, L., Mo, J., Mo, W., Ding, M., Liu, Z., 2020. Identification of Sugarcane with NDVI Time Series Based on HJ-1 CCD and MODIS Fusion. J. Indian Soc. Remote Sens. 48 https://doi.org/10.1007/s12524-019-01042-1.
Ciganda, V., Gitelson, A., Schepers, J., 2008. Vertical profile and temporal variation of chlorophyll in maize canopy: Quantitative "crop vigor" indicator by means of reflectance-based techniques. Agron. J. 100 https://doi.org/10.2134/agronj2007.0322.
Clevers, J.G.P.W., Gitelson, A.A., 2013. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and-3. Int. J. Appl. Earth Obs. Geoinf. 23, 344–351. https://doi.org/10.1016/j.jag.2012.10.008.
Corti, M., Cavalli, D., Cabassi, G., Vigoni, A., Degano, L., Marino Gallina, P., 2019. Application of a low-cost camera on a UAV to estimate maize nitrogen-related variables. Precis. Agric. 20, 675–696. https://doi.org/10.1007/s11119-018-9609-y.
Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., 2008. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. Remote Sens. Environ. 112 https://doi.org/10.1016/j.rse.2007.12.003.
Dawson, T.P., Curran, P.J., Plummer, S.E., 1998. LIBERTY - Modeling the effects of Leaf Biochemical Concentration on Reflectance Spectra. Remote Sens. Environ. 65 https://doi.org/10.1016/S0034-4257(98)00007-8.
De Castro, A.-I., Jurado-Exposito, M., Gómez-Casero, M.-T., Lopez-Granados, F., 2012. Applying neural networks to hyperspectral and multispectral field data for discrimination of cruciferous weeds in winter crops. The Scientific World Journal 2012.
de Keukelaere, L., Sterckx, S., Adriaensen, S., Knaeps, E., Reusen, I., Giardino, C., Bresciani, M., Hunter, P., Neil, C., van der Zande, D., Vaiciute, D., 2018. Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: validation for coastal and inland waters. Eur J Remote Sens 51, 525–542. https://doi.org/10.1080/22797254.2018.1457937.
Delegido, J., Verrelst, J., Alonso, L., Moreno, J., 2011. Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. Sensors 11, 7063–7081. https://doi.org/10.3390/s110707063.
Delegido, J., Van Wittenberghe, S., Verrelst, J., Ortiz, V., Veroustraete, F., Valcke, R., Samson, R., Rivera, J.P., Tenjo, C., Moreno, J., 2014. Chlorophyll content mapping of urban vegetation in the city of Valencia based on the hyperspectral NAOC index. Ecol. Ind. 40, 34–42. https://doi.org/10.1016/j.ecolind.2014.01.002.
Delegido, J., Verrelst, J., Rivera, J.P., Ruiz-Verdú, A., Moreno, J., 2015. Brown and green LAI mapping through spectral indices. Int. J. Appl. Earth Obs. Geoinf. 35, 350–358. https://doi.org/10.1016/j.jag.2014.10.001.
Delloye, C., Weiss, M., Defourny, P., 2018. Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. Remote Sens. Environ. 216, 245–261. https://doi.org/10.1016/j.rse.2018.06.037.
Dhakar, R., Sehgal, V.K., Chakraborty, D., Sahoo, R.N., Mukherjee, J., 2019. Field scale wheat LAI retrieval from multispectral Sentinel 2A-MSI and LandSat 8-OLI imagery: effect of atmospheric correction, image resolutions and inversion techniques. Geocarto Int. 1–21. https://doi.org/10.1080/10106049.2019.1687591.
Dong, T., Liu, J., Shang, J., Qian, B., Ma, B., Kovacs, J.M., Walters, D., Jiao, X., Geng, X., Shi, Y., 2019. Assessment of red-edge vegetation indices for crop leaf area index estimation. Remote Sens. Environ. 222, 133–143. https://doi.org/10.1016/j.rse.2018.12.032.
Dorigo, W., Lucieer, A., Podobnikar, T., Carni, A., 2012. Mapping invasive Fallopia japonica by combined spectral, spatial, and temporal analysis of digital orthophotos. Int. J. Appl. Earth Obs. Geoinf. 19, 185–195. https://doi.org/10.1016/j.jag.2012.05.004.
Doxani, G., Vermote, E., Roger, J.C., Gascon, F., Adriaensen, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G., Li, F., Louis, J., Mangin, A., Pahlevan, N., Pflug, B., Vanhellemont, Q., 2018. Atmospheric Correction Inter-Comparison Exercise. Remote Sens (basel) 10, 1–18. https://doi.org/10.3390/rs10020352.
Duan, S.B., Li, Z.L., Wu, H., Tang, B.H., Ma, L., Zhao, E., Li, C., 2014. Inversion of the PROSAIL model to estimate leaf area index of maize, potato, and sunflower fields from unmanned aerial vehicle hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 26, 12–20. https://doi.org/10.1016/j.jag.2013.05.007.
Elarab, M., Ticlavilca, A.M., Torres-Rua, A.F., Maslova, I., McKee, M., 2015. Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. Int. J. Appl. Earth Obs. Geoinf. 43, 32–42. https://doi.org/10.1016/j.jag.2015.03.017.
Fang, H., Wei, S., Liang, S., 2012. Validation of MODIS and CYCLOPES LAI products using global field measurement data. Remote Sens. Environ. 119, 43–54. https://doi.org/10.1016/j.rse.2011.12.006.
Fang, H., Zhang, Y., Wei, S., Li, W., Ye, Y., Sun, T., Liu, W., 2019. Validation of global moderate resolution leaf area index (LAI) products over croplands in northeastern Chinas. Remote Sens. Environ. 233, 111377 https://doi.org/10.1016/j.rse.2019.111377.
Féret, J.B., Gitelson, A.A., Noble, S.D., Jacquemoud, S., 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sens. Environ. 193, 204–215. https://doi.org/10.1016/j.rse.2017.03.004.
Fernandes, R., Weiss, M., Camacho, F., Berthelot, B., Baret, F., Duca, R., 2014. Development and assessment of leaf area index algorithms for the Sentinel-2 multispectral imager. International Geoscience and Remote Sensing Symposium (IGARSS) 3922–3925. https://doi.org/10.1109/IGARSS.2014.6947342.
Frampton, W.J., Dash, J., Watmough, G., Milton, E.J., 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS J. Photogramm. Remote Sens. 82, 83–92. https://doi.org/10.1016/j.isprsjprs.2013.04.007.
Gitelson, A.A., Gritz, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J. Plant Physiol. 160, 271–282. https://doi.org/10.1078/0176-1617-00887.
Gitelson, A.A., Viña, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll content in crops. Geophys. Res. Lett. 32, 1–4. https://doi.org/10.1029/2005GL022688.
Gitelson, A.A., Peng, Y., Huemmrich, K.F., 2014. Relationship between fraction of radiation absorbed by photosynthesizing maize and soybean canopies and NDVI from remotely sensed data taken at close range and from MODIS 250m resolution data. Remote Sens. Environ. 147, 108–120. https://doi.org/10.1016/j.rse.2014.02.014.
Giuliani, G., Chatenoux, B., de Bono, A., Rodila, D., Richard, J.P., Allenbach, K., Dao, H., Peduzzi, P., 2017. Building an Earth Observations Data Cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). Big Earth Data 1, 100–117. https://doi.org/10.1080/20964471.2017.1398903.
Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. Remote Sens. Environ. 70, 29–51.
Guindin-Garcia, N., Gitelson, A.A., Arkebauer, T.J., Shanahan, J., Weiss, A., 2012. An evaluation of MODIS 8- and 16-day composite products for monitoring maize green leaf area index. Agric. For. Meteorol. 161, 15–25. https://doi.org/10.1016/j.agrformet.2012.03.012.
Hank, T.B., Berger, K., Bach, H., Clevers, J.G.P.W., Gitelson, A., Zarco-Tejada, P., Mauser, W., 2019. Spaceborne imaging spectroscopy for sustainable agriculture: Contributions and challenges. Surv. Geophys. 40, 515–551.
Herrmann, I., Pimstein, A., Karnieli, A., Cohen, Y., Alchanatis, V., Bonfil, D.J., 2011. LAI assessment of wheat and potato crops by VENμS and Sentinel-2 bands. Remote Sens. Environ. 115, 2141–2151. https://doi.org/10.1016/j.rse.2011.04.018.
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.
Hunt, E.R., Horneck, D.A., Spinelli, C.B., Turner, R.W., Bruce, A.E., Gadler, D.J., Brungardt, J.J., Hamm, P.B., 2018. Monitoring nitrogen status of potatoes using small unmanned aerial vehicles. Precis. Agric. 19, 314–333.
Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sens. Environ. 34 https://doi.org/10.1016/0034-4257(90)90100-Z.
Jacquemoud, S., Baret, F., Andrieu, B., Danson, F.M., Jaggard, K., 1995. Extraction of vegetation biophysical parameters by inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. Remote Sens Environ 52. https://doi.org/10.1016/0034-4257(95)00018-V.
Jay, S., Maupas, F., Bendoula, R., Gorretta, N., 2017. Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. Field Crops Res 210, 33–46. https://doi.org/10.1016/j.fcr.2017.05.005.
Jia, F., Liu, G., Liu, D., Zhang, Y., Fan, W., Xing, X., 2013. Comparison of different methods for estimating nitrogen concentration in flue-cured tobacco leaves based on hyperspectral reflectance. Field Crops Res 150, 108–114. https://doi.org/10.1016/j.fcr.2013.06.009.
Kayad, A., Sozzi, M., Paraforos, D.S., Rodrigues, F.A., Cohen, Y., Fountas, S., Francisco, M.J., Pezzuolo, A., Grigolato, S., Marinello, F., 2022. How many gigabytes per hectare are available in the digital agriculture era? A digitization footprint estimation. Comput. Electron. Agric. 198 https://doi.org/10.1016/j.compag.2022.107080.
Kganyago, M., Adjorlolo, C., Mhangara, P., 2022. Exploring Transferable Techniques to Retrieve Crop Biophysical and Biochemical Variables Using Sentinel-2 Data. Remote Sens. (Basel) 14. https://doi.org/10.3390/rs14163968.
Kganyago, M., Mhangara, P., Alexandridis, T., Laneve, G., Ovakoglou, G., Mashiyi, N., 2020. Validation of sentinel-2 leaf area index (LAI) product derived from SNAP toolbox and its comparison with global LAI products in an African semi-arid agricultural landscape. Remote Sensing Letters 11, 883–892. https://doi.org/10.1080/2150704X.2020.1767823.
Kganyago, M., Ovakoglou, G., Mhangara, P., Adjorlolo, C., Alexandridis, T., Laneve, G., Beltran, J.S., 2023. Evaluating the contribution of Sentinel-2 view and illumination geometry to the accuracy of retrieving essential crop parameters. Gisci Remote Sens 60. https://doi.org/10.1080/15481603.2022.2163046.
Kimm, H., Guan, K., Jiang, C., Peng, B., Gentry, L.F., Wilkin, S.C., Wang, S., Cai, Y., Bernacchi, C.J., Peng, J., 2020. Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the US Corn Belt using Planet Labs CubeSat and STAIR fusion data. Remote Sens. Environ. 239, 111615.
Laurent, V.C.E., Schaepman, M.E., Verhoef, W., Weyermann, J., Chávez, R.O., 2014. Bayesian object-based estimation of LAI and chlorophyll from a simulated Sentinel-2

top-of-atmosphere radiance image. Remote Sens. Environ. 140, 318–329. https://doi.org/10.1016/j.rse.2013.09.005.
LI, H., CHEN, Z. xin, JIANG, Z. wei, WU, W. bin, REN, J. qiang, LIU, B., Tuya, H., 2017. Comparative analysis of GF-1, HJ-1, and Landsat-8 data for estimating the leaf area index of winter wheat. J Integr Agric 16, 266–285. https://doi.org/10.1016/S2095-3119(15)61293-X.
LI, Z. wang, XIN, X. ping, TANG, H., YANG, F., CHEN, B. rui, ZHANG, B. hui, 2017a. Estimating grassland LAI using the Random Forests approach and Landsat imagery in the meadow steppe of Hulunber, China. J Integr Agric 16, 286–297. https://doi.org/10.1016/S2095-3119(15)61303-X.
Li, X., Liu, X., Liu, M., Wang, C., Xia, X., 2015. A hyperspectral index sensitive to subtle changes in the canopy chlorophyll content under arsenic stress. Int. J. Appl. Earth Obs. Geoinf. 36, 41–53. https://doi.org/10.1016/j.jag.2014.10.017.
Liang, L., Di, L., Zhang, L., Deng, M., Qin, Z., Zhao, S., Lin, H., 2015. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. Remote Sens. Environ. 165, 123–134. https://doi.org/10.1016/j.rse.2015.04.032.
Liao, C., Wang, J., Dong, T., Shang, J., Liu, J., Song, Y., 2019. Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean. Sci. Total Environ. 650, 1707–1721. https://doi.org/10.1016/j.scitotenv.2018.09.308.
López-Granados, F., Torres-Sánchez, J., Serrano-Pérez, A., de Castro, A.I., Mesas-Carrascosa, F.J., Peña, J.M., 2016. Early season weed mapping in sunflower using UAV technology: variability of herbicide treatment maps against weed thresholds. Precis. Agric. 17, 183–199. https://doi.org/10.1007/s11119-015-9415-8.
Marinello, F., Bramley, R.G.V., Cohen, Y., Fountas, S., Guo, H., Karkee, M., Martínez-Casasnovas, J.A., Paraforos, D.S., Sartori, L., Sørensen, C.G., Stenberg, B., Sudduth, K., Tisseyre, B., Vellidis, G., Vougioukas, S.G., 2019. Agriculture and digital sustainability: A digitization footprint, in: Precision Agriculture 2019 - Papers Presented at the 12th European Conference on Precision Agriculture, ECPA 2019. https://doi.org/10.3920/978-90-8686-888-9_9.
Mulla, D.J., 2013. Twenty five years of remote sensing in precision agriculture. Biosyst. Eng. 114.
Nasir, R., Khan, M.J., Arshad, M., Khurshid, K., 2019. Convolutional neural network based regression for leaf water content estimation, in: 2019 2nd International Conference on Latest Trends in Electrical Engineering and Computing Technologies, INTELLECT 2019. https://doi.org/10.1109/INTELLECT47034.2019.8954985.
Nigam, R., Bhattacharya, B.K., Vyas, S., Oza, M.P., 2014. Retrieval of wheat leaf area index from AWiFS multispectral datausing canopy radiative transfer simulation. Int. J. Appl. Earth Obs. Geoinf. 32, 173–185. https://doi.org/10.1016/j.jag.2014.04.003.
Qian, B., Ye, H., Huang, W., Xie, Q., Pan, Y., Xing, N., Ren, Y., Guo, A., Jiao, Q., Lan, Y., 2022. A sentinel-2-based triangular vegetation index for chlorophyll content estimation. Agric. For. Meteorol. 322, 109000.
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. https://doi.org/10.1016/j.jag.2014.12.010.
Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R., Heitkönig, I.M.A., 2012. Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor. Int. J. Appl. Earth Obs. Geoinf. 19 https://doi.org/10.1016/j.jag.2012.05.009.
Rasul, A., Ibrahim, S., Onojeghuo, A.R., Balzter, H., 2020. A trend analysis of leaf area index and land surface temperature and their relationship from global to local scale. Land (basel) 9, 1–17. https://doi.org/10.3390/land9100388.
Rautiainen, M., Heiskanen, J., Korhonen, L., 2012. Seasonal changes in canopy leaf area index and MODIS vegetation products for a boreal forest site in central Finland. Boreal Environ. Res. 17, 72–84.
Richter, K., Atzberger, C., Vuolo, F., Weihs, P., D'urso, G., 2009. Experimental assessment of the Sentinel-2 band setting for RTM-based LAI retrieval of sugar beet and maize. Canadian Journal of Remote Sensing 35, 230–247. https://doi.org/10.5589/m09-010.
Richter, K., Hank, T.B., Vuolo, F., Mauser, W., D'Urso, G., 2012. Optimal exploitation of the sentinel-2 spectral capabilities for crop leaf area index mapping. Remote Sens. (Basel) 4. https://doi.org/10.3390/rs4030561.
Rivera-Caicedo, J.P., Verrelst, J., Muñoz-Marí, J., Camps-Valls, G., Moreno, J., 2017. Hyperspectral dimensionality reduction for biophysical variable statistical retrieval. ISPRS J. Photogramm. Remote Sens. 132, 88–101. https://doi.org/10.1016/j.isprsjprs.2017.08.012.
Shah, S.H., Angel, Y., Houborg, R., Ali, S., McCabe, M.F., 2019. A Random Forest Machine Learning Approach for the Retrieval of Leaf Chlorophyll Content in Wheat. Remote Sens. (Basel) 11, 920. https://doi.org/10.3390/rs11080920.
Shelestov, A., Kolotii, A., Skakun, S., Baruth, B., Lozano, R.L., Yailymov, B., 2017. Biophysical parameters mapping within the SPOT-5 take 5 initiative. Eur J Remote Sens 50, 300–309. https://doi.org/10.1080/22797254.2017.1324743.
Si, Y., Schlerf, M., Zurita-Milla, R., Skidmore, A., Wang, T., 2012. Mapping spatio-temporal variation of grassland quantity and quality using MERIS data and the PROSAIL model. Remote Sens. Environ. 121, 415–425. https://doi.org/10.1016/j.rse.2012.02.011.
Sola, I., García-Martín, A., Sandonís-Pozo, L., Álvarez-Mozos, J., Pérez-Cabello, F., González-Audícana, M., Montorio Llovería, R., 2018. Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes. Int. J. Appl. Earth Obs. Geoinf. 73, 63–76. https://doi.org/10.1016/j.jag.2018.05.020.
Sun, C., Sun, Z., Liu, T., Guo, D., Mu, S., Yang, H., Ju, W., Li, J., 2013. The validation of a model estimating the leaf area index of grasslands in southern China. Rangel. J. 35, 245–250. https://doi.org/10.1071/RJ12025.
Tian, Y.C., Gu, K.J., Chu, X., Yao, X., Cao, W.X., Zhu, Y., 2013. Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. Plant and Soil 376, 193–209. https://doi.org/10.1007/s11104-013-1937-0.
Uddling, J., Gelang-Alfredsson, J., Piikki, K., Pleijel, H., 2007. Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. Photosynth. Res. 91, 37–46. https://doi.org/10.1007/s11120-006-9077-5.
Van Der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., Su, Z., 2009. An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. Biogeosciences 6. https://doi.org/10.5194/bg-6-3109-2009.
Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. Remote Sens. Environ. 16 https://doi.org/10.1016/0034-4257(84)90057-9.
Verhoef, W., Bach, H., 2007. Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. Remote Sens. Environ. 109 https://doi.org/10.1016/j.rse.2006.12.013.
Verhoef, W., Jia, L., Xiao, Q., Su, Z., 2007. Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies. IEEE Trans. Geosci. Remote Sens. 45 https://doi.org/10.1109/TGRS.2007.895844.
Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J., Moreno, J., 2012a. Retrieval of vegetation biophysical parameters using Gaussian process techniques. IEEE Trans. Geosci. Remote Sens. 50, 1832–1843. https://doi.org/10.1109/TGRS.2011.2168962.
Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G., Moreno, J., 2012b. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. Remote Sens. Environ. 118, 127–139. https://doi.org/10.1016/j.rse.2011.11.002.
Verrelst, J., Rivera, J.P., Alonso, L., Guanter, L., Moreno, J., 2012c. Evaluating machine learning regression algorithms for operational retrieval of biophysical parameters: Opportunities for Sentinel. European Space Agency. (Special Publication) ESA SP.
Verrelst, J., Romijn, E., Kooistra, L., 2012d. Mapping vegetation density in a heterogeneous river floodplain ecosystem using pointable CHRIS/PROBA data. Remote Sens. (Basel) 4, 2866–2889. https://doi.org/10.3390/rs4092866.
Verrelst, J., Rivera, J.P., Leonenko, G., Alonso, L., Moreno, J., 2013a. Optimizing LUT-based RTM inversion for semiautomatic mapping of crop biophysical parameters from Sentinel-2 and-3 data: Role of cost functions. IEEE Trans. Geosci. Remote Sens. 52, 257–269.
Verrelst, J., Rivera, J.P., Moreno, J., Camps-Valls, G., 2013b. Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval. ISPRS J. Photogramm. Remote Sens. 86, 157–167. https://doi.org/10.1016/j.isprsjprs.2013.09.012.
Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Marí, J., Clevers, J.G.P.W., Camps-Valls, G., Moreno, J., 2015. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods - A comparison. ISPRS J. Photogramm. Remote Sens. 108, 260–272. https://doi.org/10.1016/j.isprsjprs.2015.04.013.
Verrelst, J., Rivera, J.P., Gitelson, A., Delegido, J., Moreno, J., Camps-Valls, G., 2016. Spectral band selection for vegetation properties retrieval using Gaussian processes regression. Int. J. Appl. Earth Obs. Geoinf. 52, 554–567. https://doi.org/10.1016/j.jag.2016.07.016.
Vilfan, N., Tol, C. Van Der, Yang, P., Wyber, R., Robinson, S.A., Verhoef, W., 2018. Remote Sensing of Environment Extending Fluspect to simulate xanthophyll driven leaf re fl ectance dynamics 211, 345–356. https://doi.org/10.1016/j.rse.2018.04.012.
Vilfan, N., van der Tol, C., Muller, O., Rascher, U., Verhoef, W., 2016. Fluspect-B: A model for leaf fluorescence, reflectance and transmittance spectra. Remote Sens. Environ. 186 https://doi.org/10.1016/j.rse.2016.09.017.
Viña, A., Gitelson, A.A., Nguy-Robertson, A.L.A.L., Peng, Y., 2011. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. Remote Sens. Environ. 115, 3468–3478. https://doi.org/10.1016/j.rse.2011.08.010.
Vincini, M., Calegari, F., Casa, R., 2016. Sensitivity of leaf chlorophyll empirical estimators obtained at Sentinel-2 spectral resolution for different canopy structures. Precis. Agric. 17, 313–331. https://doi.org/10.1007/s11119-015-9424-7.
Wan, L., Cen, H., Zhu, J., Zhang, J., Zhu, Y., Sun, D., Du, X., Zhai, L., Weng, H., Li, Y., 2020. Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer–a case study of small farmlands in the South of China. Agric. For. Meteorol. 291, 108096.
Wang, S., Gao, W., Ming, J., Li, L., Xu, D., Liu, S., Lu, J., 2018. A TPE based inversion of PROSAIL for estimating canopy biophysical and biochemical variables of oilseed rape. Comput. Electron. Agric. 152 https://doi.org/10.1016/j.compag.2018.07.023.
Weiss, M., Baret, F., 2016. S2ToolBox Level 2 products: LAI, FAPAR, FCOVER - Version 1.1. Sentinel2 ToolBox Level2 Products 53.
Weiss, M., Baret, F., Smith, G.J., Jonckheere, I., Coppin, P., 2004. Review of methods for in situ leaf area index (LAI) determination Part II. Estimation of LAI, errors and sampling. Agric. For. Meteorol. 121, 37–53. https://doi.org/10.1016/j.agrformet.2003.08.001.
Wen, D., Tongyu, X., Fenghua, Y., Chunling, C., 2018. Measurement of nitrogen content in rice by inversion of hyperspectral reflectance data from an unmanned aerial vehicle. Cienc. Rural 48, 1–8. https://doi.org/10.1590/0103-8478cr20180008.
Wu, C., Liu, M., Liu, X., Wang, T., Wang, L., 2019. Developing a new spectral index for detecting cadmium-induced stress in rice on a regional scale. Int. J. Environ. Res. Public Health 16. https://doi.org/10.3390/ijerph16234811.
Wu, M., Wu, C., Huang, W., Niu, Z., Wang, C., 2015. High-resolution Leaf Area Index estimation from synthetic Landsat data generated by a spatial and temporal data

fusion model. Comput. Electron. Agric. 115, 1–11. https://doi.org/10.1016/j.compag.2015.05.003.

Wu, J., Zhang, J., Lü, A., Zhou, L., 2012. An exploratory analysis of spectral indices to estimate vegetation water content using sensitivity function. Remote Sensing Letters 3, 161–169. https://doi.org/10.1080/01431161.2011.551845.

Xiaoyan, W., Zhiwei, L., Wenjun, W., Jiawei, W., 2020. Chlorophyll content for millet leaf using hyperspectral imaging and an attention-convolutional neural network. Cienc. Rural 50. https://doi.org/10.1590/0103-8478cr20190731.

Xie, Q., Dash, J., Huang, W., Peng, D., Qin, Q., Mortimer, H., Casa, R., Pignatti, S., Laneve, G., Pascucci, S., 2018. Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11, 1482–1493.

Xie, Q., Dash, J., Huete, A., Jiang, A., Yin, G., Ding, Y., Peng, D., Hall, C.C., Brown, L., Shi, Y., Ye, H., Dong, Y., Huang, W., 2019. Retrieval of crop biophysical parameters from Sentinel-2 remote sensing imagery. Int. J. Appl. Earth Obs. Geoinf. 80 https://doi.org/10.1016/j.jag.2019.04.019.

Yan, G., Hu, R., Luo, J., Weiss, M., Jiang, H., Mu, X., Xie, D., Zhang, W., Ronghai, H., Jinghui, L., Weiss, M., Hu, R., Luo, J., Weiss, M., Jiang, H., Mu, X., Xie, D., Zhang, W., 2019. Review of indirect optical measurements of leaf area index: Recent advances, challenges, and perspectives. Agric. For. Meteorol. 265, 390–411. https://doi.org/10.1016/j.agrformet.2018.11.033.

Yang, F., Sun, J., Fang, H., Yao, Z., Zhang, J., Zhu, Y., Song, K., Wang, Z., Hu, M., 2012. Comparison of different methods for corn LAI estimation over northeastern China. Int. J. Appl. Earth Obs. Geoinf. 18, 462–471. https://doi.org/10.1016/j.jag.2011.09.004.

Yuan, H., Yang, G., Li, C., Wang, Y., Liu, J., Yu, H., Feng, H., Xu, B., Zhao, X., Yang, X., 2017. Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. Remote Sens. (Basel) 9. https://doi.org/10.3390/rs9040309.

Zaroug, M.A.H., Sylla, M.B., Giorgi, F., Eltahir, E.A.B., Aggarwal, P.K., 2013. A sensitivity study on the role of the swamps of southern Sudan in the summer climate of North Africa using a regional climate model. Theor. Appl. Climatol. 113 https://doi.org/10.1007/s00704-012-0751-6.

Zhang, M., Su, W., Fu, Y., Zhu, D., Xue, J.H., Huang, J., Wang, W., Wu, J., Yao, C., 2019. Super-resolution enhancement of Sentinel-2 image for retrieving LAI and chlorophyll content of summer corn. Eur. J. Agron. 111, 125938 https://doi.org/10.1016/j.eja.2019.125938.

Zhang, M., Yuan, X., Otkin, J.A., 2020. Remote sensing of the impact of flash drought events on terrestrial carbon dynamics over China. Carbon Balance Manag. 15, 1–11. https://doi.org/10.1186/s13021-020-00156-1.

Zheng, H., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y., Cao, W., Zhu, Y., 2019. Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. Precis. Agric. 20, 611–629. https://doi.org/10.1007/s11119-018-9600-7.

Zheng, Y., Wu, B., Zhang, M., Zeng, H., 2016. Crop Phenology Detection Using High Spatio-Temporal Resolution Data Fused from SPOT5 and MODIS Products. Sensors 16, 2099. https://doi.org/10.3390/s16122099.

Zhou, J., Khot, L.R., Boydston, R.A., Miklas, P.N., Porter, L., 2018. Low altitude remote sensing technologies for crop stress monitoring: a case study on spatial and temporal monitoring of irrigated pinto bean. Precis. Agric. 19, 555–569. https://doi.org/10.1007/s11119-017-9539-0.

Zhou, Z., Plauborg, F., Kristensen, K., Andersen, M.N., 2017. Dry matter production, radiation interception and radiation use efficiency of potato in response to temperature and nitrogen application regimes. Agric. For. Meteorol. 232, 595–605. https://doi.org/10.1016/j.agrformet.2016.10.017.

Zhou, X., Wang, P., Tansey, K., Zhang, S., Li, H., Tian, H., 2020. Reconstruction of time series leaf area index for improving wheat yield estimates at field scales by fusion of Sentinel-2, -3 and MODIS imagery. Comput. Electron. Agric. 177, 105692 https://doi.org/10.1016/j.compag.2020.105692.

Zisi, T., Alexandridis, T.K., Kaplanis, S., Navrozidis, I., Tamouridou, A.A., Lagopodi, A., Moshou, D., Polychronos, V., 2018. Incorporating surface elevation information in UAV multispectral images for mapping weed patches. J Imaging 4. https://doi.org/10.3390/jimaging4110132.